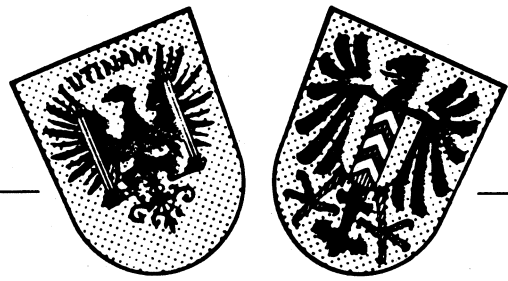


EFTF 1997

11th European Frequency and Time Forum

11ème Forum Européen Fréquence et Temps

4, 5, 6 and 7 March 1997

NEUCHÂTEL, SWITZERLAND

PROCEEDINGS

11th European Frequency and Time Forum

11ème Forum Européen Fréquence et Temps

Neuchâtel, Switzerland – March 4-7 1997

PROCEEDINGS

including proceedings of the Workshop « New Generation of Space Clocks »

Conference organizer and proceedings publisher :

FSRM

Swiss Foundation for Research in Microtechnology

Rue Jaquet-Droz 1 / Case postale 20

2007 Neuchâtel, Switzerland

Additional copies of these proceedings can be purchased from the

Permanent EFTF Secretariat

c/o FSRM

Swiss Foundation for Research in Microtechnology

Rue Jaquet-Droz 1 / Case postale 20

CH-2007 Neuchâtel, Switzerland

☎ +41 32 7 200 900 / Fax :+41 32 7 200 990

e-mail: eftf@fsrm.ch

http://www.fsrm.ch

TABLE OF CONTENTS

INTRODUCTION

✹ = ***Abstract only***

OPENING SESSION

RESONATORS I

CESIUM STANDARDS

TELECOMMUNICATIONS

OPTICAL FREQUENCY STANDARDS

SAW FILTERS

COLD CESIUM ATOMS

RESONATORS II

TIME AND FREQUENCY COMPARISONS

RUBIDIUM STANDARDS AND FLYWHEELS

MATERIALS

OSCILLATORS

TIME, SPACE, AND RELATIVITY

GPS APPLICATIONS

MEASUREMENT TECHNIQUES

POSTERS – PIEZOELECTRICITY

POSTERS
T/F TECHNIQUES AND ATOMIC STANDARDS

WORKSHOP

NEW GENERATION OF SPACE CLOCKS

SESSION 1

SESSION 2

ORGANIZATION 1997

EFTF Local Organization

Dr. Marcel Ecabert, FSRM, Neuchâtel

Carine Montandon, FSRM, Neuchâtel

Bernard Schlueter, Keres Engineering Ltd., Neuchâtel

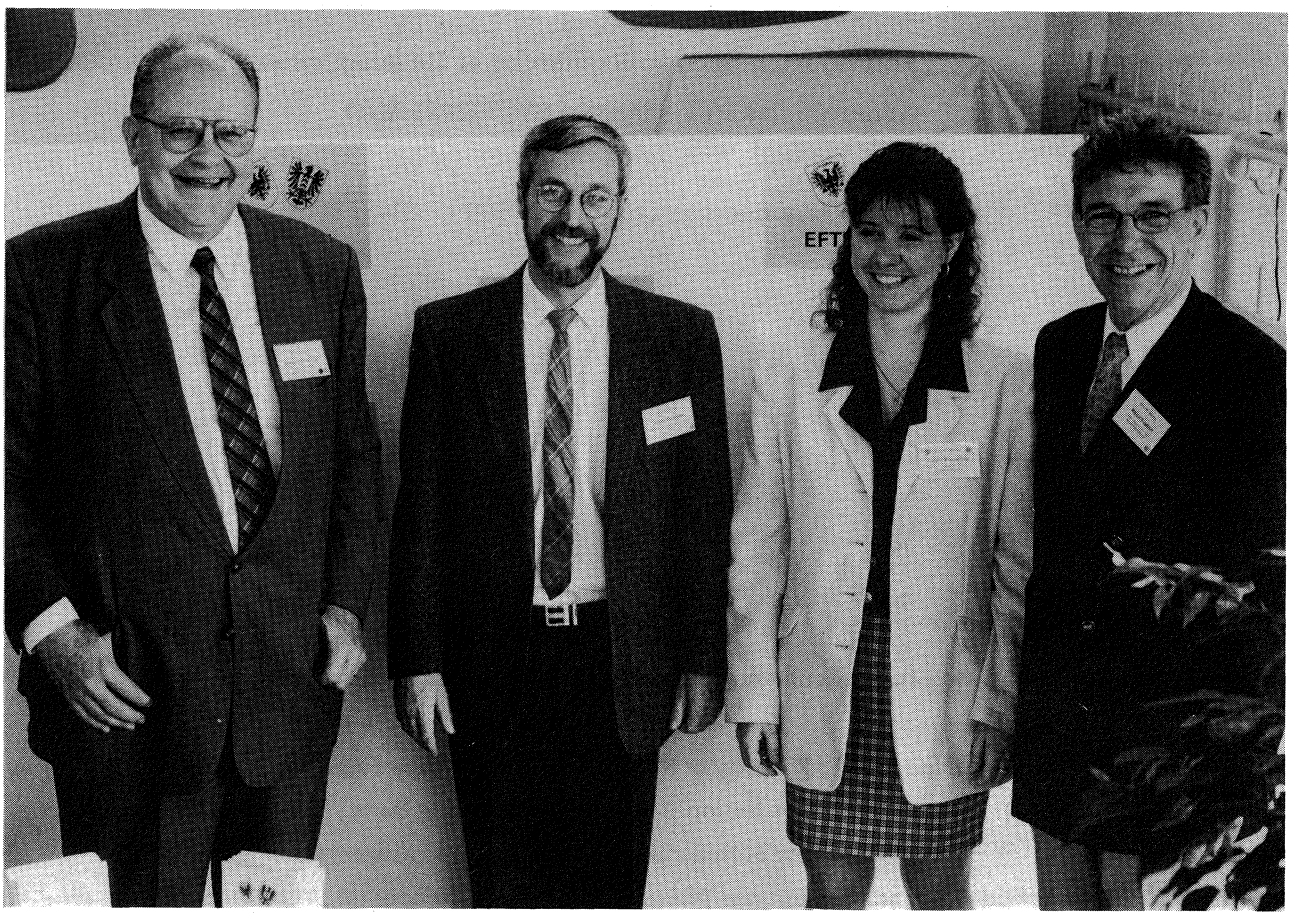
Technical Programme Chairman

Dr. Pierre Thomann, Observatoire cantonal, Neuchâtel

Workshop Organization

Dr. Philippe Laurent, BNM-LPTF, Paris - France

Dr. Pierre Thomann, Observatoire cantonal, Neuchâtel



From the left : B. Schlueter, P. Thomann, C. Montandon, M. Ecabert

EXECUTIVE COMMITTEE

Claude Audoin	Laboratoire de l'Horloge atomique, CNRS - Orsay - FRANCE
Raymond J. Besson	Laboratoire de Chronométrie, Electronique et Piézoélectricité de l'Ecole nationale Supérieure de Mécanique et des Microtechniques Besançon – FRANCE
Marcel Ecabert	Fondation Suisse pour la Recherche en Microtechnique, FSRM Neuchâtel – SWITZERLAND
Jean-Jacques Gagnepain	Département Sciences pour l'ingénieur, CNRS – Paris - FRANCE
Peter Kartaschoff	Retired from Section Recherche et Développement de la Direction Générale des PTT – Bern – SWITZERLAND
Bernard Schlueter	Keres Engineering Ltd – Neuchâtel - SWITZERLAND

Acknowledgements

We wish to thank the following institutions and companies for their sponsorship :

The Republic and Canton of Neuchâtel ; The City of Neuchâtel ; The Swiss Academy of Engineering Sciences (SATW / ASST), Zürich ; The Swiss National Science Foundation, Bern ; Le Bureau National de Métrologie, Paris – France ; The University of Neuchâtel ; L'imprimerie Krattiger, Corcelles, Neuchâtel ; PTT, Direction d'arrondissement des Télécommunications, Neuchâtel ; The Tourist Office, Neuchâtel ; The Observatory of Neuchâtel

SCIENTIFIC COMMITTEE EFTF 1997

Thomann P.	*Chairman 1997, Observatoire Cantonal, Neuchâtel*	*Switzerland*
Aubry J.-P.	Oscilloquartz SA, Neuchâtel	Switzerland
Audoin C.	LHA / CNRS, Université Paris Sud, Orsay	France
Ballato A.	US Army Research Laboratory, Fort Monmouth, N.J	U S A
Bauch A.	Physikalisch-Technische Bundesanstalt, Braunschweig	Germany
Besson R.	LCEP / ENSMM, Besançon	France
Boloix R.	Real Instituto y Observatorio de la Armada, San Fernando, Cadiz.	Spain
Brunet M.	C.N.E.S., Toulouse	France
Busca G.	Observatoire Cantonal, Neuchâtel	Switzerland
Cutler L.	Hewlett Packard Laboratories, Palo Alto, CA.	USA
De Jong G.	NMi Van Swinden Laboratory, Delft	The Netherlands
De Marchi A.	Politecnico di Torino, Torino	Italy
Defranould P.	Thomson-Microsonics, Sophia Antipolis	France
Deyzac F.	ONERA, Châtillon sous Bagneux	France
Dorenwendt K.	Physikalisch-Technische Bundesanstalt, Braunschweig	Germany
Dowsett J.	London	UK
Duchaussoy H.	D.R.E.T., Paris Armées	France
Eskelinen P.	Kotka Institute of Technology, Kotka	Finland
Feltham S.	ESA - European Space Agency ESTEC, Noordwijk	The Netherlands
Godone A.	IEN Galileo Ferraris, Torino	Italy
Graf E.	Bevaix	Switzerland
Granveaud M.	LPTF - Observatoire de Paris, Paris	France
Hartl P.	Universität Stuttgart, Stuttgart	Germany
Hauden D.	Université de Franche-Comté, Besançon	France
Heighway J.	Goulens	France
Hilty K.	Swiss Telecom PTT, Bern	Switzerland
Hruska C.	York University, Downsview, Ontario	Canada
Kalliomäki K.	University Oulu, Oulu	Finland
Kartaschoff P.	Cormondrèche	Switzerland
Kihara M.	NTT Optical Network Systems Laboratories, Yokosuka	Japan
Kirchner D.	Technische Universität Graz, Graz	Austria
Krempl P.	AVL List GmbH, Graz	Austria
Leschiutta S.	Polytecnico di Torino, Torino	Italy
Nakagiri K.	Kinki University, Nishi-Mitani	Japan
Nau H.	DLR, Wessling	Germany
Neubig B.	Tele Quarz GmbH, Neckarbischofsheim	Germany
Nieuwejaar P.	Norwegian Naval Materiel Command, Haakonsvern	Norway
Prost L.	Office Fédéral de Métrologie, Wabern	Switzerland
Quinn T.	B.I.P.M., Sèvres	France
Reid A.	British Telecom, London	UK
Schlüter W.	Institut für Angewandte Geodäsie, Kötzting	Germany
Steele J.	National Physical Laboratory, Teddington, Middlesex	UK
Uljanov A.	Kvarz Institute of Electronic Measurements, Nizhny Novgorod	Russia
Underhill M. J.	University of Surrey, Guildford	UK
Unverdross R.	Unverdross Technik, Wörthsee	Germany
Vanier J.	Ste-Foie, Québec	Canada
Walls F.	NIST, Boulder, CO	USA
Winkler G.	Innovative Solutions International, Vienna, Virginia	USA
Yannoni N.	Newton, MA	USA
Zampetti G.	Telecom Solutions, San José, CA	USA
Zingg W.	Microcrystal Div. of ETA S.A., Grenchen	Switzerland

OPENING SESSION

Chairman: Pierre Thomann

THE INTERNATIONAL SPACE STATION — A CHALLENGE FOR GLOBAL COOPERATION

Jörg Feustel-Büechl[1], Dieter Isakeit[2]

[1]Director of Manned Spaceflight and Microgravity
[2]Head of the Coordination Office, Directorate of Manned Spaceflight and Microgravity
[1][2]European Space Agency (ESA), Noordwijk, The Netherlands

ABSTRACT

The International Space Station is a global cooperation programme involving 14 countries. The paper presents the reasons for the Station, the political framework, the role of Europe and the expected utilisation benefits.

Keywords: International Space Station, European Space Agency, Utilisation

1. INTRODUCTION

Among all the SI units, the "Second" is today the unit which can be measured with the highest accuracy of all. Using the principle of atomic resonance with the element Cesium 133, the reference second can be determined with an accuracy which is by several orders of magnitude better than for example the reference meter. In principle, scientists and engineers involved in the measurement of time could therefore quietly lean back, waiting that their colleagues who are responsible for the measurement of the other SI units, such as mass, electric current or magnetic force, join the club. The fact that reality is quite the opposite and that scientists and engineers are further striving for achieving more and more precise clocks, shows that there is a real need for knowing the time with extreme precision. That is because time is often used as a reference for other measurements, like distance or velocity. Any improvement in the measurement and dissemination of time has therefore a direct impact on the precision of many other measurements.

When it comes to the improvement of time measurement and dissemination systems, space based systems can be both: a customer and a helper. A customer, because a more precise on-board clock on a spacecraft means a more precise determination of its position in space, or, like for the satellites involved in the Global Positioning System, it means the dissemination of a more precise position signal from space for air and ground vehicles on Earth.

But space systems can also be a helper: they can contribute to increasing the precision of atomic clocks, at least by one order of magnitude, as compared to atomic clocks on the ground. They can also be a platform for the dissemination of time reference signals all over the world.
On each of these aspects, there are many presentations given at this Eleventh European Frequency and Time Forum.

One particular space system which will offer new or improved possibilities for the measurement and dissemination of time is the International Space Station (Fig. 1). This Station will soon come into reality. A large number of presentations at this Forum deal with experiments or services proposed for the Station. It may therefore be interesting for the audience to know more about the reasons, nature and expected benefits of this Station.

2. REASONS FOR THE INTERNATIONAL SPACE STATION

One of the main challenges of the International Space Station is already reflected in the name itself. The Station is an international programme of cooperation that units five international partners from all aroud the world. They are the USA, Russia, Europe, Japan and Canada.

One of these five partners, namely Europe, is itself already an international partnership embodied in the European Space Agency (ESA). Out of the 14 Member States of ESA, ten are taking part in the ESA Programme for European Participation in the International Space Station. In alphabetical order, they are: Belgium, Denmark, France, Germany, Italy, the Netherlands, Norway, Spain, Sweden and Switzerland.

The assembly of the Station in space is planned to begin in late 1997/early 1998, and it should be ready in late 2002/early 2003. But even while building is going on, the scientific and technological use of the

Station will already begin one year after the launch of the first element into orbit. Full routine use will be under way by 2003, and will go on for at least ten years until 2013.

When the construction of the International Space Station will have been completed, the Station will comprise a number of pressurized modules in which an international crew of up to seven astronauts can live and work. Additionally, various external platforms will make it possible to set up observation and experimental devices outside the Station.

The Station will then provide opportunities that have never been matched in quality or quantity by any space system up to now. Its four major advantages are:

- availability over a long period,
- regularity of access,
- high performance,
- a globalised utilisation scheme.

The long operational life of the International Space Station is a political, technical and operational challenge and one of the Station's most outstanding assets. There has never before been such a continuous and pre-plannable useability. Europe in particular has been more or less reliant on sporadic Spacelab, Mir and Eureca missions, some of them years apart.

For Europe, the reasons to take part in the International Space Station lie primarily in three areas:

- making use of the new scientific and technical opportunities that the International Space Station offers,
- creating the technical and industrial capabilities needed in key technologies that will be important for Europe's role in the future; and
- fostering international cooperation.

3. UTILISATON OF THE INTERNATIONAL SPACE STATION

Spaceflight is not an end in itself. It is a means of solving problems on Earth from a higher level. It also helps us gain a better knowledge of the universe, and of our own place in it.

In the eyes of the public, the question whether spaceflight should be better conducted by manned or by unmanned systems is sometimes elevated to the level of an ideological dispute. Reality is, however, different. There is no clear border line between both means. They are in fact complementary to each other.

A manned space station is not art for art's sake: it is not only used to find out how to operate space stations themselves, and what effects living on such a station has on the astronauts. It is meant first of all for scientific research and for technical progress for people on Earth — by making skilled use of the special advantages offered by its position and its environment out in space.

Since Europe's participation in the International Space Station had definitively been approved by the ESA Member States at the Ministerial Council Meeting of Toulouse in October 1996, we are finding to our delight that there is a lot of interest in Europe in using the International Space Station. Early October 1996 we were holding in Darmstadt, at the ESA Operations Centre (ESOC), the first symposium in Europe on the utilisation of the Station. The response has been impressive. More than 400 scientists and engineers presented and discussed their own proposals and those of their colleagues for making use of the Station in the various user disciplines such as life sciences, physics, environmental sciences, Earth observation and technology [Ref. 1].

3.1 Platform outside the Earth's Atmosphere

One of the prime advantages of a space station's position is that it is outside the Earth's atmosphere. This envelope of air surrounding the Earth offers a substantial resistance to any transport from Earth out into space and back again to Earth. Overcoming this resistance calls for extraordinarily powerful and reliable means of transport; which means that the costs of space transport are the decisive cost factor when it comes to making use of space-based systems. As space is further opened up for scientific and technical applications, it is crucial that the overall costs be further reduced quite significantly.

The present-day space transportation means have however already to a great extent reached their limit. Any quantum leap, in boosting performance and dramatically bringing down costs, will so far as one can see be achieved only with fundamentally new space transport systems. But doing that is an extraordinarily long-term and high-cost affair.

So if we want to reduce the costs of space-based systems, we have not only to think of the costs of the transport, but also create and test ways of reducing the demands on the transport itself. Every good army general knows that he has to keep his supply lines as short and as uncomplicated as possible, and ideally manage as far as possible with the means available on the spot. It is just the same with space, and in the

long term we must learn, like a skilful general, to live off the land.

It would be going too far to say that the International Space Station will quickly become a sort of cross between an orbiting filling-station and a repair workshop. The various possibilities for carrying out servicing and provisioning in space directly and "on-site" first need to be developed and tested in practice. Only when that has been done will future space systems — both manned and unmanned — be even designed for being serviced and repaired in space - probably not yet by this Station, but maybe by its successor. But one cannot hang around doing nothing and waiting for it to happen.

There are voices saying that this will never be a cost-effective alternative to the present "use-once-and-throw-away" systems. But we have already seen so many prophets get it wrong in their predictions on a completely new technical potential. Who, for instance, would have believed it possible that the market for telecommunication satellites would have moved so fast in less than 25 years that the great majority of these satellites are now being built, launched and operated on a commercial basis, and that the European Ariane launcher alone would be lifting today a new telecommunications satellite into space once a month on average?

A further advantage the International Space Station will gain from its position outside the atmosphere is the opportunity to observe the Earth itself, with its atmosphere, as well as the rest of the universe.

The high inclination of the International Space Station's orbit, which makes it overfly 85% of the Earth's surface where 95% of the human population live, and the unusually low orbiting altitude of approximately 450 km, compared to conventional Earth-observation satellites which operate at twice that altitude, offer an interesting potential.

Because it is permanently ready for use and extremely flexible in operation, the Station is particularly well suited to observing unpredictable events on Earth such a natural disasters, forest fires and volcanic eruptions. Since the instruments used for Earth observation are very costly, the International Space Station facilitates the introduction of new equipment and procedures that one would otherwise not happily want to trust to an unmanned satellite that, once launched, is no longer so easy of access.

In the space sciences, there are firm ideas on using the International Space Station for solar, astronomical and astrophysical observations. There are also proposals for studying the cosmic radiation environment.

Because of its size and other capabilities, the International Space Station will also be suitable for helping detect and track Near-Earth Objects (NEO), like asteroids and comets. As it is available longterm, and its crew is constantly present, the Station can adapt at short notice to unexpected events in space. The comet impacts on Jupiter in 1994 have shown that a possible threat to Earth from collision with a foreign body in space is no mere abstract fantasy. True, such celestial bodies can be detected from Earth with optical instruments; but the presence of our atmosphere makes it hard to measure their course accurately.

The International Space Station is also valuable for less spectacular uses in the space sciences: it offers the possibility of installing instruments for high-energy astrophysics that demand a lot of room and a lot of electrical power - both of which are unavailable on the conventional unmanned satellites and platforms. On top of that, there is the possibility of taking corrective action if such new instruments malfunction.

3.2 Microgravity Laboratory

Apart from being outside the Earth's atmosphere, the International Space Station also provides physical, environmental circumstances that are not available in the same kind or to the same extent on Earth. One of the most important factors here is weightlessness, or, to be physically more correct: reduced gravity, the so-called "microgravity". One hears many critics of space activities arguing that the whole business of weightlessness doesn't matter — most people live on Earth with a gravity of 1 g, and trying to discover what zero gravity does to a couple of astronauts out in space really isn't worth all that expense. This sounds like a cogent argument — but it's a false one.

Gravity is one of the fundamental forces in physics, but not of course the only one that affects physical and chemical processes. The classic procedure for investigating unknown phenomena is to create laboratory conditions in which the various factors influencing a process are separated so that their effect can be studied in isolation. While it is not impossible to exclude the effect of gravity on Earth, it is difficult and only possible for short durations. In space, it becomes possible to achieve a better absence of gravity, and over a very much longer period.

In the life sciences field there are proposals for research projects on the International Space Station

that concentrate on physiological and biological processes. In this area the Station will undoubtedly form a major springboard for scientific progress, since it will make it possible to subject experiments to space conditions over a very long period, as well as providing sufficient experimental subjects for physiology research in the person of the Station crew itself.

In physics, new aspects of materials processing in space and of fluid physics occupy the foreground. Here, it is expected that basic research on the International Space Station will be able to provide practical help for applications on Earth. Free of the influence of gravity, many processes important for modern technology can be studied and understood better. A good example are processes involving phase transitions: on Earth they play an important role in all combustion processes — and hence in the production of energy — and in the production of high-quality materials. The fuller knowledge that can be gained in space, freed from gravity, can make a decisive contribution back on Earth to understanding and controlling the earthbound processes better.

Another example involves the growing of protein crystals. We are not talking at all, here, about going into mass production of particular protein crystals on the International Space Station; but growing and investigating particular crystals in space can help widen our theoretical and practical knowledge so much that we shall then be able to master or improve the corresponding crystal-growing processes on Earth.

There are, in physics, other proposals for using the International Space Station for experiments in fundamental physics, e.g. for experiments on how dust, aerosols and plasma particles behave when weightless.

One possible utilisation of the International Space Station which is certainly of particular interest for the audience of the EFTF, is to try out a set of atomic clocks in space. At the Space Station Utilisation Symposium in Darmstadt in October 1996 [Ref. 1], such an idea has been presented under the name: ACES - Atomic Clock Ensemble in Space [Ref. 2]. It would consist of a Cold Atomic Clock, a Hydrogen Maser and a Trapped Ion Clock which would be installed on one of the external platforms of the International Space Station (Fig. 2). These clocks would be 1000 times more accurate than what is currently available with existing space clocks.

Furthermore, this set of clocks would not only constitute an improvement over existing space clocks, but, as far as the Cold Atomic Clock is concerned, it would also lead to an improvement with regard to what, today, is possible on the ground. In fact, the Cold Atomic Clock would work more accurately under weightlessness on the Space Station than under gravity on Earth. This is related to the principle of Atomic Fountain which is used in the clock. Under the microgravity conditions on the Station, the cold atoms cloud can traverse the microwave cavity with a slower speed than on Earth. This allows to increase the interaction time between the cesium atoms and the microwave field to 5-10 seconds which in turn reduces the atomic resonance linewidth. The result would be an improvement of one to two orders of magnitude in the frequency stability and accuracy over conventional devices. A concrete proposal for such a Cold Atomic Clock will be presented later at this Forum under the name "PHARAO" [Ref. 3].

3.3 A Test-Centre for New Technologies in Space

Besides using the special environment of space for gaining a better knowledge of processes down on Earth, there is also a field in which the International Space Station will be used for advancing space technology itself: this is in the testing, trying out and perfecting of new equipment and operating procedures.

Here, once again, we have to stress that the Station is not an end in itself, merely creating the technology for further space stations: the opportunities the International Space Station offers are also of value for developing and introducing new systems for unmanned spaceflight. They can help us introduce more rapidly technical and operational improvements in unmanned systems, such as communications and Earth-observation satellites.

Why? Because the operators of these satellites are - like the owner of any valuable piece of property - by nature very cautious and conservative. None of them wants to be the first to see his own new and expensive satellites fitted with a novel and unproven piece of equipment or material; they all prefer to first see concrete proof of the usefulness and reliability of the technical or operational improvements proposed by industry, through these having been used in comparable circumstances in space. This naturally inhibits innovation. So being able to test, perfect and demonstrate new materials, instruments and procedures on the International Space Station (Fig. 2) in the immediate environment in which they will later be employed will substantially increase the ability to innovate among space firms in the countries involved in the International Space Station.

In its role as a test centre in space, the Station provides for instance the opportunity for introducing

new energy-generating systems, closed life-support systems or remote-controlled robotic systems [Reference 4].

3.4 Platform for New Services

The International Space Station also allows entirely new commercial applications that have less to do with testing and trying out new techniques than offering services of a new kind that have been unavailable so far.

A practical example are radio-controlled clocks [Reference 5]. For synchronisation with a reference clock, these radio clocks generally activitate their receiver once a day to receive a time code signal from a long wave radio transmitter. During the remaining time the clocks run on a standard quartz oscillator which is precise enough for private applications. The current technology has reached a level of miniaturisation that allows to put the whole receiver and decoding circuits into a standard wrist watch. However, the time signal can only be received within the scope of the radio transmitter which is in the order of 2000 km. Travelling with the watch to other regions would envolve a different receiver and decoder. Furthermore, the low data rate of the long wave transmission requires long periods of good receiving conditions.

To overcome this limitation, there is a proposal [Reference 4] to install a radio transmitter onboard the International Space Station that is transmitting a UTC signal with a semi-omnidirectional antenna. The radio frequency of the transmitter must be in a range where the Earth's atmosphere is transparent. If one takes into account that the signal is to be received with a wrist watch that is hiding its antenna in the bracelet, the appropriate wavelength should be around 16 cm which corresponds to a frequency of 4 GHz. This is well within the best transmission window of the Earth's atmosphere. The orbital parameters of the Station allow to receive this signal almost on the whole inhabited surface of the Earth's several times per day.

The whole set-up onboard the International Space Station would only be an internal clock, a data processor and a time signal transmitter in the 4 GHz range (Fig. 2). The internal clock would be controlled and up-dated from a ground based station. The ground station will supply the processor with synchronisation marks and precalculated information on the overflight of the International Space Station for a given region. For that purpose, the Earth's surface will be divided into sectors. The precalculated ground track information will be transmitted together with the time synchronisation signal. It allows the wrist watch to activate its receiver the next time the Station passes over. At each activation, the wrist watch receives and stores several Station encounter times in advance, in case a connection cannot be made for some time.

4. ROLE OF EUROPE

4.1 Cooperation and Competition

Spaceflight systems and space-based services do not only call for extremely high quality and reliability, but their development, production and operation necessitate ressources which are today only available through international cooperation. However, international cooperation is far from meaning that anyone who wants to, can join in. Cooperation in certain areas does not rule out competition in other areas — or even in the same. You become an interesting partner only if you have capabilities and skills of your own that are of use to the others. Programmes as important as the International Space Station do not take "hitch-hikers" on board.

Building and running a space station calls not only for the orbiting station itself, but also for the associated transport systems and ground facilities. Europe, however, has neither the political intention nor the financial means to develop and operate such an extensive system on its own. By taking part in the International Space Station, Europe is able to focus its efforts on particular key elements while at the same time being able to make use of all the possibilities the Station will have to offer.

4.2 Key Elements of the European Participation

The elements making up European participation in the International Space Station (Fig. 6) ensure both that Europe has the capacity and the know-how needed to make it a worthwhile partner in global cooperative projects, and that Europe has the systems and infrastructure needed to make maximum use of the Station.

The core element of European participation is the Columbus Orbital Facility (COF) (Fig. 3). This is a manned multipurpose laboratory, which with its modular and interchangeable internal equipment can be tailored to various experimental purposes.

All the partners in the International Space Station have to bear their share of the outlay incurred in running and using it. Instead of paying money into a common operating fund, partners have the option of supplying

operational and logistical services, such as providing transport. Europe will use the Ariane 5 launcher for such purposes.

However, a rocket-type launch vehicle like Ariane 5 cannot directly fly and dock to a space station. To carry out the necessary rendezvous and docking manoeuvres with the International Space Station, Europe is developing a transfer stage for Ariane 5, known as the Automatic Transfer Vehicle, or ATV (Figure 4). The ATV will also be able to move the whole Station up to a higher orbit, to compensate the frictional losses caused by the residual Earth atmosphere still existing at the Station's altitude.

The ability to carry astronauts out into space and bring them back to Earth is of strategic importance for the further exploration and utilisation of space. Europe is therefore also investigating the technical and financial possibilities of developing a Crew Transport Vehicle, or CTV. At present, there are two options: either a two-way CTV that, launched from Kourou on an Ariane 5 (Fig. 6), could carry astronauts to the Station and back to Earth for routine crew exchange, or what is known as a Crew Rescue Vehicle (CRV) that would fly individual astronauts back to Earth in an emergency situation.

5. INTERNATIONAL PARTNERSHIP

One extremely important aspect of the International Space Station is Russia's part in the international partnership (Fig. 7). In December 1993 the earlier four partners, i.e. the USA, the 10 participating European States, Japan and Canada, jointly invited the Russian government to take part in the planned space station. So the International Space Station now brings 14 different countries together in a global cooperation programme; the world's biggest cooperative programme in the scientific and technical field.

Russia's participation makes it possible to bring Russian space engineers and scientists out of a decades-long, self-imposed isolation and into the international structures for cooperation. At the same time it offers the Russian experts — in a critical period of profound changes — a certain stability in their professional prospects; and this can help prevent a migration of these specialists into other undesirable areas or countries, an aspect that should not be underestimated when dealing with high technology of such strategic importance.

Through this cooperation with Russia the western partners naturally share in the rich experience that Russia has gained in building and running space stations, ranging through the various Salyut-series stations to the Mir station in operation today.

Over and above the positive effects the International Space Station is having in East/West cooperation, the Station is an important pacemaker for the increasing globalisation of the political and industrial structures.

Because of the growing market globalisation, in particular after the ending of the East–West conflict and the emergence of new suppliers on the world market that this has brought about, we have reached a point where in the aerospace industry there is a need for worldwide harmonization. To a great extent, however, the political structures and practical experience needed to attain this are still absent. Cooperation on the International Space Station is an important step towards this. Just by itself, the negotiation that has gone on for years on the international agreements for the Station has — as a kind of "training on the job" — produced a whole generation of experts at various levels in the governments who have the know-how, despite of differing political interests, differing administrative and legal set-ups and differing mentalities, to arrive nevertheless at a balance of interests.

6. CONCLUSION

To summarise, we can say that, on account of its technical, operational and political framework, the International Space Station definetely constitutes a great challenge. On the other hand, we can expect that the significantly better performance potential of the International Space Station compared with the systems that have existed up to now, together with the global approach to its development, operations and utilisation, will guarantee high quality in the scientific and technical work done on board and offer opportunities that have not been possible until today.

7. REFERENCES

[1] *Utilisation of the International Space Station Symposium*, Darmstadt, Oct. 1996, Symposium Proceedings, ESA Publication SP-385, ISBN 92-9092-223-0

[2] C. Salomon, C. Veillet: *ACES: Atomic Clock Ensemble in Space*, in: [Ref. 1], p. 295-297

[3] C. Salomon, P. Lemonde, P. Laurent, E. Simon, G. Santarelli, A. Clairon, N. Dimarcq, P. Petit, C. Audoin, F. Gonzalez, F. Jamin Changeart:

PHARAO: A Cold Atomic Clock in Microgravity, in: [Ref. 1], p. 289-294

[4] F. Didot, P. Putz, S. Losito: *External Robotics and Automation Experiments Survey*, in: [Ref. 1], p. 495-500

[5] F. Huber, E. Messerschmid: *Clock Synchronisation and Time Dissemination*, in: [Reference 1], p. 491-493

[6] W. Schulte, P. Hofmann, P. Rank, H. König: *The Space Exposure Biology Assembly (SEBA)*, in: [Ref. 1], p. 167-170

[7] P.C. Galeone, R. Trucco, F. Pepe, L. Genovese: *The Hexapod Pointing System*, in: [Ref. 1], p. 643-646

8. FIGURES

(All artist's views made by David Ducros for ESA)

Figure 1: The International Space Station: a cooperative programme of 14 countries

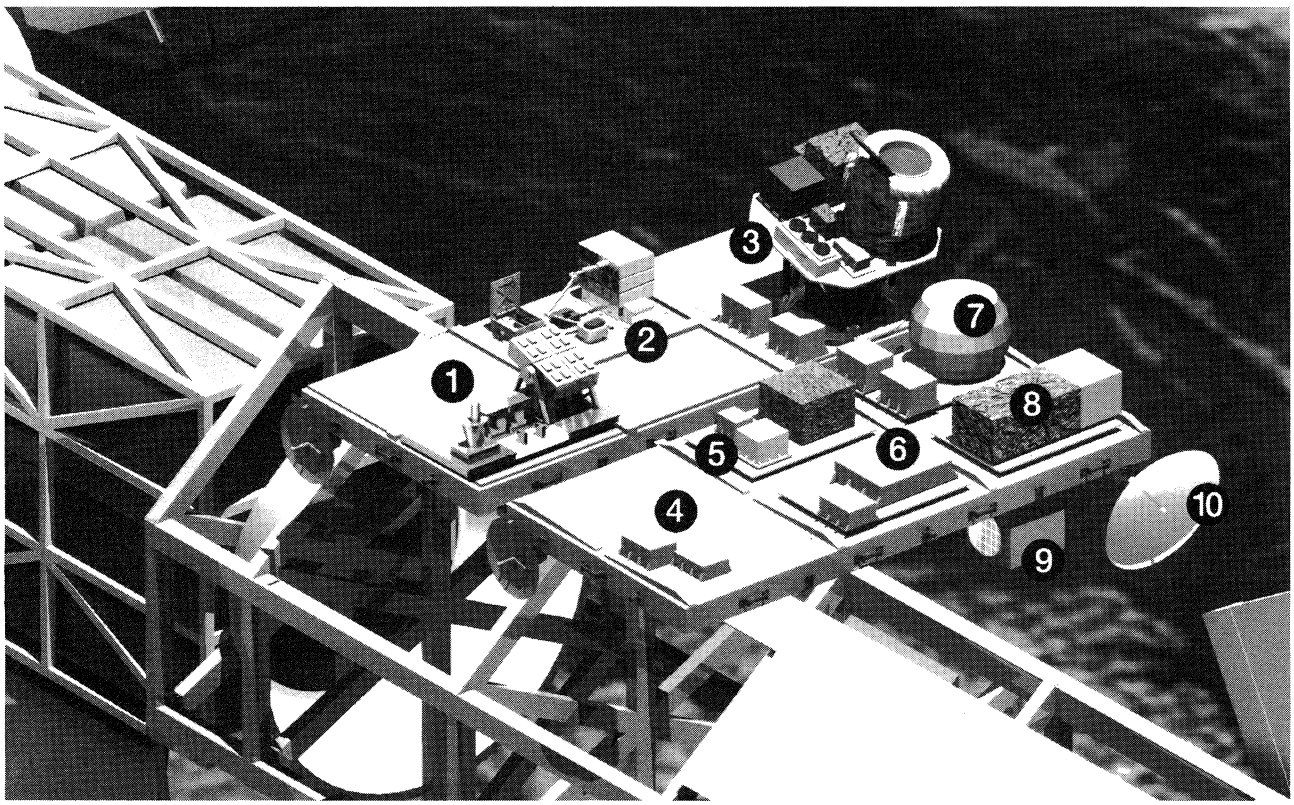


Figure 2: Some examples of research and applications payloads proposed [Ref. 1] for installation on external platforms of the International Space Station: From the left above to the right below: (1) Space exposure biology assembly [Ref. 6], (2) technology payload using a robotic manipulation system [Ref. 4], (3) telescope mounted on a Hexapod pointing system [Ref. 7], (4) Radio-controlled clock synchronisation assembly [Ref. 5], (5-10) ACES - atomic clock ensemble in space [Ref. 2] consisting of (5) a trapped ion clock, (6) the processor electronics, (7) a hydrogen maser, (8) the PHARAO cold atomic clock [Ref. 3], (9) a laser link with the ground station, and (10) a radio-frequency antenna for the dissemination of time synchronisation signals

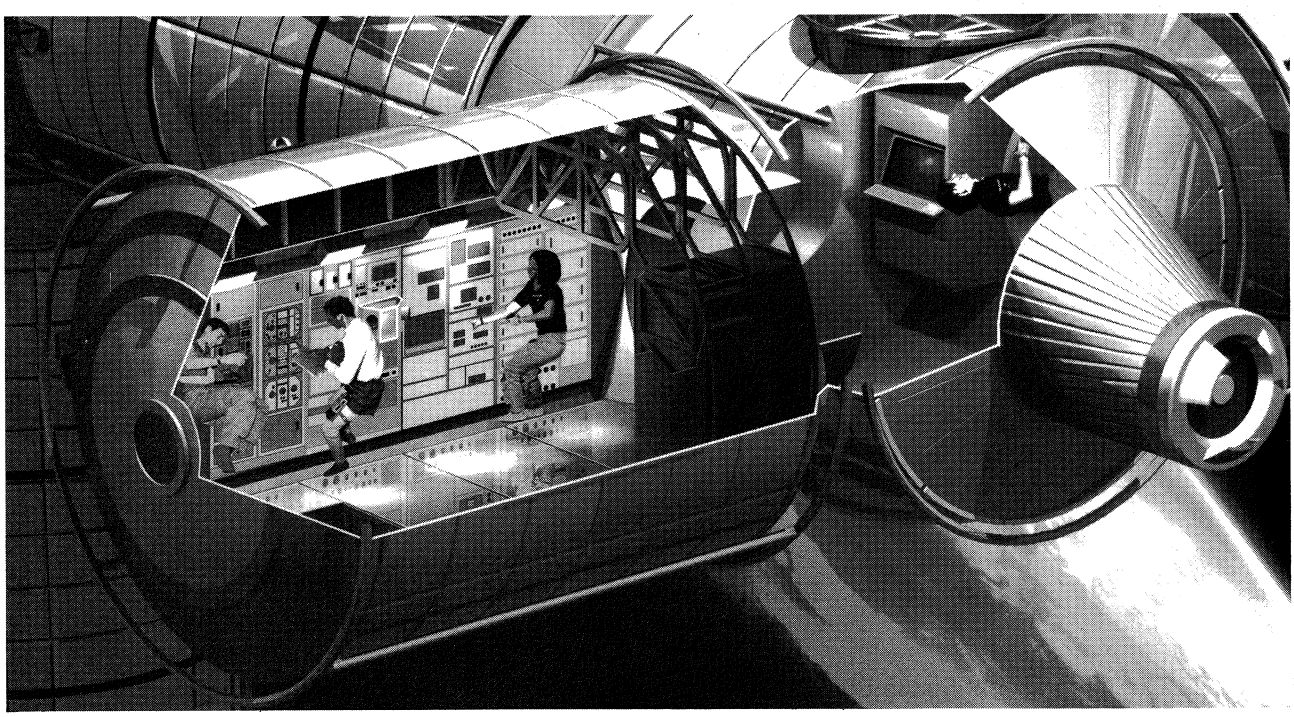


Figure 3: The Columbus Orbital Facility (COF): the core element of the European participation in the International Space Station

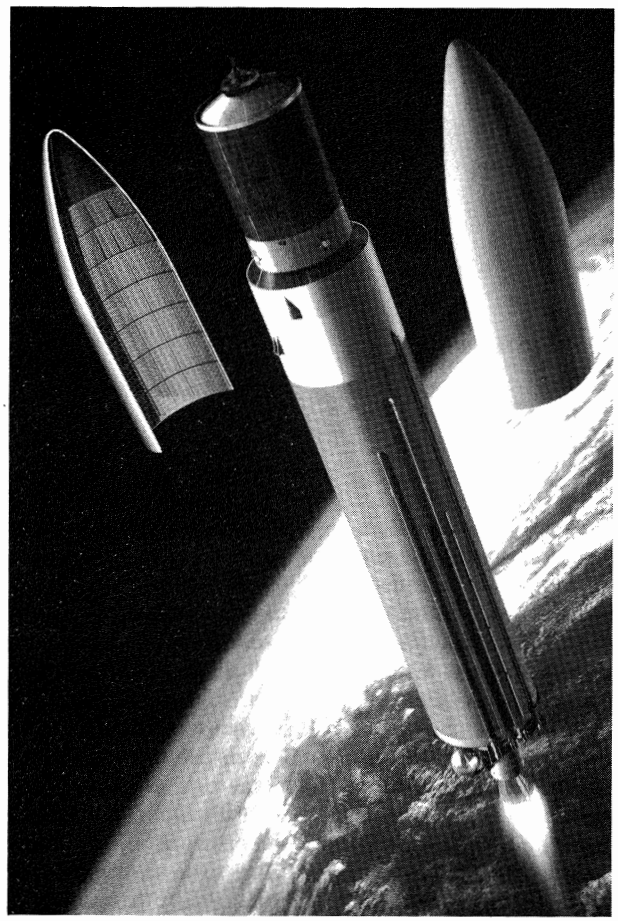


Figure 4: The Automatic Transfer Vehicle (ATV): unmanned automatic transfer stage for Ariane 5 to carry resupply goods to the Station

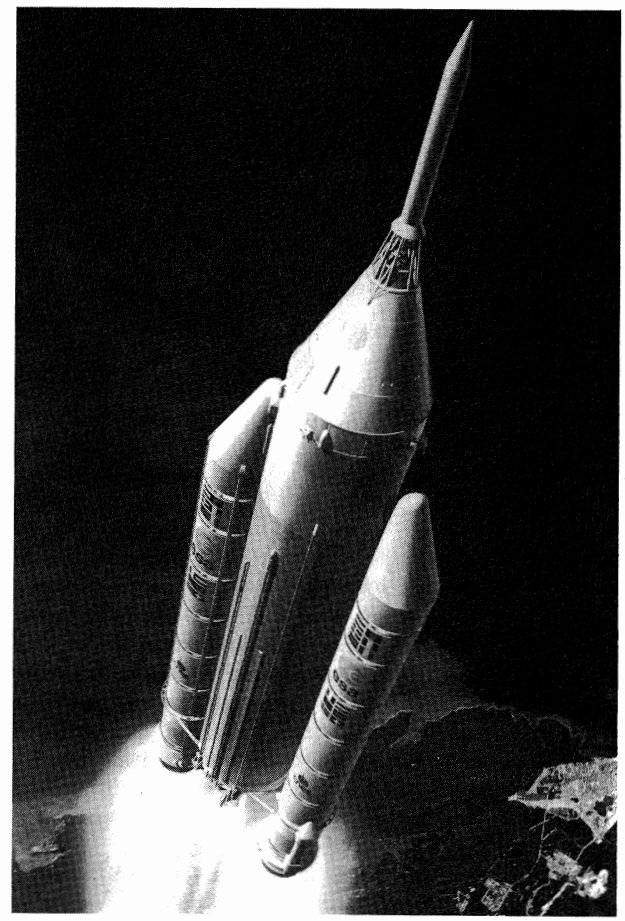


Figure 5: Artist concept of a Crew Transport Vehicle (CTV), launched from Kourou on Ariane 5

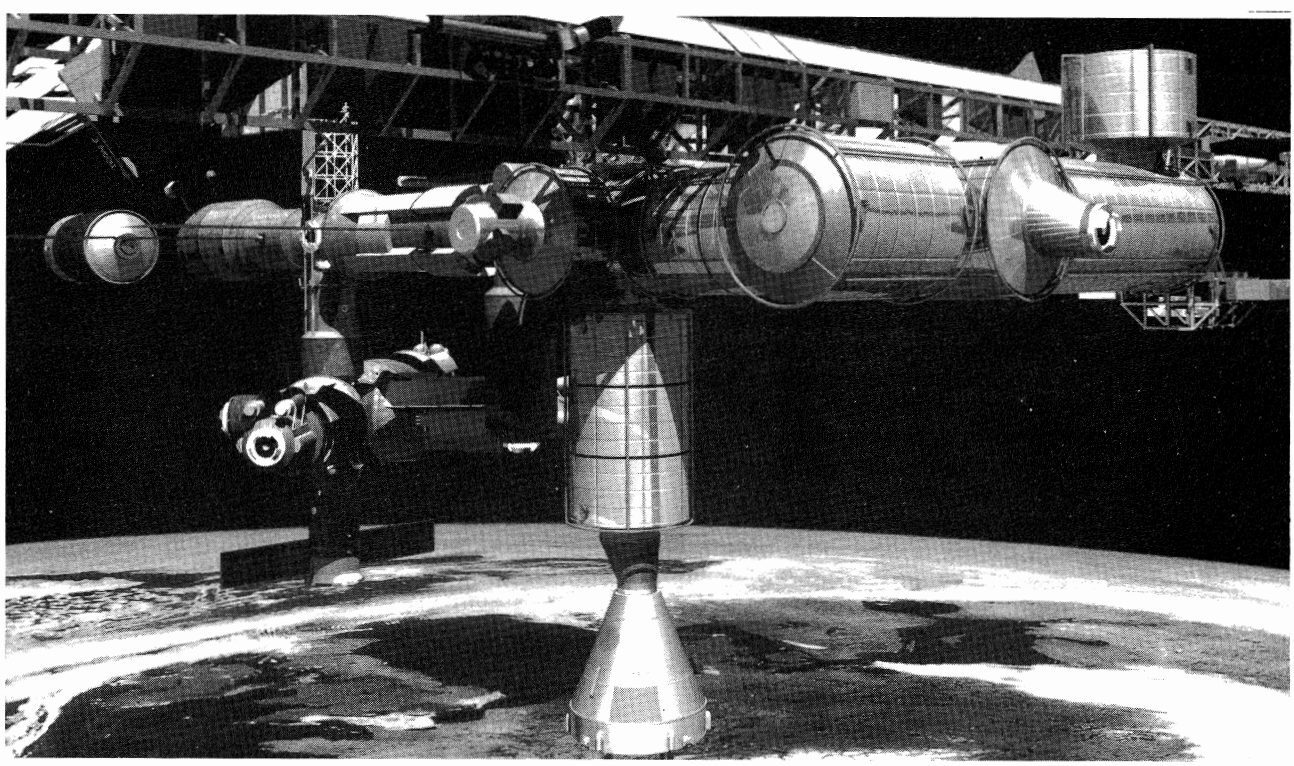


Figure 6: Columbus Orbital Facility (module on the left in the foreground), Crew Transport Vehicle (visible in the middle, docked below the vertical US Habitation Module) and Automated Transfer Vehicle (approaching the Station at the rear on the left)

METROLOGY AND RELATIVITY

Bernard Guinot

Observatoire de Paris
61, avenue de l'Observatoire, F-75014 Paris

ABSTRACT

In 1992, the Comité International des Poids et Mesures created a working group on the *Application of General Relativity to Metrology*. This paper first explains why this group was needed, then comments some of the main points raised in a report issued by the group (now in print in *Metrologia*): the role of the Einstein's principle of equivalence and the distinction between proper quantities and coordinated quantities.

Keywords: general relativity, metrology, SI units, coordinates.

1. THE GENESIS OF A WORKING GROUP ON METROLOGY AND RELATIVITY

The scientific background to this communication, the theory of general relativity, was available some 70 years ago. However, it was familiar only to a small number of theoreticians. The circle of initiates began to widen in the 1960s as a consequence of the development of theoretical astrophysics and cosmology, and of progress in time measurement and in applications of time/frequency techniques (Ref. 1).

In the fundamental metrology of time, one observes that the relativistic effects have always been taken into account correctly, as soon as this is required by reduction of the uncertainties. This is to credit of the many colleagues who explain to neophytes the practical consequences of these effects, sometimes in front of an audience barely interested in such esoteric considerations. I remember communications and papers by Hudson, Becker *et al.*, Hudson *et al.*, Reinhardt, Ashby and Allan (Refs. 2-6), but I am sure I have forgotten other pioneers and I ask them to forgive me. From this era, one should also emphasize the role of experiments based on clock transportation, among them in particular those of Hafele and Keating, Alley, Vessot (Refs. 7-9), which, in addition to their intrinsic value, drew the attention to practical aspects of the theory. Pursuit of a rigorous treatment of time measurement continues and, among the most recent papers one should mention those of Refs. 10-12, on synchronization and syntonization.

In spite of all this effort, both theoretical and practical, in spite of many didactic papers, in particular those presented at our Forum, the situation is not yet fully satisfactory. Many users know the formulae to be used to account for relativistic effects, but apply them as *corrections* to an underlying Newtonian model of space-time, with all the risk of error that this entails : an example of such an error, concerning the GPS (Ref. 13), was mentioned at a previous Forum. One may observe also that a confusion still exists between the second, as unit of proper time, and the scale unit of International Atomic Time, TAI. Among other difficulties, one might mention that the International Astronomical Union (IAU) defined in 1976 the relativistic time scales to be used in dynamical astronomy without stating explicitly the space-time coordinate systems to which they pertain. This generated difficulties and controversies to which the IAU and the International Union of Geodesy and Geophysics (IUGG) tried to put an end in 1991 by introducing coherent definitions of reference systems in the framework of general relativity (Refs. 14-16). However, some confusion still affects the dynamics of the solar system and space geodesy.

The official bodies in charge of the International System of Units (SI) have given little indication of how to use units in a relativistic framework. I believe that everyone agrees that relativity should not be introduced in the definitions of units themselves. Nevertheless, it is essential to ensure that the definitions are compatible with the theory required in the most accurate applications and that all possible ambiguities of interpretation are removed. Indeed, the Comité Consultatif pour la Définition de la Seconde (CCDS) regularly considers the relativistic aspects of the measure of time, but only in 1980, at the request of the Comité International des Radiocommunications (CCIR), did the CCDS issue a *declaration* on the relativistic definition of TAI (without stating explicitly the reference system to which it belongs) and provide formulae for time comparisons. This declaration, together with explanations on the meaning of the definition of the second and the precautions to be taken when comparing distant clocks, appears in the 6th edition of the publication on the SI (1991) by the Bureau International des Poids et Mesures (BIPM) and will be considered more extensively in the 7th edition. This text, however, is necessarily too compact to have a didactic role.

The lack of information on the metre is even more striking. The definition of the metre, however, is based on the constancy of the velocity of light, measured locally, a

consequence of a postulate of general relativity. On the other hand, the recommendation specifying the rules for the practical realization of the metre, issued by the Comité International des Poids et Mesures (CIPM) in 1983, contains an ambiguous statement on the need to correct for gravitation.

In April 1992, this situation led me to propose to the Director of the BIPM, the creation of a working group on the Application of General Relativity to Metrology with two missions : first to prepare a report on the interpretation and use of the SI units in the framework of general relativity ; then to study the consequences of the increasing accuracy with which these units are realized. This proposal was examined by the CIPM in October 1992 and it was decided to entrust the CCDS with the formation of the working group and the responsibility for running it. The composition of the group is given in the annex : the members are experts in general relativity and metrology, and representatives of scientific unions. The first mission of the group was completed recently and a report to *Metrologia* is in press. I shall now present its principal points to the Forum. Before I begin, however, I wish to make clear that this report is a collective work coming from the group.

Undoubtedly, those who have no notions on general relativity will find the report rather hard to follow. It does not, however, enter deeply into the theory. The essential is the understanding of its basic principles. These are well described in elementary books (for example, Ref. 17), but are also the subject of deeper reflections (for example, Ref. 18).

2. METROLOGY AND GENERAL RELATIVITY

From the definitions of the SI base units it appears that basic metrology is implicitly treated in the familiar framework of Newtonian absolute time and Euclidian space. This theoretical framework being insufficient, the first step is to specify the model to be adopted. Einstein's general relativity was chosen because it is the simplest theory which agrees with experiment and observation, within the limits of their uncertainties. That does not imply that this theory is more *true* than any other one, merely that it is the most convenient. It may be that we shall have to adopt another theory in the future. In the meantime, the report is built on general relativity and accepts without discussion the postulates of this theory and their consequences.

Although metrology requires a relativistic treatment only for time and length measurement, we believe that the general consideration of all types of quantities in a relativistic framework could lead to a better understanding.

It appears that there are two keys for this understanding :
- the postulate known as the Einstein's principle of equivalence,
- the distinction between quantities which are directly measurable with standards, the *proper quantities*, and the coordinates (and quantities constructed from these coordinates designated as *coordinate-quantities*).

2.1 The Einstein's principle of equivalence

One basic postulate of special relativity says that no non-gravitational experiment makes it possible to distinguish two inertial frames (or Lorentzian frames). This is an idealization and generalization of the observed invariance of the velocity of propagation of electromagnetic waves. General relativity, which is a metric theory of gravitation, was built in such a way that this postulate remains valid locally. Thus, if the same local, non-gravitational, physical experiment is performed anywhere and anytime in the universe, and if the measurement of its evolution is based on identical standards, the mathematical description of this evolution is the same. An important consequence is that the coupling quantities which enter into the laws of local physics, when referred to local standards, take the same values. Such a statement on local physics is often referred to as the Einstein's principle of equivalence (EEP). It is a very powerful statement : it justifies the existence of universal laws of physics and of physical constants, but emphasizes their local character. It implies, in particular, invariance of the velocity of light in local measurements. Whether the EEP implies invariance of the constants of atomic physics or not may be a matter of controversies. In our report, this invariance is explicitly accepted. Therefore, the suitability of atomic clocks to represent the time scales needed for the study of macroscopic systems appears as a consequence of the EEP. This is in agreement with the spirit of IAU Recommendation A4 (1991), in which coordinate times for the dynamics of the solar system are defined on the basis of the definition of the (atomic) second.

The restriction to local and non-gravitational experiments made by the EEP calls for further consideration.

A space laboratory orbiting the Earth may be a good realization of a local Lorentzian frame : the gravitational field generated by external bodies, mainly that of the Earth, can be detected only by its non-uniformity (i. e. its tidal terms) which causes effects of order 10^{-16} per metre in relative value. On the other hand, gravitational interactions within the laboratory are usually negligible. Where this is not the case, as in the measurement of the constant of gravitation, a sufficient approximation is to treat this interaction, predicted by general relativity, in the same way as other forces, for example, as electrostatic forces.

In a ground-based laboratory, the EEP can still be applied. Inertial and gravitational forces are treated in the classical manner. Thus, general relativity does change our familiar physics, providing that the experiments are local (however it must be remembered that light propagates in a straight line only in a Lorentzian frame and that the apparent curvature of the photon trajectory may have to be taken into account). We should not be surprised : were this not the case, the theory would not be viable. But, what is considered to be local is becoming smaller and smaller. Until about 1920, the full solar system could be treated as local, later one had to take account of relativistic effects. In geodesy, the Earth ceased recently to be local. In time measurement we reach the situation where the volume of an atomic clock cannot be seen

as sufficiently local, the gravitational frequency offset being 10^{-16} per metre of altitude.

To fulfil the mission of our working group, we had to check that the definitions of the SI base units are based on local physics and that the realization of these units can be accomplished locally. We will not here examine thoroughly the definitions of the seven SI base units : it is easy to observe that the second, the kilogram and the kelvin are independently defined on the basis of local properties. The other base units are derived from these independent units using physical laws in local experiments, an example being the definition of the metre. For this reason definitions are acceptable in general relativity. It should be noted that this would not have been the case with earlier definitions, as in the definition of the second as a fraction of tropical year and the definition of the metre as a fraction of the Earth's meridian. It would also not be the case if there were advantages in defining the second using pulsars.

It is nevertheless interesting to look more closely at the definitions. We take the example of the metre. The space extension of a metre is such that the definition is ambiguous at the level of 1.10^{-16}m at the surface of the Earth. What should we do when such an ambiguity becomes unacceptable ? With rigour, one can say that the definition applies only in the flat tangent space ; practically, one can consider that the definition applies at some sufficiently small sub-multiple of the metre. Consider next the kilogram. General relativity is based on the equivalence of inertial and gravitational mass ; this poses no problem in itself. But the definition of the mass of an extended body is a difficult problem and, in principle, the mass equivalent of the energy accumulated in the body must be taken in account. Fortunately, the kilogram is so small that this leads to negligible effects, relative to the uncertainty of measurements.

We arrive thus at the philosophically disturbing conclusion that, strictly speaking, metrology as we understand it usually does not work in curved space-time. Mathematically this is a consequence of the differential geometry on which general relativity is based. We can forget these difficulties only in a region of space sufficiently small that curvature can be neglected. In this region, the definition of a quantity given by the International Vocabulary of Basic and General Terms in Metrology (Ref. 19) can be employed : *an attribute of a phenomenon, body or substance that may be distinguished qualitatively and determined quantitatively.* Such a quantity does not depend on its location in space-time and, therefore, on the coordinate system used to specify this location, it is a *proper quantity.* The units being quantities taken as reference, the conclusion of this discussion is that

the SI base units are defined as proper units.

To avoid any ambiguity in the numerical value of a proper quantity, the quantity must be measured against co-moving standards of the SI units, remaining in their vicinity.

We should note that the output of an atomic clock, at its connector, does not extend in space, so the *proper time* of a clock is a well defined finite quantity. Proper time is not only the best measured quantity, it is also the best defined one. These properties are widely used in astronomy and space geodesy. They also raise the question (to which I have no answer) : is time a more fundamental quantity than the other ones ?

2.2 Coordinates

In Newtonian mechanics, coordinates are implicitly defined by the postulates of the theory, for example the principle of inertia (with freedom in orientation and translation). They are assumed to be directly measurable everywhere and anytime using realizable standards. All local clocks can be synchronized with a single ideal clock which ticks seconds timing the universe. Cartesian coordinates can be graduated by carrying an ideal metre stick along specified axes. These space-time coordinates directly provide the units needed in physic, whatever be the location of the phenomenon under consideration and its extension in space and time.

Unfortunately, we have to accept the fact that space-time is curved : locally, these convenient properties of Newtonian space-time are approximately valid, but it is impossible to find a coordinate system which makes them globally valid. In general relativity, the coordinates must be seen as a system for labelling events, devoid of physical meaning. To emphazise this fact, some authors (Ref. 20) consider that they are simple numbers, without dimension or unit. An example is the number of a house in a numbered street : no dimension, no unit... but this is a useful coordinate system ! This didactic point of view has great merits, but, as we shall see, leads to complications in practice.

Although there are no privileged coordinate systems in general relativity in the same sense as in the Newtonian model, there are systems which are more convenient for specific problems. Given the success of Newtonian mechanics, it is inviting to think that coordinates which are not very different from the absolute time and usual coordinates of Euclidian space will lead to the simplest theoretical developments. For physics on the Earth, or close to it, geocentric coordinates in a Post-Newtonian Approximation have been defined by the IAU and the IUGG, for a non-rotating system and for a system rotating with the Earth. In the domain where they are useful, these coordinates remain very close to proper quantities : the time coordinate, common to both systems and known as Geocentric Coordinate Time, TCG, and the space coordinates have graduation units which differ, on the ground, from the proper second and the proper metre by about 7.10^{-10}, in relative value. Although this difference is small, it is not negligible in precise applications. We know that very well in the domain of time. But consider another example.

In space geodesy, observers on the ground measure proper quantities : received frequencies, times of pulse arrival referred to their clock. These have first to be transformed into coordinate quantities by reference to International Atomic Time or to TCG. A global treatment provides then satellite ephemerides which express space coordinates as functions of coordinate time : these are truly relativistic coordinates, as are the geodetic coordinates also obtained. Both the ephemerides and the geodetic coordinates depend critically

on the choice of coordinate system, not only for origin and orientation, but also for scale. The coordinate system must be unambiguously defined. For example, whether one uses as coordinate time TAI or TCG, the difference on the Earth radius (a coordinate quantity) reaches 5 mm : this is important, for example, in studies of the mean sea level by satellite altimetry in an experiment such as Topex/Poseidon.

In the domain of time, coordinates play an essential role in establishing a world-wide system for timing and for time comparisons : this is well known. With some surprise, however, I discovered that, although the convention of coordinate synchronization is widely accepted, the convention for clock comparisons may be controversial and must be defined explicitly. In our report, the recommended convention is that clock comparisons express the difference of readings at instants receiving the same date in TCG (or equivalently in TAI) at two locations.

Many other examples of the use of coordinates, in various coordinate systems, can be found in astronomy. They all show the importance of making a clear distinction between coordinates and proper quantities, a distinction which will become essential for an increasing number of scientists and technicians as measurement techniques improve. One might expect that this would be widely known. Unfortunately this distinction tends to be hidden for two reasons: one is the use of scaling factors in the definition of coordinates, the other one is the use of the same units for proper and coordinate quantities. Let us examine these matters.

The use of scaling factors in coordinates was initiated by the IAU in 1976, when defining coordinate times for dynamics. The aim was to avoid a secular divergence of the time argument of astronomical ephemerides with respect to TAI, so that these ephemerides could be used as Newtonian ephemerides by non-experts who did not need their ultimate accuracy. At the time, this seemed a natural decision. It was later recognized that this raised difficulties because, in practice, this was equivalent to the introduction of new proper units differing from the SI units (see, for example, Ref. 21).

Much more surprising is the attempt of the Conventions of the International Earth Rotation Service, an important document published in 1996 (Ref. 22), to introduce scale factors in geodetic coordinates with the objective of restoring (approximately) the agreement of their scale unit with the proper metre on the geoid. The consequences of using these factors are complex and risk the creation of ambiguities at a level of 10^{-9} in relative value.

The smallness of the difference between coordinates and proper quantities encourages the use of the same units for both. Indeed, this use is justified by the rules of quantity calculus (Ref. 23) : all quantities of the same type must have the same unit. Relaxation time, life time, for example, have the second as unit. Radius, diameter, circumference, ... , are expressed in metres. Similarly, coordinate times and space coordinates are expressed in seconds and metres. A general rule is that units do not define quantities. It is nevertheless easy to see that these rules of quantity calculus lead to problems, even paradoxes, when applied to relativistic coordinates or, more generally, to scales (of time, temperature,...). This is explained in an annex to the report. I shall not develop this subject further, I simply recall that when one writes TAI – UTC = 30 s, one equals the difference of simultaneous readings to a duration (by definition, *the second is the duration of...*) ! The system for identifying the graduations of scales has, in my opinion, not yet been resolved satisfactorily (but I know that others may disagree with this).

3. GENERAL RELATIVITY AND METROLOGY OF TIME

A large part of the report is devoted to the metrology of time. We attempted to state the basic conventions clearly and to develop the usual formulae, making explicit use of these conventions, in a didactic way. Thus the report can be seen as an introduction to more complete developments. One of the objectives of the report is also to unify the notations and the sign conventions in applying the theory of general relativity. It is not our role to make recommendations on this subject, but we carefully followed the conventions of the IAU Resolution of 1991 in the hope that they will be generalized.

4. CONCLUDING REMARKS

When working on the report I have presented here, I had to look at some past developments in metrology, especially the measure of time. I became increasingly conscious of the pragmatism in these domains. To take examples in the fields I know the best, consider International Atomic Time. It was born, because it was a natural action to accumulate seconds produced by atomic standards ; but the relativistic effects were not considered until the pressure of the accuracy of the standards made it necessary. Might it have been better to define, instead of TAI, a geocentric time without scaling factor ? Could it be that the choice of the frequency of the caesium atom was not the best one ? [Another choice could have reduced the frequency of the UTC steps.] Could it be that the choice of the origin was not appropriate ? [The time offset of 32,184 s with Ephemeris Time and Terrestrial Time is inconvenient.]

Since action often preceded reflection, many examples of unfortunate decisions can be found. This was largely unavoidable, and one may also argue that too much reflection is sterilizing ! Nevertheless, it is always better to anticipate problems and to look for best solution. In the field of time, progress is so fast that we will shortly be faced with shortcomings in existing definitions and basic data. Examples are the definition of the geoid, the influence of Earth tides, the value of the gravitational potential, the definition of the data needed for a terrestrial use of space clocks. To anticipate these problems and carry out the corresponding studies is the next task of the working group. This is a difficult task, requiring extended cooperation with interested bodies, but it is an essential preparation for the future.

5. REFERENCES

1. Will C.M., Theory and experiment in gravitational physics (revised ed.), Cambridge University Press, 1993.
2. Hudson G.E., Space-time coordinate systems, Actes du Congrès international de Chronométrie, Lausanne, 1964.
3. Becker G., Fischer B., Kramer G., Müller E.K., Die Definition der Sekunde und die allgemeine Relativitätstheorie, PTB Mitt., *77*, 111, 1967.
4. Hudson G.E., Allan D.W., Barnes J.A., Hall R.G., Lavanceau J.D., Winkler G.M.R., A coordinate frequency and time system, Proc. 23rd Ann. Freq. Control Symp., USAE Com., May 1969.
5. Reinhardt V., Relativistic effects of the rotation of the Earth on remote clock synchronization, Proc. 6th PTTI, Dec. 1974.
6. Ashby N., Allan D.W., Practical applications of relativity for a global coordinate time scale, Radioscience, *14* (4), 649, 1979.
7. Hafele J.C., Keating R.E., Around the world atomic clocks, Science, *177*, 166 and 168 (two papers), 1972.
8. Alley, C.O., Relativity and clocks, Proc. 33rd Ann. Freq. Control Symp., 1979.
9. Vessot R.F.C., Levine M.W., A test of the equivalence principle using a space-borne clock, General relativity and gravitation, *10*, 181, 1979.
10. Klioner S.A., The problem of clock synchronization : a relativistic approach, Cel. Mechanics and Dynamical Astronomy, *53*, 81, 1992.
11. Petit G., Wolf P., Relativistic theory for picosecond time transfer in the vicinity of the Earth, Astron. Astrophys., *286*, 971, 1994.
12. Wolf P., Petit G., Relativistic theory for clock syntonization and the realization of geocentric coordinate times, Astron. Astrophys., *304*, 653, 1995.
13. Spallici A., Jiménez C., Prisco G., Ashby N., Relativistic effects for low-Earth orbit satellites using GPS, Proc. 6th Eur. Freq. and Time Forum, 61, 1992.
14. Proceedings of the International Astronomical Union, Vol. XXIB (J. Bergeron ed.), Kluwer, 1991 (see Resolution A4).
15. Guinot B., Atomic time scales for dynamical astronomy, Proc 6th Eur. Freq. and Time Forum, 51, 1992.
16. Willis P., IAG Newsletter, Bull. Géodésique, *65*, 189, 1991.
17. Schutz B.F., A first course in general relativity, Cambridge University Press, 1990.
18. Anderson J.L., Principles of relativity physics, Academic Press, 1967.
19. International Vocabulary of Basic and General Terms in Metrology, ISO, Genève, 1993.
20. Misner C.W., Thorne K.S., Wheeler J.A., Gravitation, Freeman, 1973.
21. Soffel, M.H., Brumberg V.A., Relativistic reference frames including time scales : questions and answers, Cel. Mechanics and Dynamical Astronomy, *52*, 355, 1991.
22. McCarthy D.D. (ed.), IERS Conventions (1996), IERS Technical Note 21, Observatoire de Paris, 1996.
23. de Boer J., On the history of quantity calculus and the international system, Metrologia, *31*, 405, 1995.

ANNEX

Membership of the CCDS Working Group on Application of General Relativity to Metrology

N. Ashby	
J. de Boer	Comité Consultatif des Unités
C. Boucher	International Union of Geodesy and Geophysics
V.A. Brumberg	
T. Damour	
T. Fukushima	International Astronomical Union
B. Guinot	(Chairman)
W. Israel	
G. Petit	
B.W. Petley	Comité Consultatif pour la Définition du Mètre
	International Union of Pure and Applied Physics
M. H. Soffel	
J..McA Steele	International Union of Radio Science
C. Thomas	(Secretary)
C.M. Will	
P. Wolf	

Other contributors : C. Audoin, N. Capitaine, P. Giacomo, P. Noerdlinger, T.J. Quinn.

RESONATORS I

Chairman: W. Zingg

INFLUENCE OF THE MOUNTING ON THE MODE SHAPE OF THE THICKNESS SHEAR CONTOURED RESONATORS.

J.Détaint*, B.Capelle, A.Zarka**.**

*** FRANCE TELECOM CNET/DMR. 196, Av. H.Ravéra; BP 107, 92225 Bagneux France,**
**** L.M.C.P. Université Paris VI & CNRS. 4, Place Jussieu; 75252 Paris Cedex 05 France,**

Abstract

To study the influence of the mounting, a model of these resonators based upon the algebraical solutions found for a particular unspherical contour was build. The solutions for this geometry are already good approximations of those existing for spherical contours. The influence of the usual boundary conditions (totally free edge, "infinite" plate, totally fixed edge) on the mode shapes and of the mounting in both x1 and x3 direction were studied. It was observed that the solution should, then, include higher order terms so that the mode shape can become more intricate. In a second part, from computed results obtained with a numerical model based upon the previous solutions, we show that the same effects exist also, for conventional resonators having a spherical contour. The third part we report experiments which confirm that the predicted effects of the fixation does exist. However, a larger amplitude than expected from the computations was in sometimes found for these effect that may result of slightly actual conditions at the mounting. Other features, probably of the same nature, but that can only be explained in the frame of more general theories were also observed.

INTRODUCTION.

The bevelled, the bi-convexe and the plano-convexe resonators are of fundamental importance for frequency generation and for filtering. Many theoretical works were recently devoted to further understand their properties and to improve their characteristics [1 to 7]. Most of these recent developments were based upon two principal types of theory: the theory of the essentially thickness mode [8] and the plate theories [9-10]. In this communication our purpose is to investigate further the influence of the boundary conditions at the edge of the contoured resonators and particularly to determine the effects of fixing the resonator along two segments of its perimeter. To this end, in the frame of the theory of the essentially thickness mode, we have chosen simple cases for which it is easy to detect the influence of different types of boundary conditions on the mode shape and on the properties. From the theoretical point of view, one of the most simple case, which is actually very close of the usual plano-convexe resonators is constituted by resonators with unspherical contours having thickness variations related to the lateral anisotropy of the crystalline plates. For this geometry, the solutions with a quite simple algebraical expression exist particularly for totally electroded elliptical plates having an axis ratio related to the lateral anisotropy. So, it is possible to observe accurately the effects of the fixation on the mode shape. In a second part we consider the usual resonators with spherical contours. A numerical model based upon the previous solutions is described and applied to compute the effect of the fixation. In a third part, we report experimental results showing some cases for which effects of particular intensities were observed.

I RESONATORS WITH UN-SPHERICAL CONTOURS

I.1 Model for un-spherical contours.

The mathematics of the model are substantially similar to those we have previously presented [3] for the bevelled resonators and only the most important facts useful to understand the results will be recalled. This model is based upon the theory of the essentially thickness shear modes established by H.F.Tiersten and D Stevens [8]. In this theory, the lateral variations the main thickness displacement (or more exactly the lateral variations of a transformed displacement) are the solutions of a partial derivative scalar equation which is valid for the modes (anharmonics) situated in the vicinity of one overtone (noted n) of a given thickness mode (noted (1)).

$$M'_n \tilde{u}^n_{1,11} + P'_n \tilde{u}^n_{1,33} + (\rho\omega^2 - n^2\pi^2 c^*/4h(x_1,x_3)^2)\tilde{u}^n_1 = (-1)^{(n-1)/2} 4\rho\omega^2 e_{26} V_0 \exp(j\omega t)/(c^{(1)} n^2\pi^2) \quad \{1\}$$

$\tilde{u}_1(x_1,x_3)$ is the lateral dependence of $\hat{u}_1(x_1,x_2,x_3)$ the transformed displacement: $\hat{u}_1(x_1,x_2,x_3)=\tilde{u}_1(x_1,x_3)\sin(n\pi x_2/2h)$. C* is either $\overline{C}^{(1)}$ for the unelectroded part of the resonator or $\hat{C}^{(1)}$ for the electroded part. $\overline{C}^{(1)}$ is the stiffened elastic constant relative to the corresponding one dimensional mode. $\hat{C}^{(1)} = \overline{C}^{(1)}(1-2\hat{R}-8k^{(1)2}/n^2\pi^2)$ is a constant that includes the electrical and mechanical effects of the metallization. $(4k^{(1)2}/n^2\pi^2)$ is the relative frequency lowering due to the electrical effect of the metallization (thickness 2x2h'); $\hat{R} = 2.2h''/2h'$ is the relative frequency lowering due to the inertia of the electrodes. $C^{(1)} = \overline{C}^{(1)}(1-k^2)$ is a pseudo-ordinary elastic constant. $k^{(1)}$ is the coupling coefficient of the corresponding one-dimensional mode.M'_n and P'_n are intricate functions of the material constants and of the plate orientation [8].

We consider un-spherical contours having lateral thickness variations such that their two principal curvatures are in the same ratio as the two first coefficients of the equation: The lateral variations of the thickness can then be approximated by the expression [1]

$$\left(\frac{1}{2h}\right)^2 = \left(\frac{1}{2h_0}\right)^2 \left[1 + \frac{x_1^2}{2R'_1 h_0} + \frac{x_3^2}{2R'_3 h_0}\right] \quad \{2\}$$

with $R'_1/R'_3=M'_n/P'_n$. The contours represented by this relation are surfaces of the fourth degree that can be approximated by quadrics such as elliptical paraboloids or ellipsoids. Their intersections with planes normal to x_2 are ellipses, the intersections of two surfaces of this kind (with different 2ho and R') are also ellipses, so we can consider un-spherical resonators of different kinds (plano-convexe, bevelled, and devices with two "curvatures").

The eigen mode at V=0 (short-circuit modes) are the solutions of the homogeneous form of the equation {1}. For the eigen modes: $\hat{u}_1 \equiv u_1$. Some algebra allows to introduce characteristic (cut-off) frequencies in the equations:

$$M'_n \tilde{u}^n_{1,11} + P'_n \tilde{u}^n_{1,33} + 4\pi^2\rho\left[f^2 - f^{*2}\left(1+\frac{x_1^2}{2R'_1 h_0}+\frac{x_3^2}{2R'_3 h_0}\right)\right]\tilde{u}^n_1 = 0$$

Where: $f^* = \hat{f}^n = \frac{n}{4h_0}\sqrt{\frac{\hat{c}^{(1)}}{\rho}}$ (electroded regions),

or $f^* = \bar{f}^n = \frac{n}{4h_0}\sqrt{\frac{\bar{c}^{(1)}}{\rho}}$ (unelectroded regions).

We define the quantity: $R_{ns}' = \left(\frac{C^*}{M_n}\right)R_1 = \left(\frac{C^*}{P_n}\right)R_3$

The coordinate transformations:

$$\begin{cases} X_1 = \left(\frac{C^*}{M'_n}\right)^{1/4}\left(\frac{C^*}{M'_n}\right)^{1/4}\gamma_{ns}.x_1 \\ X_3 = \left(\frac{C^*}{P'_n}\right)^{1/4}\left(\frac{C^*}{P'_n}\right)^{1/4}\gamma_{ns}.x_3 \end{cases} \quad \text{with } \gamma_{ns}^4 = \frac{n^2\pi^2}{2R'_{ns}h^3} \qquad \{3\}$$

lead to an isotropic separable form of the equation:

$$\begin{cases} \tilde{u}^n_{1,rr} + \frac{1}{r}\tilde{u}^n_{1,r} + \frac{1}{r^2}\tilde{u}^n_{1,tt} + \left(A^* - \frac{r^2}{2}\right)\tilde{u}^n_1 = 0 \\ \text{with } A^* = \frac{n\pi}{2}\left(\frac{R'_{ns}}{2h_0}\right)^{1/2}\left(\frac{f^2 - f^{*2}}{f^{*2}}\right) \end{cases} \qquad \{4\}$$

In this transformation, the un-spherical contour is transformed into a revolution surface. As previously noticed the present formulation is directly a very good approximation for the cases with spherical curvatures when the lateral anisotropy of the plate is small (Mn'# Pn'). This is the case for the 3rd overtone of the AT cut and for the third and the fifth overtone of the SC cut. Equation {4} can be separated as $\tilde{u}_1(r,t) = R(r)T(t)$ into two ordinary differential equations:

$$\begin{cases} R'' + \frac{1}{r}R' + R\left(A^* - \frac{r^2}{4} - \frac{\mu^2}{r^2}\right) = 0 \\ T'' + \mu^2 T = 0 \end{cases} \qquad \{5\}$$

The separation constant must be an integer $\mu=m$, since the solution of the first equation has to be periodic with a period $2m\pi$. The transformation R(r)=(1/r) W(r2/2) leads to the Whittaker equation:

$$W'' + \left(-\frac{1}{4} + \frac{A^*/2}{r^2/2} + \frac{1/4 - m^2/4}{(r^2/2)^2}\right)W = 0. \qquad \{6\}$$

Among the solutions of this equation [11-12] the Whittaker $M_{A^*/2,m/2}(r^2/2)$, and $W_{A^*/2,m/2}$ $(r^2/2)$ functions of index $k=A^*/2$, and $\mu=\pm m/2$ are the most interesting ones for our purposes. The M functions are bounded at zero and the W functions tend towards zero when the argument becomes infinite. They are very appropriate respectively for the internal region (centre) and for the most external region in the case of "infinite" plate condition. The most general symmetrical solutions (with respect to the x1 and to the x3 axis) bounded at r=0, is constituted of a linear combination of the successive order symmetrical solution of equation {4}. It can be decomposed in two sub-expressions: a general solution bounded at zero and a general symmetrical solution bounded at r infinite:

$$\begin{aligned} \tilde{u}_1 &= \sum_m B_m\left(\frac{1}{r}\right)M_{A^*/2,m/2}\left(\frac{r^2}{2}\right).\cos mt \\ \tilde{u}_1 &= \sum_m C_m\left(\frac{1}{r}\right)W_{A^*/2,m/2}\left(\frac{r^2}{2}\right).\cos mt \end{aligned} \qquad \{7\}$$

where m is an even integer.

On the lateral surface limiting regions with different curvatures or limiting the electroded and the un-electroded regions we have to specify the continuity of u_1 and of its normal derivative (approximation of the continuity of $T_{ij}n_j$). On the external edges of the resonator, the solution has to verify a traction free condition or a zero displacement condition.

I.2 Electroded plate with elliptical shape (free or fixed edge).

In this simplest geometry (isotropic and circular in the r,t coordinates), for a totally free or a totally clamped edge, the solution contain only one term of the series with the M_{km} functions. For each overtone rank the eigen modes are given by:

$$\tilde{u}_1^m = \sqrt{2}\left(\frac{1}{r}\right)M_{A^*/2,m/2}\left(\frac{r^2}{2}\right).\cos mt, \qquad \{8\}$$

with a value of A*/2 which makes that, respectively, either du_1/dr or u_1 is zero at the edge. For well designed resonator the corresponding value of A*/2, for each mode is close to half an integer. This latter value corresponds to the "infinite" case where both boundary conditions are simultaneously achieve at infinity. The fundamental mode or its overtones correspond to the function:

$$\tilde{u}_1^{(n)} = \sqrt{2}\left(\frac{1}{r}\right)M_{A^*/2,0}\left(\frac{r^2}{2}\right), \qquad \{9\}$$

which have a revolution symmetry in the r,t coordinate.

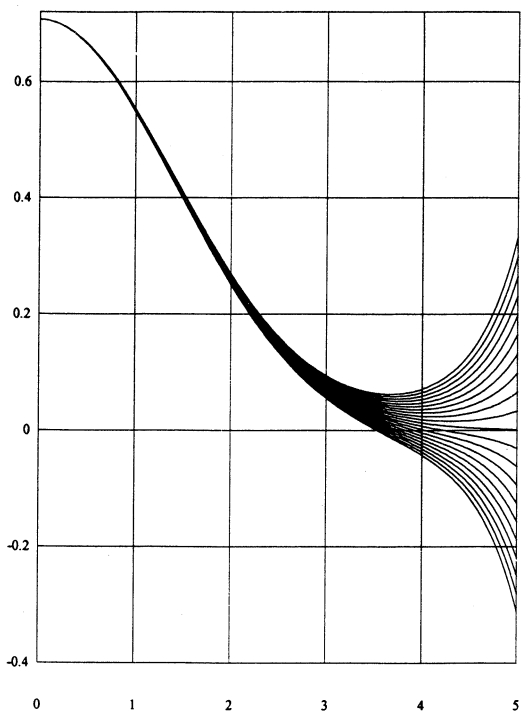


Figure 1: Mode shape for the fundamental mode of a resonator with 2ho=1.117mm, Rc1=41.mm, diameter along x1=15mm.

The variation of the mode shape with the reduced coordinate r are represented in figure 1 for the case of a fundamental mode resonator with R1=41mm, $2h_0$=1.117mm a total dimension along x1=15.mm. In this figure the A*/2 parameter was varied around the values found in the free edge case (0.499910) and the value for the fixed edge case (0.500072). The displacement at the edge (r#4.87) in the first case is 5 thousandth of the value at the centre of the plate. In figure 2, the mode shape near the edge has been enlarged.

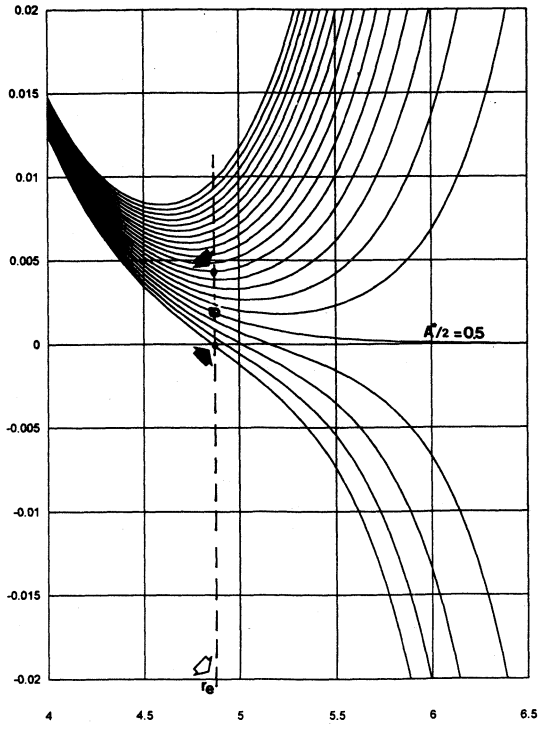


Figure 2: Detail of the mode shape for the fundamental mode resonator of figure 1.

For the third overtone of a resonator having the parameters: 2h=1.320mm, R1=200mm diameter along x1=15mm; the values of the A*/2 parameter are for the free case A*/2=.499996 and for the clamped case A*/2=.500008. The displacement at the edge (r#5.528) in the free case is of about 1/1000 of the value at the centre of the plate.

I.3. Un-spherical resonator fixed along two segments.

We have considered the case of un-spherical resonators in which a small fraction of the edge is fixed either in the x1 or in the x3 direction (1/24 or 1/32 or 1/36 of the perimeter for each fixation). A discretization of the boundary condition was used for the numerical solution (two consecutive points around the x_1 or the x_3 axis were fixed for each side of the corresponding axis). Then the resonance frequencies are slightly different of the previous ones but the most important fact is the non negligible contribution of several terms of the series (cosmt angular variations). The results are given in tables I and II. In these tables which concern mainly the un-spherical resonators totally electroded having the first geometry previously considered above (R1=41mm, 2h=1.117mm, $2ax_1$=15mm) we have included the case of an unspherical plate with a circular fixed edge. This shape which does not respect the lateral anisotropy induces also a non negligible contribution of the higher terms of the series.

For the three last cases, in table I, the higher order coefficients seem to be quite low but the values of the function can be very high so that the contribution can be non negligible at the edge of the plate. Due to the very great sensitivity of the Mkm functions to their parameters and argument the contributions can vary fastly with small variations of the design parameters. These contributions depend of the angle t and it is extremum for each term when cosmt is maximum (m=n*180°/m). So, the effect of the fixation can be to destroy the radial symmetry of the mode and to increase significantly the energy at the plate edge. We can observe in table II that, as known from the experimental practice, the effect seem to be more important when the mounting clips are situated in the x_1 direction.

Similar effects can be computed for the overtone mode resonators. But for resonators having design parameters near the usual values for such devices, these effects are weaker mostly due to the overtone rank dependence of A* and of the larger $R_{ns}/2h_0$ ratio existing in such resonators. In table III, we have represented the frequencies computed for the un-spherical overtone mode resonator previously considered (2ho=1.320mm, Rc1=200mm diameter along x1=15mm). The coefficients of the series are listed in table IV .

Table III: Resonance frequencies for the 3rd overtone resonator

Fr	El* Free	El. Fixed	El. X1fix	El . X3fix
(Hz)	3819268.4	3819269.2	3819268.6	3819268.6

(*) El=Elliptical Un-spherical plate totally electroded.

Table IV: Coefficients of the series for the third overtone resonators.

Coef.	A0	A2	A4	A6	A8	A10
Free	1.000	0 (3e-12)	0 (7e-15)	0 (-1e-14)	0 (-5e-16)	0 (1e-16)
Fixed	1.000	0 (2e-12)	0 (-2e-14)	0 (2.e-14)	0 (4e-15)	0 (-1e-16)
X1 clips.	1.000	-.133 e-5	-.117 e-5	-.847 e-8	-.463 e-9	-.158 e-10
X3 clips.	1.000	.122 e-5	-.111 e-6	.839 e-8	-.506 e-9	.234 e-10

*for the two first lines the A2,A4,...A10 coef. are also theoretically 0; the values found by the numerical model are given ().

On figure 3 we have represented the lateral variations of the mode shape for the cases considered in table I and II. Figure 3a is relative to the totally fixed edge, figure 3a is relative to the totally free edge, figure 3c consider the case where the

TABLE I: Resonance frequencies for non spherical contours

frequency	El. Free	El.Fixed	El.X1fix	Ro.fixed
(Hz)	1621294.	1621338	1621303.	1621324.

(*) El=Elliptical plate totally electroded Ro=Round plate totally electroded

TABLE II: Coefficients of the series for the fundamental mode resonator.

Coef*	A0	A2	A4	A6	A8	A10
Free	1.000	0 (7e-13)	0 (-2e-12)	0 (-4e-13)	0 (-6e-14)	0 (-2e-16)
Fixed	1.000	0 (7e-13)	0 (1e-12)	0 (3e-13)	0 (7e-14)	0 (4e-16)
X1 clips	1.000	-.367 e-4	-.320 e-5	-.219 e-6	-.105 e-7	-.274 e-9
X3 clips	1.000	+.257 e-4	-.246 e-5	+.234 e-5	-.114 e-6	+.189 e-8
Ro fixed	1.000	-.175 e-4	-.910 e-6	-.291 e-7	-.596 e-9	-.714 e-11

*for the two first lines the A2,A4,...A10 coef. are theoretically 0; the values found by the numerical model are displayed. They give an order of magnitude of the errors in the numerical computations of the coefficients.

elliptical plate is fixed along two segments in the x_1 direction while figure 3d is relative to the case of a round plate with an un-spherical contour fixed along to segments in the x_1 direction. In figure 3d we can observe that the displacement becomes negative near the regions where the edge is free, thus nodal lines joining the mounting segments appears in the mode shape. In figures 3b,c,d the displacement is multiplied by 10 to better observe its value near the edge.

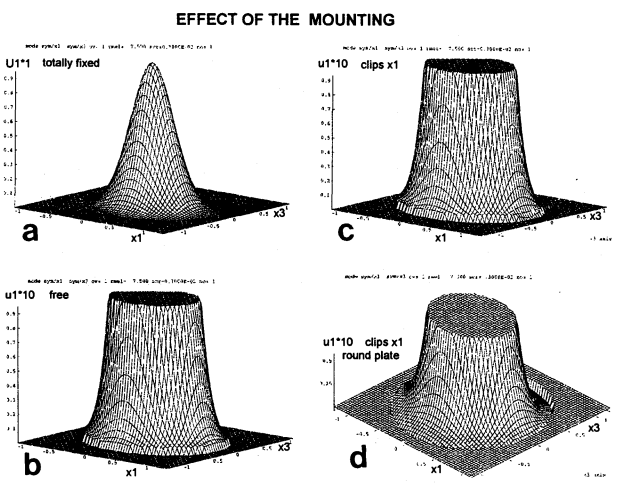


Figure 3: Computed mode shape of unspherical resonators

II RESONATORS WITH SPHERICAL CONTOURS.

We have previously shown [1-3] that the un-spherical case considered constitutes already a good approximation of the spherical case provided certain relations between the parameters of the resonators are satisfied. A model for devices with spherical contour was derived from the solutions shown in paragraph I. In this model we divide the resonator in regions (characterised by a curvature, an electrode, etc.) and then each region in sectors. We approximate each sector with spherical curvature by a sector of an un-spherical resonator characterised by R_{ns} and $2h_0$ parameters such that the curvature and the thickness are the same at the centre of the plate and that the thickness is also exact at each limit of regions (electroded region, bevel, edge, etc..). Fundamentally, in this approximation the spherical curvature is approximated, in the limiting case of an infinite number of sectors, by an un-spherical contour having the R_{ns} (and in some cases the 2ho) parameters continuously varying with the polar angle so as it is spherical in the usual coordinates. For example in the sectors of the electroded region of plano-convexe resonators we approximate the spherical curvature by the relations:

$$\left(\frac{1}{2h_0}\right)_{centre} = \left(\frac{1}{2h'_0}\right)_{centre} \quad et$$

$$\left(\frac{1}{2h'_b}\right)_{edge} = \left(\frac{1}{2h'_0}\right)\left(1+\frac{x^2_{1edge}C^*}{M'_n R_{ns}.2h'_0}+\frac{x^2_{3edge}C^*}{P'_n R_{ns}.2h'_0}\right)$$

where h'_b, $x_{1_{edge}}(\rho,\theta)$, $x_{3_{edge}}(\rho,\theta)$ are the values of the thickness and of the untransformed coordinates at the corners of the sector at edge of the plate. Discretization scheme involving other approximations are used for the external contoured regions to ensure a good representation of the curvature [1].

This numerical model was applied to resonators having geometries close to those previously considered in the case of the un-spherical contours. In order to compare with the previous case, an approximation considering for the developments "mean coefficients" between the different sectors was used. In the table V the results are given for a 3rd overtone mode resonator: (Plano-convexe Rc=200mm, 2ho=1.320mm, diameter 15mm, totally electroded). Six terms of the development using the M_{km} functions (with m=0 to 10) are used

TABLE V: Coefficients of the solutions.

Coef.	A0	A2	A4	A6	A8	A10
Free	1.0	-.639 e-2	.110 e-4	-.117 e-7	.129 e-12	.305 e-12
Fixed	1.0	-.645 e-2	.113 e-4	-.118 e-7	-.118 e-11	.369 e-12
X1 clips	1.0	-.641 e-2	.920 e-5	-.142 e-6	-.693 e-8	.226 e-9
X3 clips	1.0	-.641 e-2	.127 e-4	-.134 e-6	.702 e-8	-.305 e-9

In this case, the effect of the boundary condition, should also be evaluated using the "Free" or the "Fixed" cases as a reference since then the corresponding "mean" coefficients take into account a different anisotropy of the mode shape. Again the effect of fixing only a small part of the edge is to increase (algebraically) the coefficients. We can observe here that the modification is much greater for the highest order terms of the series, so that the effect of the fixation is mainly to add terms with very fast angular variations. Here the effect of a different plate geometry, is to modify very slightly the mode shape.

The computed resonance frequencies for the different cases are listed in table VI. We observe the same phenomena as for the un-spherical cases for the effect of a "partial" fixation of the resonator.

TABLE VI: Resonance frequencies for spherical contours

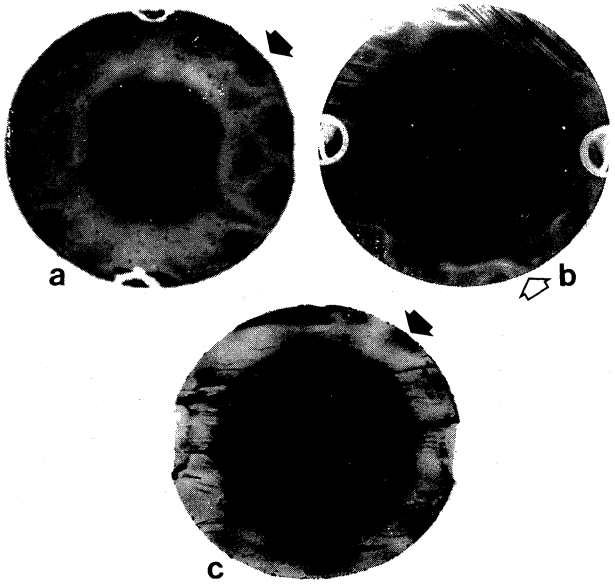
Fr	R. Free	R. Fixed	R.X1clip	R.X3clip
(Hz)	3822662.0	3822662.5	3822664.3	3822659.8

The comparison between the spherical and the un-spherical case indicates that the latter provides qualitatively and quantitatively correct predictions of the phenomena resulting of the various types of boundary conditions.

III X-RAY TOPOGRAPHY OBSERVATIONS

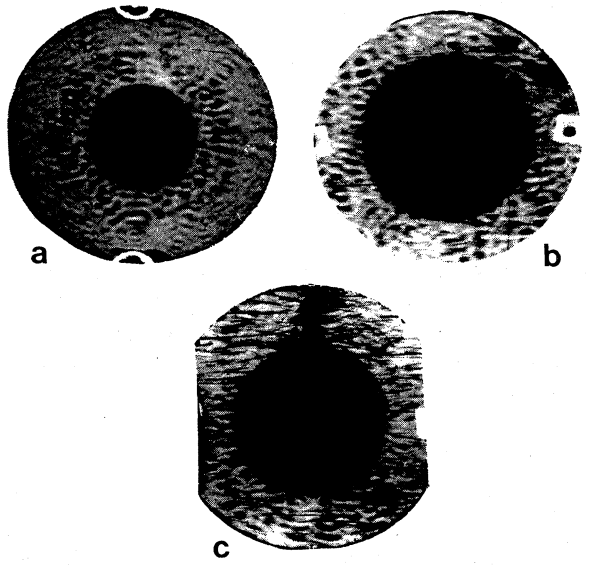
The mode shapes of resonators with parameters similar to those considered in the previous paragraphs were imaged by x-ray topography. The expected deformations of the mode were observed with the appropriate angular variations. Such a case can be seen in topography 1a. which concerns a fundamental mode resonator very similar to those considered in the previous § but with electrodes smaller than the plate (Rc=41mm, electrode diameter 8mm plate diameter 15mm, mounting clips in the x1 direction). The maximum value of the displacement at the edge exceed 1% of that at the centre (the topography was over-exposed to better observe the small amplitude contributions). The contributions with angular variations in cosmt have larger amplitudes than expected from the calculation, the most important ones being those in cos2t, cos4t and cos6t which are found (except near the mounting clips) to be extremal at the expected angles and nodal lines appears between the extrema which possess alternate signs.. It is much probable that the large values of their amplitudes result of properties of the mounting not taken into account in the calculations. The topography 1b is relative to a fundamental mode bevelled resonator which displays, in the (external) contoured region, another phenomena expected from the calculation; the apparition of nodal lines. They can be observed near the free edge of the plate but in this case, their geometry is disturbed by a strong interaction of the mode with the Growth bands existing in the plate. The topography of figure 1c

represents a fundamental mode bi-convexe resonator. Here again contributions with fast angular variations are observed near the edge of the plate. we can observe that the mode shape is disturbed by interactions with dislocations.



Topographies 1a,b,c: Fundamental mode resonators.

The mode shapes of the third overtone of plano-convexe resonators with one or two curvatures, have often shown some extra-features which can be fundamentally of same nature (related to the fixations of the plate or to the existence of the external boundary or of an internal one separating two zones) but, eventually, in the frame of a more general theory. In facts, in each case, the small u1 component with fast lateral variations which extend over the whole plate was found to be very dependant of the fixation conditions and it has disappeared or was greatly modified when the type of mounting clips was changed. In some instances a weak u3 component with similar and related variations was observed. Such observations which may also be related to couplings with plate modes are presently the object of in depth investigations. Topographies 2 display the observed phenomena in different contoured resonators.



Topographies 2a,b,c: Third overtone resonators

CONCLUSION.

The theoretical study of the non-spherical resonators with curvatures and a lateral geometry respecting the lateral anisotropy of the crystal plate has shown new features which also exist in all contoured resonators having a spherical contour. The modes of contoured resonators with a totally free lateral boundary can have a non negligible amplitude at the edge of the plate. This amplitude decrease with overtone rank in a manner well described by the properties of the index A* of the Whittaker functions, For the anharmonics of these overtones, this amplitude increase with their rank.

When the plate is fixed on small portions of its perimeter, higher terms of the series with fast angular variations are included in the displacement in the non-spherical case (and also in the spherical case) so that the vibration amplitude can vary (increase) near the edge of the plate and acquire fast angular variations. The phenomena is more intense for the fundamental mode than for the overtones but it exists in any cases. The effect of the plate geometry on the mode shape can be non-negligible. Particularly, the calculation have shown a noticeable effect for the case of plate geometry (circular) which leads to the apparition of nodal lines. In this case, the higher order coefficients of the series are also increased. The experiments have confirmed the existence of such effects. They have also shown that other (or larger) contributions to the real vibration modes resulting from the actual boundary conditions corresponding to the mounting, can be observed.

We believe that all the effects observed here does have consequences on the ultimate frequency stability that can be achieved with the contoured crystal resonators and also on all other properties which depend strongly on the mode shape. Further investigations and computations are being made to precise the influence of several factors on the mode shape and on the most important properties of such resonators.

REFERENCES:

[1] J.Détaint, H.Carru, J.Schwartzel, B.Capelle, A.Zarka; Proc. of the 42nd Annual Frequency Control Symposium pp.19-28 (1988).
[2] P.C.Y. Lee, Ji. Wang. Proc. of the 1995 IEEE Int. Frequency Control Symposium. p.705 (1995).
[3] J.Détaint, B.Capelle, A.Zarka, Y.Zheng. Proc. of the 1995 IEEE Int. Frequency Control Symposium p.716 (1995).
[4] H.F.Tiersten, B.J.Lwo, B.Dulmet. Proc. of the 1994 IEEE Int. Frequency Control Symposium pp.172-183 (1994)
[5] J.S.Yang, H.F.Tiersten. Proc. 1995 Int. IEEE Frequency Control Symposium p.727 (1995).
[6] J.Détaint, A.Zarka, B.Capelle. to appear in Proc. 1996 Int. IEEE Frequency Control Symposium (1996)
[7] J.S.Yang, H.F.Tiersten. to appear in Proc. 1996 Int. IEEE Frequency Control Symposium (1996)
[8] D.S. Stevens , H.F. Tiersten. J. Acoust. Soc. Am. Vol. 79 nø 6 p.1826-1811 (1986).
[9] R.D. Mindlin. Int. J. Solids Structures v.21 n°6 p.597-607 (1985).
[10] P.C.Y. Lee, S.Syngellakis, J.P.Hou. J. appl.phys. v.61 n°4 p1249-1262 (1987).
[11] M. Abramowitz; I. Stegun. Handbook of Mathematical functions. Dover (1972)
[12] E.T. Whittaker; G.N. Watson. A Course of Modern Analysis. Cambridge Press 1927.

FEM MODELING OF SPECTRUM AND HIGH ORDER THERMAL SENSITIVITIES OF FLAT PLATES RESONATORS OPERATING IN FUNDAMENTAL MODE OF ESSENTIALLY THICKNESS SHEAR

*F. RAOELIJAONA, *B. DULMET, °J. T. STEWART, °D. S. STEVENS

*ENSMM/ LCEP 26 Chemin de l'Epitaphe F 25030 Besançon Cedex, France
°Vectron Technologies Inc, 267 Lowell Road, Hudson, NH 03051-0917, USA

ABSTRACT

A new finite element software has been developed for the modeling of plate resonators operating in fundamental mode of essentially thickness shear. Investigated designs consist of parallelepipedic thin plates, in the vicinity of AT-cut, with thin metallic electrodes deposited on facing-together surfaces. The plates have various length-to-width and width-to-thickness ratios. We consider that the vibration is trapped by mass loading effect along one in-plane axis (*l*) of the crystal, while trapping effect is absent or poor along the second in-plane axis (*w*). FEM provides with an efficient means to model dimensions-dependent couplings between thickness-shear and various other modes, mainly high overtones of flexure, which arise from the boundary conditions at free edges. In this paper, after recalling the main features of our program, we particularly investigate the effect of *w/t* ratio (*t*=thickness) on the frequency spectrum and high order temperature behavior for values in the range [10–15] of *w/t*.

Keywords : strip resonator, finite element, quartz, temperature sensitivities.

1. INTRODUCTION

Strip resonators have been a subject of constant interest since their introduction [1]. They are suitable for low cost, mass production resonators and can operate at a much higher frequency than flexure mode resonators, although their dimensions are quite comparable, i.e. small in compare with trapped energy square or circular plates operating in quasi pure thickness shear mode. Since early works from Mindlin [2], it has been known that the resonant modes occuring in narrow rectangular AT strip quartz resonators exhibit definite mechanical couplings between thickness shear mode, still dominant, and flexural and other modes such as face shear or extensional. This couplings, which strongly depend on the *w/t* ratio, determine a marked oscillating behavior of the U1 and U2 components of mechanical displacement, respectively corresponding to fundamental overtone of slow shear and flexure waves propagating along the width. The vibration of strip resonators is trapped along the larger dimension of crystal (*l*), with help of piezoelectric coupling and mass loading effect of electrodes. Although the analytical modeling of such resonators can be refined [3], it becomes then intricate, and the accuracy of prediction may seem insufficient, especially for wide *w/t* values, involving coupling of TS with high overtone of flexure, a technically interesting configuration since it allows for better C1/C0 ratios than very narrow strips. So, starting from a program existing at LCEP, and initially designed for the modeling of eigenmodes in contoured resonators, we have redesigned a new program with suitable features for the analysis of resonant modes in not-so-narrow strip resonators (plates, in fact), including the prediction of static frequency-temperature characteristics. Due to the weaker oscillating behavior along the thickness, in compare with high overtone thickness shear resonators, and the relatively simple shape of crystals, it is possible to consider really dense discretisations along in-plane directions, letting expect a good accuracy while the generated problem remains quite treatable with nowadays available hardware.

2. THEORETICAL POINT OF VIEW

The reader can find more extensive details on topics considered in this paragraph at Ref [4]. Key points considered on building a FEM program were the choice of a pertinent variational formulation, the matrix form of this formulation, and discretisation techniques (including interpolating functions and element geometry which strongly affect efficiency). We just recall the main equations and we shortly mention some details about the two new elements built specifically to study the considered resonators.

2.1 Variational principle in lagrangian formulation

$$\int_{t_o}^{t} dt \left[\int_{V_o} \left(K_{L\alpha} \delta u_{\alpha,L} + D_L \delta \phi_{,L} + \rho_o \omega_\alpha^2 \delta u_\alpha \right) dV_o - \int_{S_o} \left(F_\alpha \delta u_\alpha + \sigma \delta \phi \right) dS_o \right] = 0 \qquad (1)$$

- **K** is the Piola Kirchoff stress tensor,
- **D** is the material electric displacement,
- ρ_O, V_O, S_O denote the fixed mass density, volume and outer surface of the crystal in the stress free natural state (reference temperature 25°C),
- **F** is an applied static force,
- σ is the net charge imposed on non metallized surfaces.

2.2 Matrix form of variational formulation

$$\int_V \left(\langle \delta U' \rangle [G] \{U'\} + \langle \delta U' \rangle [R] \{\phi'\} \right) dV + \int_V \left(\langle \delta \phi' \rangle [R] \{U'\} - \langle \delta \phi' \rangle [N] \{\phi'\} \right) dV + \int_V \rho_O \omega^2 \langle \delta U \rangle \{U\} dV = O \qquad (2)$$

where G, R, N respectively denote the sets of so-called effective material elastic, piezoelectric and dielectric constants refered to the natural state. These constants, which have special symmetries and are respectively stored in 9x9, 3x9 and 3x3 matrixes, were already defined and discussed in previous papers [5-6].

2.3 Discretization of the problem

The quadratic interpolations associated with well known 20-nodes or 32-nodes hexaedric elements are not optimal to fit the oscillating behavior of the normal displacement along the width. Since the trigonometric interpolations proposed in previous

papers [4] were more specifically designed for high overtone resonators, we decided to design two kinds of new 48-nodes hexaedric elements with polynomial interpolations. The first kind uses cubic interpolations in every direction, while the second one uses cubic interpolations along the thickness OY and the width OX, and quadratic ones along the length OZ. We give on Fig. 1 the distribution of the nodes in each element (For legibility, we just show the nodes along the base plane XOZ and the thickness).

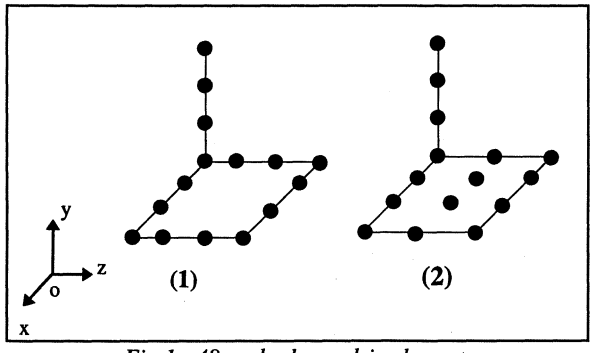


Fig 1 : 48-nodes hexaedric elements

2.4 Effect of the electrodes

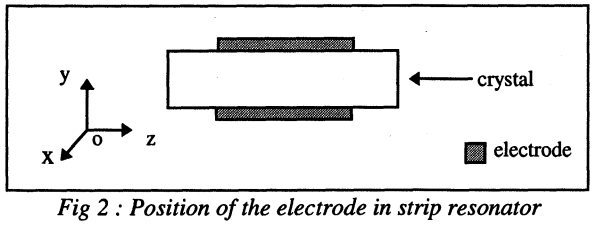


Fig 2 : Position of the electrode in strip resonator

Our software takes into account the elastic, dielectric and inertial effects induced by thin metal electrodes deposited on facing-together surfaces of the resonator. This is obtained through adding the following term into the time integral of the variational formulation in Eq(1):

$$h_e \int_{S_e} \left(\rho_e \omega^2 u_\alpha \delta u_\alpha + T_{L\alpha} \delta u_{\alpha,L}\right) dS$$

- h_e: thickness of the electrode
- ρ_e : mass density of electrode material
- S_e: surface of the electrode

This expression includes both the mass loading and the stiffness effects. We use shape functions expressed in terms of the parametric coordinates *(r,s)* to facilitate the computation of integrals, thereby requiring the computation of the relevant Jacobian J of transformation from real to parametric coordinates.

mass loading terms

$$[M^e] = h_e \int_{S_e} \rho_e [H(r,s)]^t [H(r,s)] \, det(J) ds$$

$$[H] = \begin{bmatrix} H^1 & o & o & o & \dots \\ o & H^1 & o & o & \\ o & o & H^1 & o & \end{bmatrix}_{(3X(4Xk))}$$

stiffness terms

$$[K^e] = h_e \int_{S_e} [B(r,s)]^t [C^e] [B(r,s)] det(J) ds$$

$$[B] = \begin{bmatrix} H^1_{,1} & o & o & o & \dots \\ o & H^1_{,2} & o & o & \dots \\ o & o & H^1_{,3} & o & \dots \\ o & H^1_{,3} & o & o & \\ o & o & H^1_{,1} & o & \\ H^1_{,2} & o & o & o & \\ o & o & H^1_{,2} & o & \\ H^1_{,3} & o & o & o & \\ o & H^1_{,1} & o & o & \end{bmatrix}_{(9X(4Xk))}$$

$$[C^e] = \frac{E}{(1-\nu^2)} \begin{bmatrix} 1 & o & \nu & o & o & o & o & o & o \\ & o & o & o & o & o & o & o & o \\ & & 1 & o & o & o & o & o & o \\ & & & o & o & o & o & o & o \\ & & & & & o & o & o & \gamma & o \\ & & & & & & o & o & o & o \\ & & & & & & & o & o & o \\ & & & & & & & & \gamma & o \\ & & & & & & & & & o \end{bmatrix}$$

where E and ν respectively denote Young's modulus and Poisson's ratio of the supposedly isotropic electrode. The above expression for C^e assumes that the thickness is along the OY axis. The size of this matrix is 9X9 in order to follow the adopted formulation for the stiffness matrix of the volumic elements meshing the piezoelectric crystal [4].

2.5 Temperature sensitivities

Since the writing of the variational principle is mapped onto the coordinates of the reference state (lagrangian formalism), the geometry and mass density remain fixed along temperature variations. We consider that only the elastic coefficients are affected by temperature, although the program would require very little changes to take into account thermal sensitivities of piezoelectric and dielectric constants. The temperature derivatives of the G_{LM} elastic constants can be either determined from frequency-temperature characteristics associated with analytic models for the vibration, or they can be deduced from existing values of the temperature derivatives of elastic constants in the « usual » sense (BBL) through formulas that only involve thermal expansions, and are given at Ref [6]. Two methods have been adopted to compute the frequency-temperature characteristics.

- In the first one, we perform successive runs upon updating the computation of all elementary stiffness' for each temperature step. This is what we call the direct method.
- In the second one, we adopt a perturbation process considering that the results of a first preliminary run yields a subset of mode patterns which can be used as an initial assumption for the description of eigenmodes when temperature varies. The variations of stiffness coming from temperature variations (up to the third order) are considered as additional perturbation terms in the variational formulation. Since the dynamic eigenvalue/eigenvector problem is solved with subspace method, this is possible as long as the vectors in the subspace at current temperature do not become orthogonal to the initial subset. Also, the initial subset must be large enough, since only its elements can contribute to the thermal sensitivity of the whole problem.

3. DESCRIPTION OF THE SOFTWARE

Our program has the usual structure encountered in finite elements programs. All routines are written in ANSI C language, with use of *typedef* prototyping, checking of passing parameters to functions, and dynamic memory allocation. The program compiles and executes well with Sparcworks add-on *acc* compiler under SunOS-4.1.3 and with FSF-GNU *gcc* compiler release 2.7.2 under SunOS-4.1.3 and Solaris 2.5 on Sparc-based systems, and under Linux 1.2.13 and 2.0.0. on Intel Pentium systems. The package was compiled with Visual C++, installed and tested at Vectron Technologies Inc. on a DEC-Alpha computer running Windows NT4.0 . Independent external modules are devoted to the pre- and post-processing.

3.1 Mesh generation (pre-processing)

According to the specific design of the parallelepipedic resonators made by VTI, a specific module was developed to automatically mesh such structure, while accepting sufficient construction parameters. This program provides an ASCII file containing all data in the suitable format for our FEM program, *e.g.* :

- the geometric characteristics of the resonator (length, width, electrode length, electrode width, position of the electrode, thickness, electrode thickness, position of the electrodes),
- a choice between meshing all of the resonator or the half of it, according to symmetry considerations,
- the kind of element retained to mesh the structure (20 nodes, 32 nodes, or either kind of 48 nodes element), the number of the elements along in-plane axes (OX and OZ) and the number of layers of elements along the thickness,
- the physical characteristics of the crystals : temperature, cut-angles (Φ, Θ, Ψ),

The results of the mesh can be visualized with the graphic tools of MAPLE.

3.2 Organization of FEM program

The software allows for three kinds of analyses : frequency computation, thermal sensitivity with direct method and thermal sensitivity with perturbation method.
The first task performed, after reading the input file which must be properly formatted, consists in elaborating a direct access binary file containing all connectivity information. It is written after the nodes are renumbered with the Cuthill and Mac-Kee algorithm, which was found to typically yield a decrease of the initial bandwidth of the global matrixes by a factor 1/3.
Since the global stiffness and mass matrixes are symmetric, it is only required to store the upper triangular part of them. In this way the memory requirements for the storage of the global stiffness equations is approximately halved. A further lowering of the memory requirements is gained upon adopting a *skyline storage*: it consists of only storing in memory the part of each matrix column which is located between the diagonal and the farthest non zero element. Thus, for the cases that we considered in practice, the memory requirements are roughly reduced to one tenth of the size of the initial full matrix.
After assembling elementary stiffness and mass matrices into K and M we solve the eigenvalues problem $KU=\lambda MU$ with the *subspace iteration method*, which is suitable for the treatment of large problems, for which the investigated frequency window is not necessarily located at the lower or higher end of the eigenvalues spectrum.

3.3 Vizualisation (post-processing)

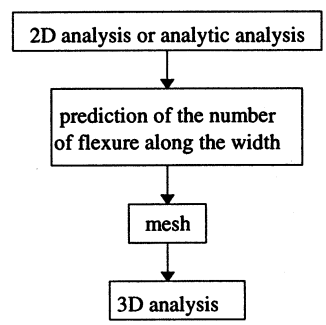
In order to conveniently identify the modes of interest, a post-treatement was designed to visualize the results. First an independent C-language module generates data files from the finite elements calculation. Those data are pre-formatted for subsequent visualization with help of popular graphical programs (Gnuplot or Maple).

4. NUMERICAL RESULTS

In this part we present some analyzes performed for the well-known patented 4.2MHz Zumsteg resonator and for a 10MHz resonator designed at VTI.

4.1 Process of resolution

We start from the data provided by a simpler program which computes the dispersion curves according to Ref [7]. This allows for an estimation of the number of flexure wavelengths along the width X, thereby predicting the minimum amount of elements in this direction. A 2D finite element software is also available at VTI which is helpful to predict the appropriated mesh for 3D analysis while requiring less memory than our 3D program. The following organigram summarizes the process of resolution.



4.2 Narrow strip resonator 4.2MHz

In this section, we investigate a well-known 4.2MHz strip resonator design. The test configuration is exactly that one at Ref [1]. In this design, the electrode covers the whole width of the resonator. Dimensions are shown on Fig 3.

4.2.1 Comparison of different elements

We have performed some analyses in order to compare the efficiency of the different hexaedric elements available in our software, mainly the 32 nodes (cubic along thickness, parabolic otherwise) and the two kinds of 48-nodes elements. In all cases, we have meshed the structure with 8 elements along OX, 12 elements along OZ, 1 layer along the thickness. The following table recapitulates the results of these computations :

Kind of element	Nb. of the nodes	Frequency (Hz)
32	1316	4.21170653e6
48 (kind 1)	2164	4.20357006e6
48 (kind 2)	2500	4.20353101e6

Those results show a frequency decrease when the number of nodes increases. Fig. 4 shows the pattern of U1 in the base plane of the resonator Since it has the same appearance for all tested elements, we just show it for the 48-hexaedric elements. For the visualization, we have cut the structure along OZ in order to see the behavior along OX (width of the resonator). The behavior along OZ is symmetric in this cut. The strong coupling between thickness shear U1 and flexure waves U2 propagating along X1 is readily seen.

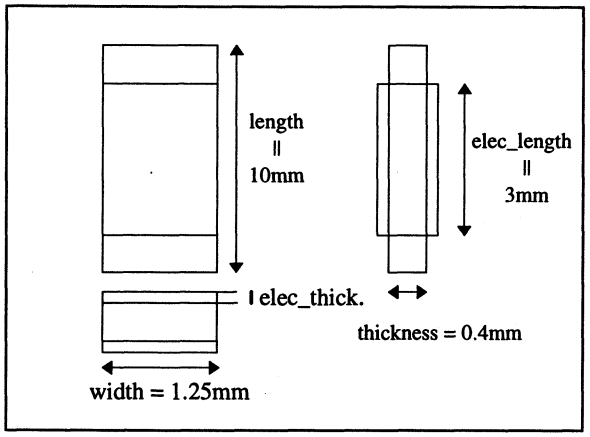


Fig 3 : 4.2MHz AT-Cut resonator

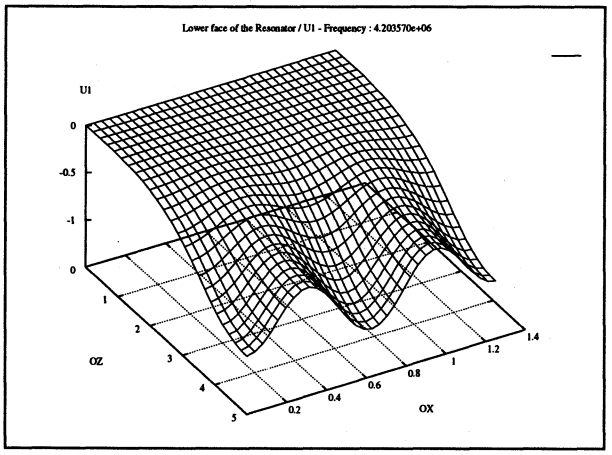


Fig 4 : Distribution of U1 in the lower surface

4.2.2 Temperature sensitivities

The thermal sensitivity has been investigated in the abovementionned configuration. We have tested both available methods (direct and perturbation) for each kind of elements (32, 48(1), 48(2)). All analyses are performed in the range of temperature [10°C - 180°C]. The whole set of obtained curves are given on Figs 5 and 6.

The change from 32-nodes to 48-nodes elements has significant impact on the frequency temperature characteristics. The curve labelled «half_32.direct» corresponds to results obtained when we only mesh the half of the strip resonator (cut along the width). Also, the predicted high order temperature sensitivities with the direct method are neatly different from the ones obtained with the perturbation method, which are close to sensitivities predicted for unidimensional infinite plates. We attribute the following reason to this behavior : to be accurate, the perturbation method requires the initial subset to act as a *complete* basis, thereby being large enough. So, we think that the number of asked modes should be *strongly* increased for the perturbation method to become accurate enough to account for the contribution of flexure coupling onto the global sensitivity. Nevertheless, assuming that this guess is correct, we have not yet investigated how (expectedly) increased accuracy associated to a larger subspace might be balanced by a reduction of the speed advantage of perturbation in compare with direct method. As long as this issue is not settled, we prefer to stick to the direct method.

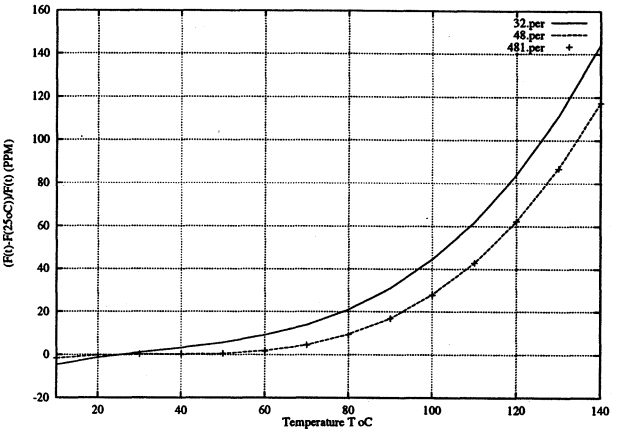


Fig 5 : Perturbation method - 4.2MHz

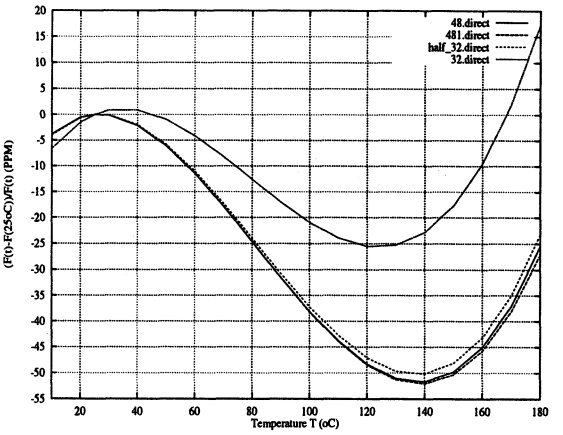


Fig 6 : Direct method - 4.2MHz

4.3 Wide strip resonator 10MHz (AT-Cut)

This configuration has been proposed by VTI and has been extensively investigated. Those investigations include the convergence checking, frequency temperature sensitivities and dispersion curves. In compare with 4.2Mhz design, this design requires a serious increase of computer resources. Corresponding analyses were performed at LCEP on a Ultra 1 Sparc workstation with 224 Mb of RAM and at VTI on a DEC Alpha workstation with 1Gb of RAM. We give in the following table the complete characteristics of the resonator, concerning both the crystal and the electrodes.

Crystal characteristics		Electrode characteristics	
length	5.588	Gold	
width	21 varies	ρ	18.5e-9
clamping zone	0.762	length	2.032
thickness	0.16346	width	1.7272
theta	35° 10' 0	thickness	2.804e-4
ρ	2.648e-9		

We mesh the structure with 48-hexaedric elements (12 nodes in the plane XOZ, 4 nodes along the thickness OY), we have 10 elements along the axis OZ, 1 layer of element along the thickness and 18 elements along OX

4.3.1 Dispersion curve

w	w/thickness	Frequency (Hz)
2.190364	13.4	10.1021305e6
2.223056	13.6	10.0882805e6
2.255748	13.8	10.0787017e6
2.288440	14.0	10.0709116e6
2.321132	14.2	10.0633815e6
2.353824	14.4	10.0538837e6

4.3.2 Shape mode

To give an insight of the mode patterns, Figs 7 and 8 present the distribution of U1 and U2 along the base plane of the resonator.

4.3.3Temperature sensitivities

Fig 9 shows the temperature sensitivities curves corresponding of the values of w/t varying between 13.8 and 14.4. (Temperature derivatives of elastic constants from Ref[5])

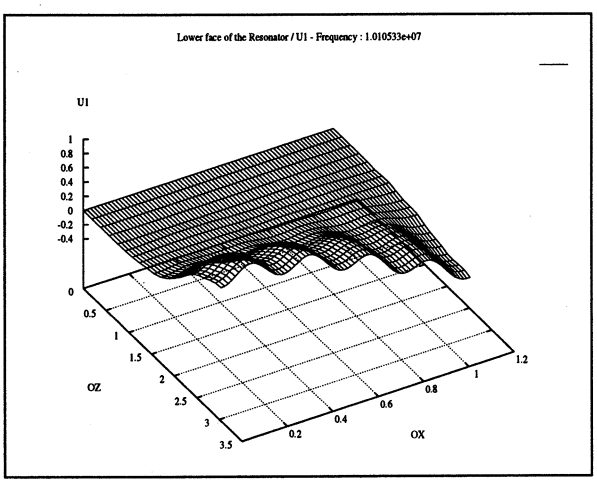


Fig 7 : Distribution of U1 along the base plane

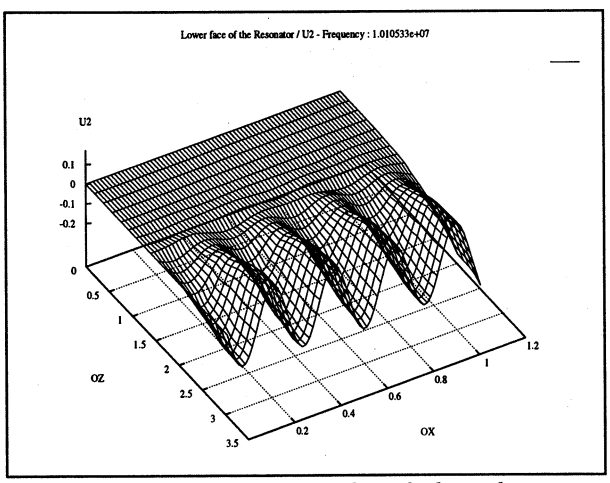


Fig 8 : Distribution of U2 along the base plane

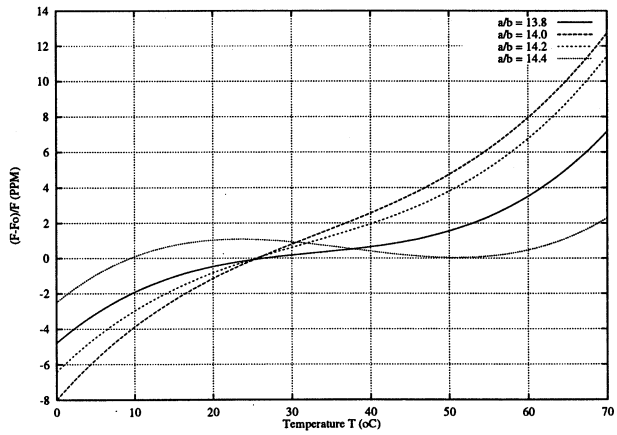


Fig 9 : Direct method - 10MHz

4.3.5 Analysis performed at VTI

The last tables recapitulate some analyses performed at VTI. They concern the configuration 10MHz with the ratio w/t = 14.4 and clearly illustrate the amount of memory required for such analyses (T^nG_{IJ} from James).

	Frequency (MHz)	
0	10.0509837	10.0538705
25	10.0509218	10.0538030
50	10.0508595	10.0537367
70	10.0508470	10.0537241
85	10.0508846	10.0537620
100		10.0538592
120		10.0541087
Duration	26h 46mn	3h 32mn
Mesh	2 layers 18 (OX) 10 (OZ)	1 layer 18 (OX) 10 (OZ)
RAM	750Mb.- 20modes	250Mb.- 10modes

Freq. (MHz)	Duration	RAM	Mesh
10.0537565	3h 50mn	600Mb. - 15modes	1 layer 18 (OX) 15 (OZ)
10.0538390	1h 30mn	346Mb. - 15 modes	1 layer 27 (OX) 10 (OZ)
10.053708	1h 40mn	358Mb. - 10 modes	1 layer 28 (OX) 10 (OZ)

References

[1] A. E. Zumsteg, P. Suda, Proc 30th A.S.F.C., pp 42-48, 1975.
[2] R. D. Mindlin,Journal of Applied Physics, Vol 22, N 3, pp 316-323, March 1951.
[3] R. F. Milsom, D. T. Elliott, S. Terry-Wood, M. Redwood, IEEE Trans SU, Vol 30, N3, pp 140-155, May 1983.
[4] F. Raoelijaona, B. Dulmet, Proc of 9th E.F.T.F., pp 121-126, March 1995.
[5] B. Dulmet, R. Bourquin, Proc of IEEE Ultrasonics Symp, pp 331-335, Nov 1994.
[6] B. Dulmet, R. Bourquin, Proc of 3rd E.F.T.F., pp 305-312, March 1989.
[7] R.D. Mindlin and D.C. Gazis, Proc of 4-th US National Congress of Applied mechanics, pp 305-310, 1962.
[8] K Trümpy, W. Zingg, Proc of 7th E.F.T.F., pp 45-50, March 1993.

ANALYSIS OF QUARTZ RESONATOR ACTIVITY DIPS WITH HIGH-DEGREE FEM

K.Trümpy

Micro Crystal, Div.of ETA SA, 2540 Grenchen, Switzerland

I. ABSTRACT

The coupling of quartz resonator main mode of vibration with spurious modes of vibration occuring at certain temperatures, causing the so called activity dips, is analysed with finite element method. High degree prismatic elements with up to septic Hermite interpolation were developed in order to obtain the high precision necessary to resolve the dense modal spectrum of AT-cut quartz resonators with fundamental frequencies of 16.8 and 50 MHz. Frequency dips resulting at maximum mode coupling are interpolated from values yielded at neighbouring temperatures. Motional resistances are calculated from the mechanical losses in the electrodes and in the bonding epoxy resin. The undamped dip is modified by means of the electrical equivalent circuit of the damped coupled double oscillator. Values for resonators of varying geometry are presented and compared to experimental data of AT-cut resonators.

KEYWORDS: High degree Hermite interpolation FEM, motional resistance, frequency dip, damped double oscillator, AT-cut resonator.

II. INTRODUCTION

A finite element program including triangular and rectangular prismatic elements with conforming cubic, quintic and septic Hermite interpolation is developed. Because in Hermite interpolation differentiated variables are used as node values, special measures are taken to allow non continuous gradients of the primary variables. The primary unknown variables are the extensive mechanical displacements and the intensive electrical potential. With displacement method for the mechanical variables and force method for the electrical variable, it is a mixed finite element approach. For quartz resonators with predominant mechanical energy, conforming interpolation still yields descending frequencies with increasing discretization level.

Except for the small electrode connections, the modelled resonators have a length symmetry plane and also for AT cut the anisotropic material constants are invariant against reflexion on this symmetry plane. Modelling only one half of the crystal structure by imposing boundary conditions of antisymmetry for nodes located on the symmetry plane gives the modes of interest at a considerably reduced caculation time. The resonator characteristics can be determined for varying temperatures. Electrodes are modelled by taking into account the mass, elasticity and electrical conductivity.

So far dissipation terms were neglected in the FE-model. In the sense of a perturbation calculus of first order we disregard the influence of dissipation on the eigenvector, because appreciable internal friction only occurs in a small part of the volume of the modelled resonator, i.e. in the electrodes and in the thin epoxy resin layer bonding tne resonator to the ceramic package. For given Q-factors of these materials, motional resistance of each mode can be calculated. Together with the frequency and the motional capacitance values of each mode, impedance and phase diagrams of the resonator can be determined.

Every mathematical mode of the dense modal spectrum creates at coincidence of frequencies a main mode frequency dip which is interpreted as the beat frequency of a coupled double oscillator. By means of the electrical equivalent circuit of the damped coupled double oscillator, where the coupling strength is determined from the value of the undamped frequency dip, a modified frequency dip and motional resistance value of the main mode is determined in function of the spurious mode motional resitance.

Due to the high discretization level, the finite element analysis leads to enormous caculation times. For a 50 MHz fundamental shear mode AT strip quartz resonator modelled with septic elements only few results are presented, however systematic results are presented for an analogous 16.8 MHz resonator [1] modelled with quintic elements. Some specific results of quartz crystal parameters are presented. The main part consists in the presentation of the temperature dependent spurious mode behaviour, which is systematically analyzed for different widths of the resonators modelled with varying thicknesses of the electrodes and of the epoxy resin layer. The theoretical values are compared with experimental data of chemically milled miniature AT-cut resonators.

III. HIGH-DEGREE FINITE ELEMENT ANALYSIS

Hermite interpolation leads to a much lower total number of degrees of freedom as compared to Lagrange interpolation, because at least for the rectangular prismatic element all nodal values can be concentrated at element corner nodes.

1. Basic FE-Equations of the Piezoelectric Model

The form function $\vec{N}$ links the primary variable nodal values in an unequivocal way to the values in the entire element volume. For the displacements u,v,w and the electrical potential ϕ, the relations

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} \vec{N}^T & 0 & 0 \\ 0 & \vec{N}^T & 0 \\ 0 & 0 & \vec{N}^T \end{pmatrix} \begin{pmatrix} \vec{s}_1 \\ \vec{s}_2 \\ \vec{s}_3 \end{pmatrix} = \mathbf{N}\,\vec{s} \quad ; \qquad \phi = \vec{N}^T \vec{p}$$

yield, with the Hermite interpolation function $\vec{N}(x,y,z)$, the element nodal displacements $\vec{s} = (\vec{s}_1, \vec{s}_2, \vec{s}_2)^T$ and electrical potential $\vec{p}$.

In matrix notation the mechanical strain $\vec{\varepsilon}$ and the quasi-static electric field $\vec{E}$ are related to the corresponding nodal values by the differentiated form functions $\vec{N}_x, \vec{N}_y, \vec{N}_z$: $\vec{\varepsilon} = \mathbf{B}_S\,\vec{s}$; $\vec{E} = -\mathbf{B}_\phi\,\vec{p}$.

$$\mathbf{B}_S = \begin{pmatrix} \vec{N}_x & 0 & 0 & 0 & \vec{N}_z & \vec{N}_y \\ 0 & \vec{N}_y & 0 & \vec{N}_z & 0 & \vec{N}_x \\ 0 & 0 & \vec{N}_z & \vec{N}_y & \vec{N}_x & 0 \end{pmatrix}^T \quad ; \quad \mathbf{B}_\phi = \begin{pmatrix} \vec{N}_x^T \\ \vec{N}_y^T \\ \vec{N}_z^T \end{pmatrix}$$

The symmetric linear constitutive relation of piezoelectricity with the mechanical stress $\vec{\sigma}$ and the dielectric displacement $\vec{D}$ is

$$\begin{pmatrix} \vec{\sigma} \\ \vec{D} \end{pmatrix} = \begin{pmatrix} \mathbf{c}^{\vec{E}} & \mathbf{z} \\ \mathbf{z}^T & -\mathbf{d}^{\vec{\varepsilon}} \end{pmatrix} \begin{pmatrix} \vec{\varepsilon} \\ -\vec{E} \end{pmatrix}$$

For the independent field variabes $\vec{\varepsilon}$ and $\vec{E}$ the elastic stiffness $\mathbf{c}$ has to be determined at constant $\vec{E}$, the dielectric permittivity $\mathbf{d}$ at constant $\vec{\varepsilon}$, and the coupling matrix $\mathbf{z}$ is the piezoelectric matrix.

The element stiffness and mass matrices, whose quadratic forms correspond to elastic, dielectric, mutual and kinetic element energies, are :

$$\mathbf{k}_{SS} = \int \mathbf{B}_S^T\,\mathbf{c}\,\mathbf{B}_S\,dV \;,\quad \mathbf{k}_{\phi\phi} = -\int \mathbf{B}_\phi^T\,\mathbf{d}\,\mathbf{B}_\phi\,dV \;,\quad \mathbf{k}_{S\phi} = \int \mathbf{B}_S^T\,\mathbf{z}\,\mathbf{B}_\phi\,dV$$

$$\mathbf{m}_{SS} = \varrho \int \mathbf{N}^T\,\mathbf{N}\,dV \;,\quad (\varrho = \text{mass density}).$$

The global matrices $\mathbf{K}$ and $\mathbf{M}$ are derived from the assembly of all element matrices. In the case of undamped stationary vibrations with the angular frequencies ω, the following global eigenvalue equation yields

$$\left[\begin{pmatrix} \mathbf{K}_{SS} & \mathbf{K}_{S\phi} \\ \mathbf{K}_{S\phi}^T & \mathbf{K}_{\phi\phi} \end{pmatrix} - \omega^2 \begin{pmatrix} \mathbf{M}_{SS} & 0 \\ 0 & 0 \end{pmatrix} \right] \begin{pmatrix} \vec{S} \\ \vec{P} \end{pmatrix} = \vec{0} \qquad (1)$$

with the global nodal vectors $\vec{S}$, $\vec{P}$ of displacements, electrical potential.

The consistent mass matrix is compressed to about 1/15 of the size of the stiffness matrix (skyline storage) by storing only one element of each diagonal 3x3 submatrix. The eigenmodes are calculated by inverse vector iteration with overrelaxation acceleration of convergence [1].

2. High-Degree Hermite Interpolation

The method for plane triangular elements with cubic Hermite interpol. in [2] is generalized. A polynomial expansion of function f (i.e. u,v,w,ϕ) is : $f(x,y,z)=\vec{\pi}^T\vec{\alpha}$, $\vec{\alpha}=(\alpha_1,\alpha_2,\cdots)$ (vector of polynomial coeff..), $\vec{\pi}^T=(1,x,y,z,x^2,xy,xz,y^2,yz,z^2,x^3,x^2y,x^2z,xy^2,xyz,xz^2,y^3,y^2z,\cdots)$ With $\vec{d}^T=(f_1,f_2,f_3,\cdots)$ and the nodal values $f_i=f(x_i,y_i,z_i)$ at node i the interpolation condition $f_i=\vec{\pi}^T(x_i,y_i,z_i)\ \vec{\alpha}$ results. With

$$\vec{d}=\mathbf{P}\,\vec{\alpha}\ ,\quad \mathbf{P}=\begin{pmatrix}\vec{\pi}^T(x_1,y_1,z_1)\\ \vec{\pi}^T(x_2,y_2,z_2)\\ \vdots\end{pmatrix}\ ;\quad f=\vec{N}^T\vec{d}$$

$f=\vec{\pi}^T\vec{\alpha}=\vec{\pi}^T\mathbf{P}^{-1}\vec{d}$, the form function $\vec{N}=\mathbf{P}^{-1^T}\vec{\pi}$ is obtained.

In Hermite interpolation of cubic, quintic or septic degree the nodal vector $\vec{d}^{(k)}$ contains besides function- also differentiated function-values :

$\vec{d}^{(3)^T}=(f_1,f_1',\cdots)$, $\vec{d}^{(5)^T}=(f_1,f_1',f_1'',\cdots)$, $\vec{d}^{(7)^T}=(f_1,f_1',f_1'',f_1''',\cdots)$

$f'=(f_x,f_y,f_z)$, $f''=(f_{xx},f_{yy},f_{zz},f_{xy},f_{xz},f_{yz})$

$f'''=(f_{xxx},f_{yyy},f_{zzz},f_{xxy},f_{xxz},f_{yyx},f_{yyz},f_{zzx},f_{zzy},f_{xyz})$

Local co-ordinates ξ,η,ζ are used for the further discussion. For the unit rectangular ($0\le\xi,\eta,\zeta\le 1$) and the unit triangular ($\xi+\eta\le 1$) prismatic element, polynomial terms of the vector $\vec{\pi}$ must be carefully selected. Three conditions must be fulfilled: 1. invertability of interpolation matrix $\mathbf{P}$, 2. geometrical isotropy of $\vec{\pi}$ (if e.g. $\xi^2\eta\zeta$ then also $\xi\eta^2\zeta$ and for the rectangular prism also $\xi\eta\zeta^2$) , 3. conformity of interpolation. These conditions determine the high degree terms additional to the complete polynomials in 3 variables with 20, 56 and 120 terms for cubic, quintic and septic interpolation respectively.

Triangular prism :

3: $\xi^3\zeta$ $\xi\zeta^3\eta^3\zeta$ $\eta\zeta^3$, $\xi^2\eta\zeta$ $\xi\eta^2\zeta$ (*or*: $4*\xi^3\zeta$, $2*\xi^2\eta\zeta$)

5: $4*\xi^5\zeta$, $2*\xi^3\zeta^3$, $4*\xi^3\eta^2\zeta$, $2*\xi^4\eta\zeta$

7: $4*\xi^7\zeta$, $4*\xi^5\zeta^3$, $2*\xi^6\eta\zeta$, $4*\xi^5\eta^2\zeta$, $4*\xi^4\eta^3\zeta$, $2*\xi^3\eta^2\zeta^3$

Rectangular prism :

3: $6*\xi^3\eta$, $3*\xi^2\eta\zeta$; $3*\xi^3\eta\zeta$

5: $6*\xi^5\eta$, $6*\xi^3\eta^2\zeta$, $3*\xi^3\eta^3$, $3*\xi^4\eta\zeta$; $3*\xi^5\eta\zeta$, $3*\xi^3\eta^3\zeta$

7: $6*\xi^7\eta$, $6*\xi^5\eta^3$, $3*\xi^6\eta\zeta$, $6*\xi^5\eta^2\zeta$, $6*\xi^4\eta^3\zeta$, $3*\xi^3\eta^3\zeta^2$;
$3*\xi^7\eta\zeta$, $6*\xi^5\eta^3\zeta$, $\xi^3\eta^3\zeta^3$

These polynomials are called conforming because for constant ξ, η or ζ or $\xi+\eta=1$ for the triangular prism, the remainder polynomial in two variables reduces to the conforming "serendipity" interpolation over the rectangular or triangular domain. Conformity guarantees continuity of f at the element intersections! The complete interpolation polynomials over the triangular domain contain 10, 21 and 36 terms. The additional high degree terms of Hermite interpolation over the rectanguar domain are:

3: $2*\xi^3\eta$ 5: $2*\xi^5\eta$, $\xi^3\eta^3$ 7: $2*\xi^7\eta$, $2*\xi^5\eta^3$

The interpolation is determined through the node values $\vec{d}^{(k)}$ in the six or eight prism corners and $\vec{d}^{(k-2)}$ in the two triangle centres of gravity.
(26, 68, 140 ; 32, 80, 160 values for the triangular- ; rectangular prism)

Trilinear interpolation is chosen for the co-ordinates (subparametric element). With the nodal x-ordinates $x_i=x(\xi_i,\eta_i,\zeta_i)$, the nodal vector $\vec{k}_1^T=(x_1,x_2,\cdots,x_8)$, form function $\vec{G}=\mathbf{P}_k^{-1^T}(1,\xi,\eta,\zeta,\xi\eta,\xi\zeta,\eta\zeta,\xi\eta\zeta)^T$ relation $x(\xi,\eta,\zeta)=\vec{G}^T\vec{k}_1$ and corresponding expressions for y, z result.

The Jacobi matrix **J**, which contains the derivatives

$\partial x/\partial\xi=\vec{G}_\xi^T\vec{k}_1$, $\partial y/\partial\xi=\vec{G}_\xi^T\vec{k}_2$, $\cdots$,$\partial z/\partial\zeta=\vec{G}_\zeta^T\vec{k}_3$

determines with $\vec{N}^T(\xi,\eta,\zeta)=(N_1,N_2,\cdots)$ the substitutions

$(\partial N_i/\partial x,\partial N_i/\partial y,\partial N_i/\partial z)^T=\mathbf{J}^{-1}(\partial N_i/\partial\xi,\partial N_i/\eta,\partial N_i/\partial\zeta)^T$,

which allow to caculate $\vec{N}_x,\vec{N}_y,\vec{N}_z$.

The integration with the transformation $\int_V F\,dx\,dy\,dz=\int_{V0}F\,|\mathbf{J}|\,d\xi\,d\eta\,d\zeta$

is carried out by numerical Gauss quadrature. High degree quadrature rules for the triangle are taken from [3].

So far form function $\vec{N}$ is related to the node variables $\vec{d}$ containing terms partially differentiated with respect to the local variables ξ,η,ζ . In order to assemble the global matrices for distorted and arbitrarily oriented elements the appropriate variables are the global variables x,y,z :

In the cubic case chain rule $f_\xi=f_x x_\xi+f_y y_\xi+f_z z_\xi$ gives for node 1

$$f(\xi,\eta,\zeta)=\begin{pmatrix}N_1\\N_2\\N_3\\N_4\\\vdots\end{pmatrix}^T\begin{pmatrix}f_1\\f_{1\xi}\\f_{1\eta}\\f_{1\zeta}\\\vdots\end{pmatrix}=\begin{pmatrix}N_1\\x_\xi N_2+x_\eta N_3+x_\zeta N_4\\y_\xi N_2+y_\eta N_3+y_\zeta N_4\\z_\xi N_2+z_\eta N_3+z_\zeta N_4\\\vdots\end{pmatrix}^T\begin{pmatrix}f_1\\f_{1x}\\f_{1y}\\f_{1z}\\\vdots\end{pmatrix}$$

Linear combinations for six or eight nodes i are built with $\mathbf{J}^T(\xi_i,\eta_i,\zeta_i)$, and the same combinations are applied to f_x,f_y,f_z.

For second and third order derivatives chain rule yields

$f_{\xi\eta}=x_{\xi\eta}f_x+y_{\xi\eta}f_y+z_{\xi\eta}f_z+x_\xi x_\eta f_{xx}+y_\xi y_\eta f_{yy}+z_\xi z_\eta f_{zz}+$
$(x_\xi y_\eta+y_\xi x_\eta)f_{xy}+(x_\xi z_\eta+z_\xi x_\eta)f_{xz}+(y_\xi z_\eta+z_\xi y_\eta)f_{yz}$

$f_{\xi\eta\zeta}=x_{\xi\eta\zeta}f_x+y_{\xi\eta\zeta}f_y+z_{\xi\eta\zeta}f_z+$
$(x_\xi x_{\eta\zeta}+x_\eta x_{\xi\zeta}+x_\zeta x_{\xi\eta})f_{xx}+(\cdot)f_{yy}+(\cdot)f_{zz}+$
$(x_\xi y_{\eta\zeta}+y_\xi x_{\eta\zeta}+x_\eta y_{\xi\zeta}+y_\eta x_{\xi\zeta}+x_\zeta y_{\xi\eta}+y_\zeta x_{\xi\eta})f_{xy}+(\cdot)f_{xz}+(\cdot)f_{yz}+$
$x_\xi x_\eta x_\zeta f_{xxx}+y_\xi y_\eta y_\zeta f_{yyy}+z_\xi z_\eta z_\zeta f_{zzz}+$
$(x_\xi x_\eta y_\zeta+x_\xi y_\eta x_\zeta+y_\xi x_\eta x_\zeta)f_{xxy}+(\cdot)f_{xxz}+(\cdot)f_{yyx}+\cdots+(\cdot)f_{zzy}+$
$(x_\xi y_\eta z_\zeta+x_\xi z_\eta y_\zeta+y_\xi x_\eta z_\zeta+y_\xi z_\eta x_\zeta+z_\xi x_\eta y_\zeta+z_\xi y_\eta x_\zeta)f_{xyz}$

Thus for quintic and especially septic interpolation complex linear combinations of form functions result, although for trilinear co-ordinate interpolation besides $x_\xi\cdots$ only mixed derivatives $x_{\xi\eta}\cdots$, $x_{\xi\eta\zeta}$ appear.

3. Discontinuous Primary-Variable Gradients

The class of admissible functions in FE-method are those that are continuous with piecewise continuous first derivatives, i.e. among others functions of the Lagrange typ. At the intersection of materials with different elasticity, mechanical strain is non-continuous and the same is valid for the electric field at electrode borders. Discontinuities of these fields can not occur because they are used as node variables in Hermite interpolation.

For element intersections perpendicular e.g. to the z-axis it is possible to admit discontinuous normal-component of gradients. In the cubic case one has then to assign to the variables u_z,v_z,w_z,ϕ_z at the corresponding nodes of the two intersecting elements two distinct places in the system of global equations. Functions in the intersecting plane are independent of these variables, thus conformity of interpolation is ensured. For arbitrarily directed element intersections, it is necessary to rotate the variables.

One also has to consider non-continuous strain occuring at geometrical discontinuities of structure in order not to loose the high precision obtainable with Hermite interpolation. For the free length extensional resonator shown in Fig.1, results obtained with cubic Hermite interpolation are compared to those with Lagrange interpolation quadratic in x-y and cubic in z, where every cubic element is subdivided into four quadratic elements. At the marked nodes discontinuous strain is admitted or supressed.

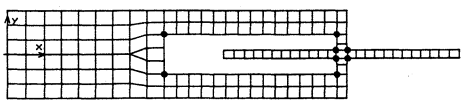


Fig.1 FE-mesh cubic Hermite interpolation, length extensional resonator

Mode:	Lagrange:	Hermite discont.:	Hermite cont.:
36 z-flexion	1161.70 kHz	1164.54 kHz	1166.87 kHz
37 length ext.	1229.24 "	1230.75 "	1233.63 "
38 xy-flex.	1237.23 "	1238.19 "	1239.57 "

Tab.1 Frequencies of the free unclamped length extensional resonator

With Lagrange interpolation the total number of degrees of freedom is 24'456 and 5'430 or 5'334 for hermite interpolation (mechanical model).

4. Resonator Specific Relations

Mass loading of thin electrodes is considered by : $\mathbf{m}'_{SS}=\mathbf{m}_{SS}+\mathbf{m}_E$,

$\mathbf{m}_E=\varrho_E\,t\int\mathbf{N}^T\mathbf{N}(x,y,z=\pm h/2)\,dxdy$

with the electrode mass density ϱ_E , the thickness t of electrodes in the x-y plane and the quartz resonator thickness h.

Stiffness of thin elecrodes is taken into account by : $\mathbf{k}'_{SS}=\mathbf{k}_{SS}+\mathbf{k}_E$,

$\mathbf{k}_E=t\int\mathbf{B}_S^T\,\mathbf{c}_E\,\mathbf{B}_S\,(x,y,z=\pm h/2)\,dxdy$

where $\mathbf{c}_E$ is the plane stress isotropic stress-strain relation.

For arbitrarily oriented electrodes, matrix **Bs** is transformed by the matrix

of direction cosines from global to local co-ordinates [4] and dxdy with the absolute value of the vector cross product $(x_\xi, y_\xi, z_\xi) \times (x_\eta, y_\eta, z_\eta)$.

Temperature effects are considered by the variations of the material coefficients and of the resonator shape. For stiffness coefficients, the relation $c_{ij}(T) = c_{ij}(25°C)(1 + Tc_{ij}^{(1)}\Delta T + Tc_{ij}^{(2)}\Delta T^2 + Tc_{ij}^{(3)}\Delta T^3)$ is given, and corresponding relations are valid for dielectric and piezoelectric coefficients. The expansion coefficients $\alpha_{ii}^{(k)}$ determine the principal expansions $a_{ii}(T) = \alpha_{ii}^{(1)}\Delta T + \alpha_{ii}^{(2)}\Delta T^2 + \alpha_{ii}^{(3)}\Delta T^3$ forming the diagonal matrix **a**. Mass density is $\varrho(T) = \varrho(25°C) / ((1 + a_{11}(T))(1 + a_{22}(T))(1 + a_{33}(T)))$. Rotating **a** according to the quartz resonator cutting angles to $\mathbf{a}' = [a'_{ij}]$, gives the node co-ordinates $\bar{K}(T) = \bar{K}(25°C) + \mathbf{a}'(T)\bar{K}(25°C)$ and the mass density of thin electrodes $\varrho_E(T) = \varrho_E(25°C)(1 - a'_{11} - a'_{22})$. Temperature coefficients Tc of gold electrodes are also considered.

IV. MOTIONAL PARAMETERS, DAMPED FREQUENCY DIP

1. Motional Parameters Cm and Rm

Motional capacitance C_m is determined by equating the kinetic energy $T = \frac{1}{2}\omega^2 \bar{S}^T \mathbf{M}_{SS} \bar{S}$ (Eq.1) to the electromagnetic energy stored in an equivalent inductance L_m created by an electric current $I = \omega Q$. In the resonance case $\omega^2 = 1/L_m \cdot C_m$, $2T = L_m I^2 = L_m \omega^2 Q^2 = Q^2 / C_m$.

$$C_m = Q^2 / 2 \cdot T \qquad (2)$$

Q is the sum of the node charges of the grounded electrodes.

Motional resistance R_m is derived from the model of hysteretic or Lorentz damping. The damping forces are considered as independent of frequency and proportional to the elastic forces. Internal friction produces a phase lag δ of strain behind stress. The energy loss per period is [5] $W_L \approx 2\pi E_L \delta \approx 2\pi E_L / Q_L$, with the elastic energy E_L and the quality factor Q_L which characterizes the mechanical losses. The definition of the quality factor of a resonator $Q_R = 2\pi T / W_L$ and the relation $R = 1/(\omega C Q_R)$ of an electric resonant circuit yield

$$R_m = E_L / (T Q_L \cdot \omega C) \qquad (3)$$

with the total resonator energy T. E_L corresponds to the elastic energies

$$B_m = \tfrac{1}{2}\int_{V_B} \bar{\varepsilon}^T \bar{\sigma}\, dV, \qquad E_m = \tfrac{1}{2}\mathrm{t}\int \bar{\varepsilon}^T \bar{\sigma}(x, y, z = \pm h/2)\, dxdy$$

of the bonding layer Bm and of the electrodes Em.

By determining n modes adjacent to the main mode, with Eq.1,Eq.2 and Eq.3 all parameters of the electric equivalent circuit of the multi mode resonator in Fig.2 are known (static capacitance $\hat{C}$).

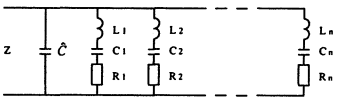


Fig.2 Electric equivalent circuit of a multi mode resonator

The |Z| impedance diagram and the $\varphi = Arctg(\Im(Z)/\Re(Z))$ phase diagram of the resonator can be determined from these values.

2 Interpolation of Cm and Rm Peaks

The course of the Cm, Bm, Em peaks is interpolated from values obtained with FEM at temperatures T1 and T2 near To. The temperature of maximum mode coupling To is obtained at the intersection of the main-mode fM(T) curve and the spurious-mode fs(T) lines of Fig.3.

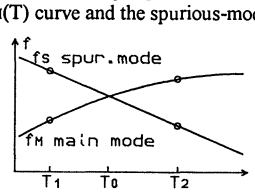


Fig.3 Intersection of frequencies at To

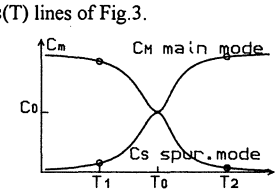


Fig.4 Motional capacitance peaks

At T=To it is no longer possible to distinguish the two modes! The following relations characterize the motional capacitance peaks of Fig.4 $C_M(T_0) = C_S(T_0) = C_0$, $C_M(T) + C_S(T) = const.$, and

$$C_S(T) = \frac{C_0}{1 + \alpha(T - T_0)^2} \qquad (4)$$

$$C_0 = (C_M(T_1) + C_S(T_1) + C_M(T_2) + C_S(T_2))/4 \qquad (4a)$$

$$\alpha_j = (C_0 / C_S(T_j) - 1)/(T_j - T_0)^2 \qquad j = 1,2 \qquad (4a)$$

The bonding layer (Fig.5a) and electrode deformation (Fig.5b) energy peaks are represented by the same curves, but with exchanged roles of the main and spurious modes.

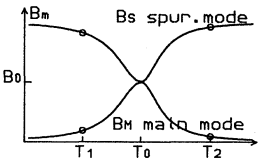


Fig.5a Bonding layer deform.energy peak

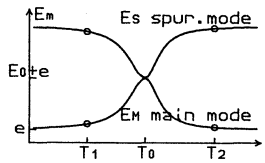


Fig.5b Electrode defor.ener. p.

$$B_0 = (B_M(T_1) + B_S(T_1) + B_M(T_2) + B_S(T_2))/4$$

$$E_0 = (E_M(T_1) + E_S(T_1) + E_M(T_2) + E_S(T_2) - 2e)/4$$

In Fig.5b the zero line energy e of Em has to be taken into account, for which the smallest EM value in the total temperature interval is taken.

$$R_0 = (B_0 / Q_B + E_0 / Q_E)/(T \cdot \omega C) \qquad (5)$$

corresponds then to the motional resistance peak amplitude (Eq.3).

3 Undamped Frequency Dip

The relation from [6] for a coupled vibration is represented in Fig.6 $(f^2 - f_M^2)(f^2 - f_S^2) = \Gamma f_M^2 f_S^2$ (coupling constant $\Gamma = 2\Delta f / f_0$).

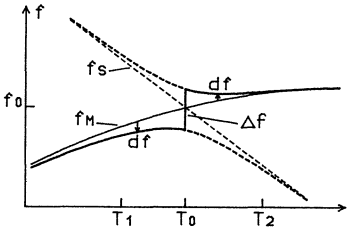


Fig.6 Coupled vibration, frequency deviation df, frequency dip Δf

The frequency deviation df from the uncoupled main mode fM was found as proportional to the slope of the spurious mode

$$df(T) = (J(T)/2C_0) \cdot (f_S(T_2) - f_S(T_1))/(T_2 - T_1)$$

and the integrated motional capacitance peak. With Eq.4

$$J(T) = \int_{-\infty}^{T} C_S(\tau)\, d\tau = C_0 \left(Arctg(\sqrt{\alpha_1}(T - T_0)) + \frac{\pi}{2}\right)/\sqrt{\alpha_1} \qquad (T \le T_0)$$

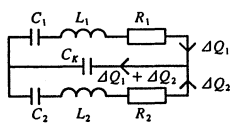
A corresponding relation yields for T≥To. This allows to determine the value of the frequency dip Δf. With Eq.4a, the relative frequency dip

$$\frac{\Delta f}{f} = -\frac{\pi}{4 f_0} \cdot \left(\frac{1}{\sqrt{\alpha_1}} + \frac{1}{\sqrt{\alpha_2}}\right) \cdot \frac{f_S(T_2) - f_S(T_1)}{T_2 - T_1} \qquad (6)$$

is determined from the FE-analysis at two arbirary temperatures near To.

4. Damped Double Oscillator, Modified Frequency Dip

The influence of damping on the beat frequency of the free stationary oscillation of a double oscillator is investigated by means of the electric equivalent circuit of the damped double oscillator, as represented in Fig.7.



$$\frac{Q_j}{C_j} + R_j \frac{dQ_j}{dt} + L_j \frac{d^2 Q_j}{dt^2} + \frac{Q_1}{C_K} + \frac{Q_2}{C_K} = 0 \quad , \quad j = 1,2$$

Fig.7 Electric equivalent circuit of the damped double oscillator. (Short-circuited L1-C1-R1-CK and L2-C2-R2 capacitively coupled with CK)

With the approach $Q_j = q_j \cdot e^{i\Omega t}$ (complex q, Ω), and equating the ratios q_1 / q_2 of the two equations yields the following relation

$$\left(\frac{1}{S_1} - \Omega^2 L_1 + i\Omega R_1\right)\left(\frac{1}{S_2} - \Omega^2 L_2 + i\Omega R_2\right) - \frac{1}{C_K^2} = 0 \quad , \quad S_j = \frac{C_j \cdot C_K}{C_j + C_K}$$

This equation is first resolved for real $\Omega = \omega$, R1=R2=0 and L1C1=L2C2.

The quadratic equation with two solutions and $\omega_1^2 - \omega_2^2 \approx 2\omega\Delta\omega$ give the relative beat frequency $\Delta f / f = (C_1 + C_2)/(2C_K)$ (7)

For the special case R1=R2≠0, C1=C2, L1=L2 the same result is obtained.

The general case $\Omega = \omega + i\delta$ leads to $a\Omega^4 - ib\Omega^3 - c\Omega^2 + id\Omega + e = 0$

$a = L_1 L_2$, $b = L_1 R_2 + L_2 R_1$, $c = R_1 R_2 + L_2 / S_1 + L_1 / S_2$

$d = R_2 / S_1 + R_1 / S_2$, $e = (C_1 + C_2 + C_K)/(C_1 \cdot C_2 \cdot C_K)$

Separation into real and imaginary part yields

$real: \quad a\omega^4 - (6a\delta^2 - 3b\delta + c)\omega^2 + (a\delta^4 - b\delta^3 + c\delta^2 - d\delta + e) = 0$

$imag.: \quad (b - 4a\delta)\omega^2 + (4a\delta^3 - 3b\delta^2 + 2c\delta - d) = 0$

The system of nonlinear equations is resolved by numerical bisection :

(Discriminant $D = (6a\delta^2 - 3b\delta + c)^2 - 4a(a\delta^4 - b\delta^3 + c\delta^2 - d\delta + e)$)

$D \geq 0$: For given δ calculate ω_1^2 from the quadratic equation and vary δ until the second equation is fulfilled. The same procedure for ω_2.

$D < 0$: For given δ calculate ω^2 from the second equation and vary δ until the quadratic equation is fulfilled. Search for δ_1 and δ_2.

The relative frequency dip $\Delta f / f$ of Eq.7 is then modified to the relative damped frequency dip $\Delta f' / f = (\omega_1^2(\delta_1) - \omega_2^2(\delta_2)) / 2\omega^2$ (8)

For $C_1 = C_2$ the relation $\delta_1(\omega_1) / \delta_2(\omega_2) = R_1' / R_2'$ results. R_0 from Eq.5 and relation $R_1 + R_2 = R_1' + R_2' = 2R_0$ give the modified main mode motional resistance peak $R_1' = 2R_0 \cdot (\delta_1 / \delta_2) / (1 + \delta_1 / \delta_2)$ (9)

For C=C1=C2, L1=L2 and $R_1 \leq R_2 << 1/\omega C$, the dip- and damping-ratio are represented in Fig.8, Fig.9 as a function of the reduced variables k , n.

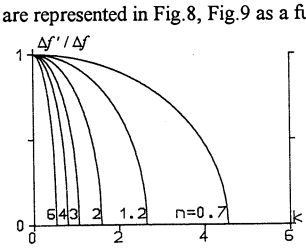


Fig.8 $\Delta f' / \Delta f$ in function of k, n

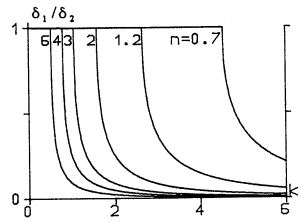


Fig.9 δ_1 / δ_2 in function of k, n.

$k = (R_2 - R_1) \cdot C_K / C = \Delta R \cdot f / \Delta f$, $n = f C$ (k: Ω/ppm, n: $F \cdot s^{-1} 10^7$)

The ratio of damped to undamped frequency dip $\Delta f' / \Delta f$ drops to zero at D=0, attains in a sharp bend the value ≈ -0.002 and for D<0 raises slowly to zero. The damping ratio $\delta_1 / \delta_2 \approx 1$ for D>0, and falls at D=0 from a sharp bend asymptotically to zero (R_1=0).

For weak coupling the spurious mode motional resistance peak is $2R_0$. This situation is prolonged to the state of maximum coupling! C_K is determined from Eq.7, Eq.6 and C_0 from Eq.4a, Eq.2. Summarizing

$C_1 = C_2 = C_0$, $L_1 = L_2 = 1/C_0\omega^2$, $C_K = C_0 \cdot f / \Delta f$, $R_1 = 0, R_2 = 2R_0$

are the values taken for the damped double oscillator. The modified frequency dip and motional resistance peak values of Eq.8 and Eq.9 are calculated with this heuristic model.

With growing spurious mode damping R_2 the frequency dip falls rapidly to zero (Fig.8), and the main mode motional resistance (Fig.9, Eq.9) remains at R_0+R(e) (R(e) in Fig.5b). With even more growing R_2 the modified frequency dip is practicaly zero whereas the motional resistance peak remains at an appreciable level. This corresponds to a situation frequently observed in practice. The fact that the dip ratio does not totally vanish for strongly damped spurious modes could explain the phenomenon of very small frequency perturbations [7].

V. RESULTS

The modelled miniature AT-cut quartz resonators are of rectangular shape with length along z'-axis and width parallel to x-axis. They are clamped with a thin isotropic epoxy resin layer onto a ceramic package. This bonding epoxy resin layer is modelled with finite elements. The elastic and dielectric constants of quartz are from [8].

1. AT Strip Quartz Resonator 16.8 MHz

The resonator length is 5.90 mm, the width w varies between 1.46 and 1.56 mm and the thickness h is about 0.096mm [1]. The AT-cutting angle θ is 35.20°. The resonator structure is subdivided in 72 • 10 quintic elements and the bonding epoxy resin layer in 6 • 7 quintic elements 0.50mm • ≈0.55mm • 0.02mm, ($E = 7000\,N/mm^2$, $\nu = 0.35$, $\varrho = 2.6\,g/cm^3$). The rectangular Au-electrodes are 3.00mm • 1.33mm • 0.2 μm, ($E = 77500\,N/mm^2$, $Tc^{(1)} = -1.69 \cdot 10^{-4}/°C$, $Tc^{(2)} = -1.93 \cdot 10^{-6}/°C^2$ [9]; $\nu = 0.3$, $\varrho = 18.5\,g/cm^3$). The structure of the resonator is represented in Fig.10 with the thickness shear fundamental mode.

In Fig.10 half of the modelled resonator is represented with the finite element mesh. The symmetry plane is at the bottom border. The bonding epoxy resin is at the upper left corner (0.02 mm thick).

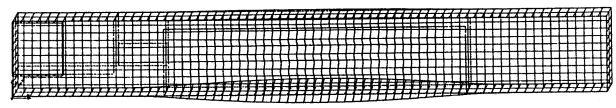


Fig.10 Upper half of the resonator. Thickness shear fundamental mode. width =1.500, displacements (magn. 50 , resonator energy 2 mWs)

The coupling of the thickness-shear mode with the width-flexural mode is investigated. The resonator was additionally modelled with cubic elements by subdividing every quintic element into four cubic elements.

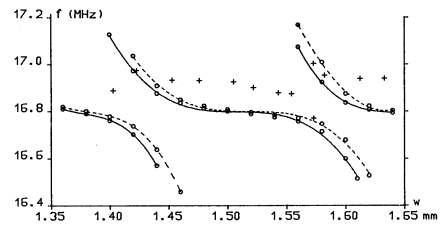


Fig.11 Coupling of thickness-shear with width-flexural mode, w = width
—— quintic interpolation, ····· cubic interpolation, + experiment
Reson.thickness: calculation h=0.096562, experiment h=0.0960 mm

Quintic interpolation leads to a better discretization, because mode splitting occurs at smaller resonator widths w, and this at a considerably reduced number of degrees of freedom of 56'847 compared to 82'885 for cubic interpolation! The number of stored elements of the stiffness matrix is 53'829'000 for quintic and 54'276'000 for cubic interpolation.

The calculated impedance and phase diagram of a resonator with strong activity dips are represented in Fig.12. Bonding layer Q_B=1000, electrode Q_E=3000. The spurious mode peaks are hidden by the main mode peaks!

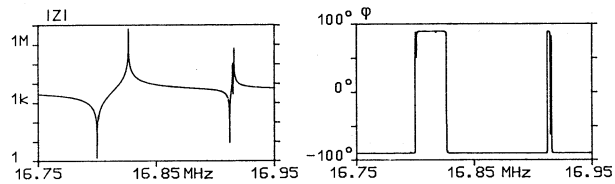


Fig.12 Impedance and phase diagram, w=1.47mm; Q_B=1000, Q_E=3000, $\hat{C}$=2.2 pF. Main mode Cm ≈ 8 fF, 2.anh.overtone Cm ≈ 0.7 fF.

Of the 24 modes in the frequency interval 16.75-16.95MHz, besides the fundamental thickness shear mode (mode 891) and the second anharmonic thickness shear overtone, just 4 weak modes appear.

Calculated frequency/temperature curves including spurious modes and motional-capacitance and -resistance curves are plotted in Fig.13, Fig.14.

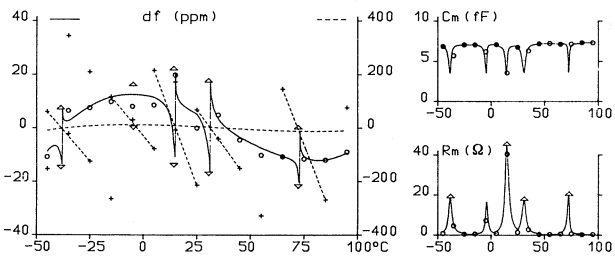


Fig.13: f/T-curve with spurious modes, Cm/T-, Rm/T-curve. w=1.470mm

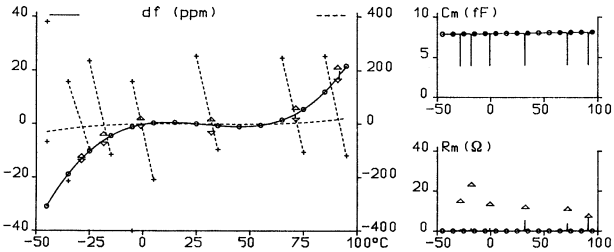


Fig.14: f/T-curve with spurious modes, Cm/T-, Rm/T-curve. w=1.500mm. Plotted are the damped peaks according to Eq.8 and Eq.9, Q_B=1000, Q_E= 3000. The triangles represent the undamped peaks according to Eq.6,Eq.5.

In Fig.13 large Cm peaks appear, which produce strong frequency dip and Rm peaks. A spurious mode which produces a strong dip at 30°C is plot-

ted in Fig.15. A regular flexural vibration modulated along length and a less regular width-, length-extensional vibration are superposed.

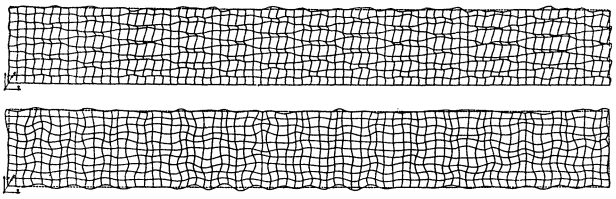


Fig.15 Spurious mode at 45°C, w=1.47. Mid-plane surface displacements
above: deformations perpendicular to mid-plane (magn. 50)
below: deformations parallel to mid-plane (magn.100)

Spurious modes with irregular vibration patterns only produce narrow Cm peaks and large bonding layer deformations, thus small frequency dips. Just one small damped frequency dip of 3.4 ppm and strongly damped resistance peaks appear in Fig.14. It is remarkable to notice that all spurious modes have negative slopes $\Delta f/\Delta T$ (-10 to -45 ppm/°C).

A measured impedance and phase curve is represented in Fig.16. The small dip in the center is a result of the first anharmonic thickness shear overtone, which theoretically is electrically inactive but appears due to the inevitable symmetry errors of the manufactured resonators.
Two experimental temperature curves of frequency and motional resistance are represented in Fig.17.

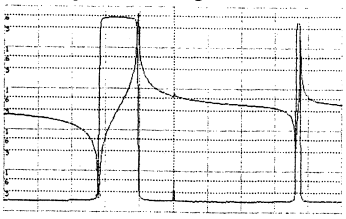


Fig.16 Measured impedance, phase diagrams, w=1,47mm (span 200kHz)

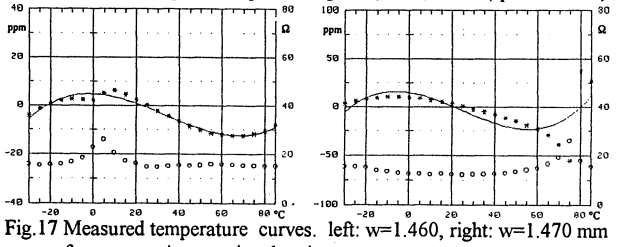


Fig.17 Measured temperature curves. left: w=1.460, right: w=1.470 mm
frequency: * ; motional-resistance: o

Effects of coupling with the width-flexural mode are represented in Fig.18, the inflexion temperature Ti and the linear coefficient α of the relation
$f(T)=f(Ti)(1+\alpha(T-Ti)+\gamma(T-Ti)^3)$, versus width w.

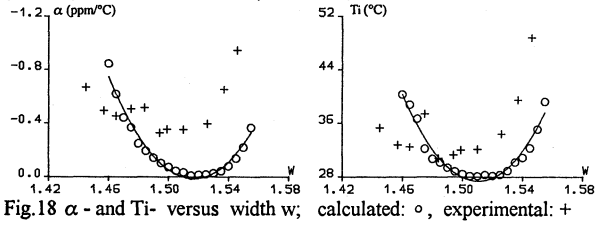


Fig.18 α - and Ti- versus width w; calculated: o , experimental: +
For w=1.50mm: $\gamma = 0.091$; 0.095 $ppb/°C^3$ calc. ; exper.

Only qualitative comparison with experimental results is possible because experimental values are with constant resonator thickness h (Fig.11) but values h of the calculations are adjusted to obtain the frequency 16.8MHz: w=1.46: h=0.096802, 1.49: 0.096605, 1.52: 0.096488, 1.55: 0.096350.
With t=0.4 μm : f=16.8MHz, w=1.49, Ti=28.6 °C, γ =0.092 ppb/°C^3 , w=1.44: h=0.093947, α =-0.797ppm/°C, 1.47: 0.093782, -0.110, 1.50: 0.093680, -0.088, 1.53: 0.093546, -0.500. (t=0.4 μm see below).

The influence of thicknesses of the bonding epoxy resin and of the electrodes on the frequency dip and motional resistance behaviour are investigated further. In these extensive investigations each configuration was modelled with 20 different resonator widths w with increments of 5 μm. Besides the configuration with the epoxy thickness b=0.02mm and the Au-electrode thickness t=0.2 μm three additional configurations: b=0.00mm, t=0.2 μm ; b=0.03mm, t=0.2 μm and b=0.02mm, t=0.4 μm are analysed. For each configuration are represented in Fig.19- Fig.22 with varying width w the maximum frequency dip (undamped Eq.6, damped Eq.8) and the maximum motional resistance (undamped Eq.5, damped Eq.9) appearing within the temperature intervall -45°C ·· 95°C, for different bonding epoxy resin quality factors Q_B. The influence of Au-electrode quality factor Q_E is of minor concern and is set to 3000. The frequencies at T=25°C are 16.8MHz ± 250Hz.

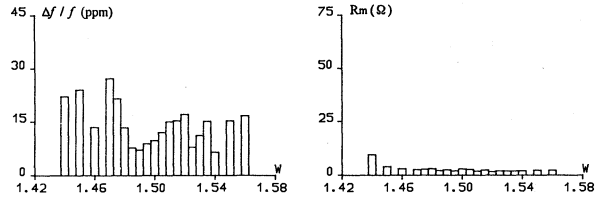


Fig.19 Max: $\Delta f/f$, Rm versus width w. b=0.00mm, t=0.2 μm

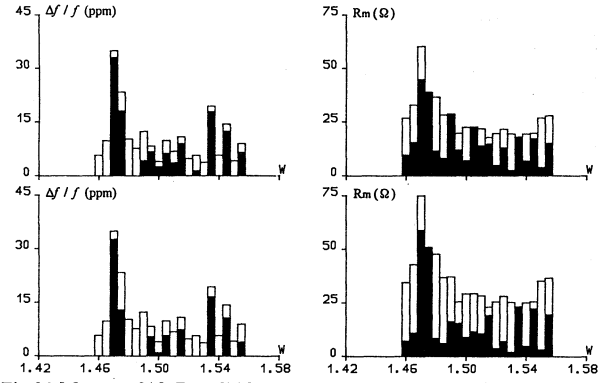


Fig.20 Max: $\Delta f/f$ □, $\Delta f'/f$ ■; Rm □, R'm ■. b=0.02mm, t=0.2 μm
above : Q_B=1000 ; below : Q_B=750

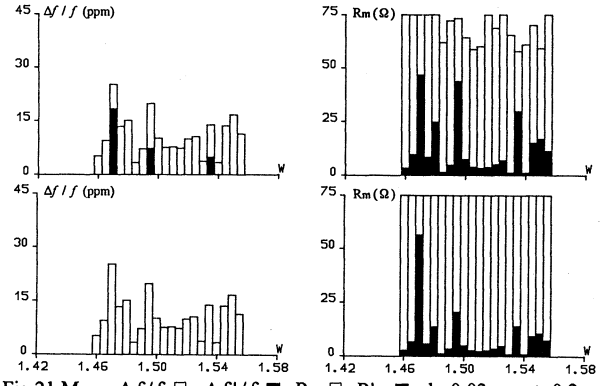


Fig.21 Max: $\Delta f/f$ □, $\Delta f'/f$ ■; Rm □, R'm ■. b=0.03mm, t=0.2 μm
above : Q_B=1500 ; below : Q_B=1000

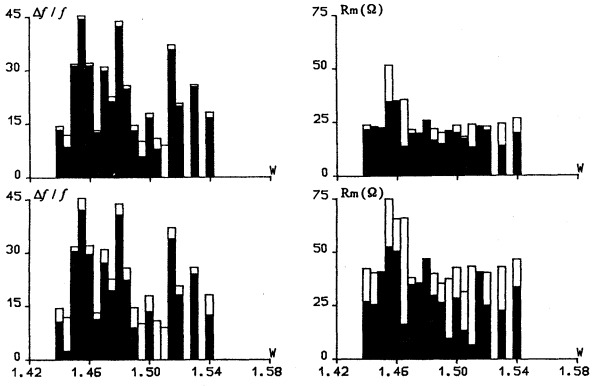


Fig.22 Max: $\Delta f/f$ □, $\Delta f'/f$ ■; Rm □, R'm ■. b=0.02mm, t=0.4 μm
above : Q_B=1000 ; below : Q_B=500

Instead of representing the maximum values in Fig.19- 22 one could also represent the sum of the dip and peak values appearing in the temperature interval -45°C, 95°C each, but qualitatively the same results would yield.
Comparing the undamped frequency dips of Fig.19 obtained with a rigid clamping of the resonator (discontinuous mech.strains are taken into ac-

count, III.3.) with those of Fig.20, Fig.21, it is remarkable to notice that the undamped frequency dips generally diminish with growing bonding layer thickness (compliance). Another striking feature is the high spurious mode coupling obtained with thick electrodes due to a stronger trapping effect, which leads to the large undamped frequency dips of Fig.22. This effect is clearly confirmed in practice.
For constant Q_B the undamped resistance values grow with larger bonding thickness (damping), but with the model IV.4. both the damped frequency dip and resistance peak values of Fig.21 are smaller than those of Fig20. In Fig.21 all dips are damped away with Q_B=1000, whereas the dips of Fig.22 are, due to strong coupling, resistent to damping.
The mechanical quality factor Q_B of the epoxy resin depends to a large extent on the drying temperature and is, as well as the exact value of layer thickness, actually unknown. The distribution of the damped dips of Fig.20 with Q_B=1000 seams to be near to reality. The actual version in production is with width w=1.505mm.

2. AT-Cut 50 MHz High Frequency Fundamental Quartz Resonator

The crystal geometry is an AT strip structure of 5.8mm length, 1.87 mm width and 0.125 mm thickness, with a recessed active area thinned down by chemial etching to a membrane of 0.033mm thickness, surrounded by a frame of 0.1mm width [10]. To consider vibrations in the thick part of the resonator, quintic interpolation is not sufficient thus septic Hermite interpolation is chosen for the whole structure (12 • 50 elements). The Al-electrodes of 0.09 μ m thickness ($1mm^2$, E=70'000N/mm^2, ϱ =2.70 g/cm^3) are connected electrically through Au-tracks (0.2 μ m). The epoxy resin layer is 1.1mm • 0.7mm • 0.02mm. The AT-cuting angle θ =35.23°.

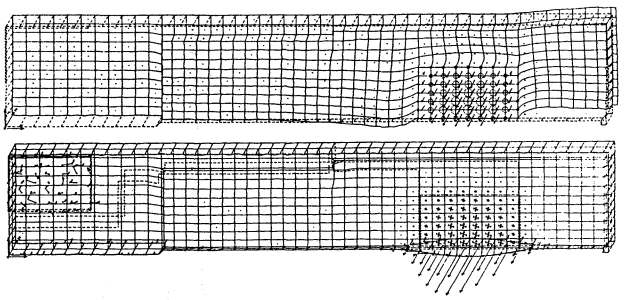


Fig.23 Upper half of the resonator. Fundamental thickness shear mode
above: electrical potential and charges, dielectric displacements
below: Displacements (magn.50), forces, principal tensions.

Charges represented with circles appear at the grounded electrode. Electrical potential gradient is discontinuous at electrode borders (III.3.), (mechanical strain is discontinuous at the intersection to the frame).
The effect of coupling with width-flexure is represented in Fig.24. Total resonator width w is varied with constant frame and electrode widths.

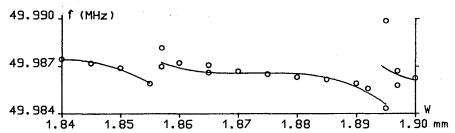


Fig.24 Coupling of thickness-shear with width-flexural mode, w = width

Due to the irregular shape (variable frame width) of the thin membrane, the coupling to width-flexural mode is not very pronounced. In Fig.25 the width-flexural component with w=1.90mm is represented. The irregular width-flexure pattern with very small wavelengths is clearly recognizable.

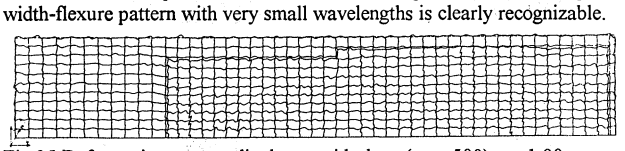


Fig.25 Deformations perpendicular to mid-plane (mag.500), w=1.90.

The variation of the frame structure thickness (0.125mm) has only very little influence on the fundamental mode frequency: 0.115: 49.98677MHz, 0.123: 49.98672MHz, 0.125: 49.98658MHz, 0.129: 49.98651MHz.

A calculated frequency/temperature curve with the very dense spectrum of spurious modes is represented in Fig.26, (Temperature increment 5°C). Main mode is mode 12'010! The maximum undamped frequency dip yields 6.65 ppm. With Q_E=10'000 and Q_B=4000 the dip value is 3.97 ppm and drops to zero with Q_B=3000.
Ti=22.0°C, $\gamma = 0.098 ppb/°C^3$. Main mode Cm ≈ 6 fF.

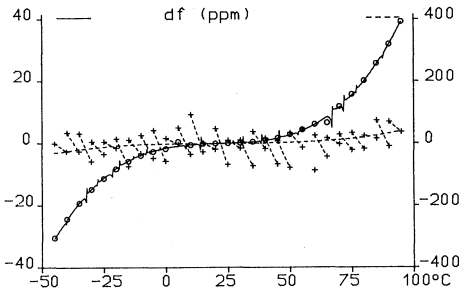


Fig.26 f/T curve (undamped dips) with spurious modes, w=1.87mm.

Due to the tremendous size of the eigenvalue problem (100'244 degrees of freedom, 235'952'000 places stiffness matrix i.e. 1.89 GB!) arising from the high discretization level, more detailled investigations are hardly feasible (Digital Alpha Station 255/300, 256 MB Ram). Nevertheless some principal insights have been obtained.

VI. CONCLUSION

The developed finite element program with high degree Hermite interpolation has proven to be numerically stable for the huge equations resulting from the analysis of high frequency quartz crystal resonators. Some principal insights of the spurious mode behaviour are gained. Besides the electrode thickness, the bonding layer compliance influences the strength of coupling with spurious modes. A method to determine the influence of epoxy resin layer and electrode damping onto the spurious mode coupling strength is developed. Apparently more regular frequency temperature curves are obtained with light electrodes and thick bonding epoxy resin layers. Transfer of quantitative results to reality is less obvious. The spectrum of mathematical modes of the modelled high frequency crystals is very dense and also chemical milling of miniature resonators leads to small (constant) shape irregularities which cannot be modelled. Besides the presented ambitious analysis, the finite element program is also successfully applied to the analysis of many different lower frequency resonators described in [11].
The autor wishes to thank Dr. W. Zingg from Micro Crystal for the helpful discussions and the supply of the various experimental results of the presented crystals.

VII. BIBLIOGRAPHIE

1. K.Trümpy, W.Zingg, Finite Element Analysis of AT Strip Quartz Resonators, Proc. 7th EFTF, Neuchâtel 1993, pp 45-50.
2. H.R.Schwarz, Methode der Finiten Elemente, Teubner, Stuttgart 1991.
3. D.A.Dunavant, High Degree Efficient Symm. Gaussian Quadrature Rules for Triangle, Int.J.Num.Meth.Eng. Vol.21 1985, pp 1129-1148.
4. O.C.Zienkiewicz, Methode der Finiten Elemente, C.Hanser, Wien1975.
5. Traité d'électricité Vol. XXI, Editions PPR, Lausanne 1986.
6. T.Ikeda, Fundamentals of Piezoelectricity, Oxford Univ. Press, 1996.
7. P.E.Morley, J.Dowsett, F.Ross, Small Freq./Temp. Perturbations in Resonators for Crystal Oscil., 7th EFTF, Neuchâtel 1993, pp181-185.
8. B.J.James, A New Measurement of the Basic Elastic and Dielectric Constants of Quartz, 42th AFCF 1988, pp 146-154.
9. International Critical Tables, 1927.
10. B.Studer, W.Zingg, S.Dalla Piazza, Char. of High Freq. Fundam. Rectang. Quartz Crystal Res., Proc. 10th EFTF, Brighton 1996, pp 14-20.
11. B.Studer, W.Zingg, Technology and Characteristics of Chemically Milled Miniature Quartz Crystals, 4th EFTF, Neuchâtel 1990, pp 653-658

EFFECT OF C-AXIS MISALIGNEMENT ON SAPPHIRE DISK RESONATOR PERFORMANCES

O. Di Monaco, W. Daniau, H. Merigoux[#], Y. Kersalé, V. Giordano

Laboratoire de Physique et de Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté Besançon
32, avenue de l'Observatoire- 25044 Besançon Cedex - France

[#] Laboratoire de Cristallographie
de l'Université de Franche-Comté Besançon
La Bouloie- 25000 Besançon - France

1. ABSTRACT

Very high-Q factors could be obtained in microwave frequency range using monocrystal sapphire resonator associated with a special field configuration (whispering gallery mode). Such a resonator is constituted by a sapphire disk placed in a center of a cylindrical copper cavity. Generally it is assumed that the C-axis of the sapphire monocrystal is parallel to the cylinder axis. In that case, Maxwell field equations could be separeted and solved with good accuracy in order to calculate frequencies and Q-factors of the main resonance modes. Experimental Q-factors better than 30 millions in X-band have been obtained at LPMO with well oriented crystals.
We had the opportunity to test several sapphire diskes with different C-axis orientations. On each of them, we have performed permittivity measurements showing significant result variations correlated with C-axis misalignment. Q-factors in whispering mode operation have been also measured. Information about C-axis tolerances for high performances have been derived.

Keywords: Microwave resonator, Whispering Gallery Mode Resonators, Low phase noise microwave sources.

2. INTRODUCTION

A Whispering Gallery Mode Resonator (WGMR) is constituted by a sapphire disk enclosed in a cylindrical copper cavity. The special mode configuration assures the confinement of the electromagnetic energy in the sapphire it self. As a consequence of the low dielectric losses of the sapphire monocrystal, very high-Q values at microwave frequencies could be obtained with this configuration. The figure 1 shows one WGMR prototype operating near 9 GHz. The sapphire is 36 mm diameter and 9 mm high.

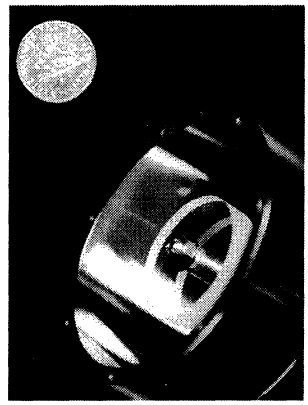


Figure 1. WGMR at 9 GHz

The sapphire monocrystal is an anisotropic dielectric. The permittivity tensor could simply expressed in the cylindrical coordinate system where z-axis is assumed to be parallel to the C-axis of the crystal. In such condition we have:

$$[\varepsilon] = \begin{bmatrix} \varepsilon_t & 0 & 0 \\ 0 & \varepsilon_t & 0 \\ 0 & 0 & \varepsilon_z \end{bmatrix}$$

In that case, Maxwell field equations could be separated and solved with a good accuracy [1]. We find two whispering gallery mode families, i.e. WGH which are quasi-TM modes and WGE (quasi-TE modes). Numerical computations give resonance frequencies with an accuracy better than 1%. Q factor could be also evaluted within 20% accuracy.
Practically, the C-axis of the sapphire monocrystal is not perfectly aligned with the z-axis of the cylinder. Typical misalignements for good quality crystals are less than 1°, but higher values can be found depending on the origin of the samples. During our experimental studies related to WGMR, we found some departures from the theory that could be explain only from a misalignement of the sapphire crystal. We give in this paper experimental results showing that in some circumstances, misalignement affects the WGMR properties.

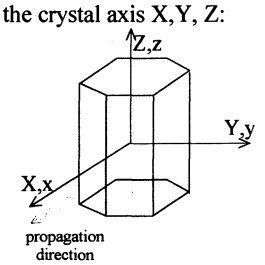
3. PERMITTIVITY MEASUREMENTS

Permittivity and dielectric losses (tgδ) have to be known with a good accuracy in order to fix the WGMR dimensions to obtain the wanted frequency and to evaluate the Q-factor. The Courtney resonator method [2] is generally used to determine the components ε_t and ε_z of the relative permittivity tensor. A sapphire disk is insered between two parallel copper plates. The measurement of the TE_{011} and TM_{011} modes frequencies allows the determination of ε_t and ε_z [3]. In the table 1, we give results for 3 different sapphires at 300 K.

	Fabrication method	Crystal orientation Θ	ε_t	ε_z	tgδ
S1	Verneuil no optical polishing	0.9°	9.43	11.42	1.10^{-5}
S2	Verneuil optical polishing on flat surfaces	5.0°	9.43	11.32	6.10^{-6}
S3	Hemex[ ] optical polishing on all surfaces	0.25°	9.43	11.54	4.10^{-6}

Table 1. Permittivity measurement results

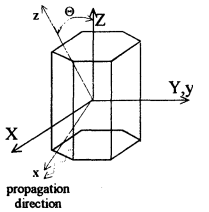
It should be noted that for the 3 sapphires, the measure gives the same value of the component ε_t. On the contrary significant discrepancies in the value of ε_z have been observed. Great care have been take in order to minimize the instrumental errors which can not give such departures. The total instrumental uncertainty is of the order of ±0.01.
The sapphire crystals differ in the growth method used for their elaboration but also in crystal orientation. The later has been measured by one of us using classical crystallography techniques. The angle Θ corresponds to the misalignement between crystal C-axis and geometrical z-axis.
It is clear that the misalignement could give erroneous values of the permittivity. Nevertheless, the exact relationship between Θ and the measured parameters is not evident to establish.
In order to establish a simple model, lets us consider one plane wave propagating along the x-direction in a hexagonal crystal media [4] where the coordinate axis x,y,z are chosen to coincide with the crystal axis X,Y, Z:



In that case the [ε] tensor takes the form:

$$[\varepsilon]=\begin{bmatrix}\varepsilon_t & 0 & 0\\ 0 & \varepsilon_t & 0\\ 0 & 0 & \varepsilon_z\end{bmatrix}$$

It is easy to find that the resolution of the Maxwell equation gives two solutions. The first has the **E** field oriented in the y-direction (corresponding to a TE mode) with a phase velocity equal to $(\mu_0 \varepsilon_t)^{-1/2}$. The second one with the **E** field oriented in the z-direction is sensible to the ε_z component of the permittivity tensor.



In the case where the Z-direction of the crystal is no longer parallel to the z-axis, one obtain again a propagation with **E** polarized along the y-axis with a permittivity ε_t. This mode is then independant of the angle Θ. However, the second propagation solution corresponds to a quasitransverse wave with **E** polarized in the XZ plane. In that case, the effective permittivity is :

$$\varepsilon_z' = \varepsilon_t \sin^2\Theta + \varepsilon_z \cos^2\Theta - \frac{\left((\varepsilon_t-\varepsilon_z)\cos\Theta\sin\Theta\right)^2}{\varepsilon_t\cos^2\Theta+\varepsilon_z\sin^2\Theta}$$

with an angle Θ = 5° and ε_t=9.43 and ε_z=11.55, we found:

$$\varepsilon_t' = \varepsilon_t = 9.43$$
$$\varepsilon_z' = 11.52$$

This simple model gives the right tendancy but not the actual departure. In the Courtney resonator, TE_{011} and TM_{011} dispersion equations are established thanks to the separability of the Maxwell equations. The misalignement complicates drastically the problem and no analytical solution could be found. Perhaps performant Finite Elements Model could bring usefull information.

We have also determined the dielectric loss, i.e. tgδ, for the 3 crystals. We measure the Q-factor of the sapphire disk used in a WGMR and operating in a high order mode. In such condition the Q-factor is only limited by the losses occuring in or on the crystal. Generally it's assumed that $Q_0 = 1/tg\delta$. It appears that the variation of the Q-factor could not be correlated with the misalignement of the crystal. In fact at ambiant temperature the parameter which limits the Q-factor seems to be principally the surface preparation. Optical polishing gives the highest Q-factor.

4. FREQUENCIES AND Q_0 AT 300 K.

We have at our disposal 3 sapphires diskes A, B, C of the same provenance elaborated with a Verneuil method. The geometrical characteristics of these sapphires are given in table 2.

	diameter	h	Θ
A	37.1 mm	9.2 mm	9.2°
B	37.1 mm	9.2 mm	1.6°
C	37.1 mm	9.2 mm	<0.1°

Table 2. Characteristics of A,B,C sapphire diskes.

We placed successively the 3 sapphires in the same copper cavity and measured the frequency and Q-factor of the $WGH_{m,0,0}$ (quasi-TM) modes for m = 4 to 10.

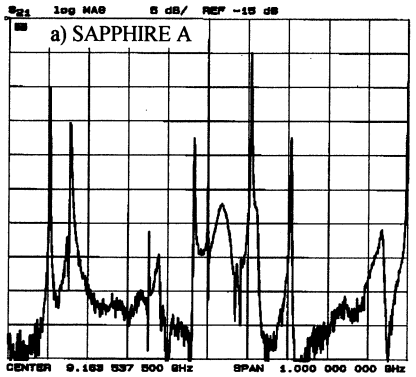
m	theory	A	B	C
4	6.84	6.94	6.36	6.36
5	7.57	7.68	7.69	7.68
6	8.30	8.42	8.41	8.41
7	9.04	9.16	9.14	9.14
8	9.80	9.99	9.87	9.88
9	10.56	10.61	10.52	10.52
10	11.33	11.37	perturbed	11.38

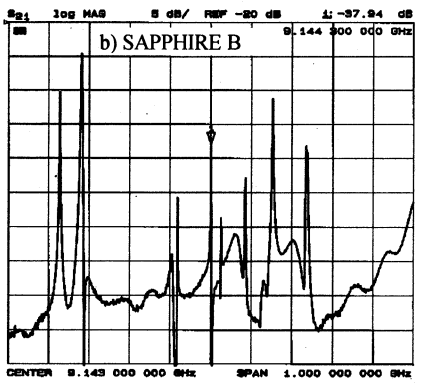
Table 3. $WGH_{m,0,0}$ mode frequencies (GHz)

m	A	B	C
4	29,000	15,000	15,000
5	78,000	60,000	60,000
6	92,000	98,000	90,000
7	120,000	130,000	140,000
8	100,000	135,000	125,000
9	135,000	120,000	80,000
10	130,000	-	110,000

Table 4. $WGH_{m,0,0}$ unloaded Q-factors.

In view of these results, it's clear that misalignement don't affect significantly the performances of the WGMR. Unloaded Q-factor are quasi independant of Θ. Nevertheless, we note some differences in frequency between A and the two other diskes. On this account the $WGH_{7,0,0}$ is representative. In figures a,b,c, one can see the transmission coefficients on the 3 WGMR arround the $WGH_{7,0,0}$ resonance.





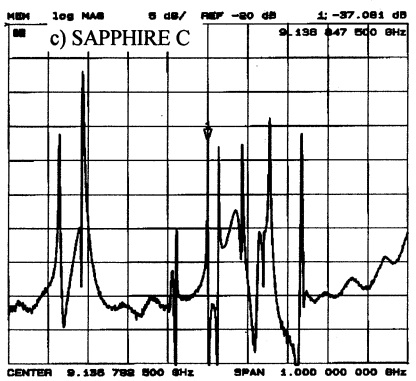


Figure 2. S_{21} of the WGMR A,B and C.

Each record is centered on the highest Q resonance. The B and C spectra are quite similar. On the other hand, in the A spectrum the highest Q resonance does not correspond to the same mode. There is, in a way, an exchange of properties with the nearest mode as we can see more precisely in the figure 3.

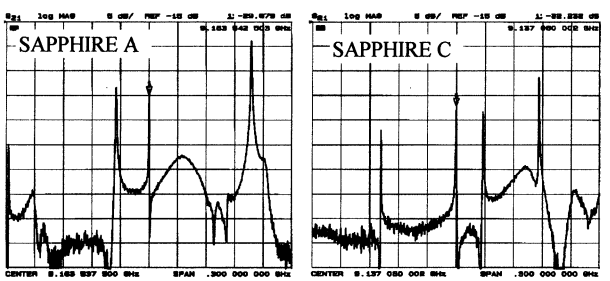


Figure 3. Comparison between WGMR A and C

5. Q-FACTOR AT 77K

At 77 Kelvin A and C WGMR present about the same unloaded Q-factor:

	frequency	Q_0
A	9.25 GHz	850,000
C	9.22 GHz	900,000

It should be noted that higher value of Q-factor are obtained with a high quality crystal. As an example, with the S3 sapphire we have $Q_0=30.10^6$ at 7.4 GHz.

6. MODAL SELECTION

To prevent oscillation on lower Q spurious mode, we have to implement a mode selector. We use as related in a preceding papers [5] thin metallic lines deposited on the top of the sapphires disk. The principle of a such mode selector is recalled in figure 4.

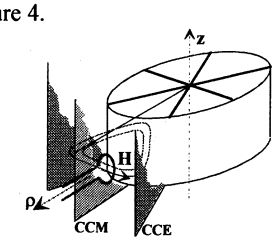


Figure 4. Principle of the modal selection

In the WGMR, the electromagnetic field components have periodical variations along the azimutal direction φ. Assuming that $\varphi=0$ corresponds to the location of the magnetic probe plane as shows on figure 3, we have for the $WGH_{m,0,0}$ mode:

$$\begin{vmatrix} H_z \\ H_\rho \\ E_\varphi \end{vmatrix} \propto \sin m\varphi \quad \text{and} \quad \begin{vmatrix} E_z \\ E_\rho \\ H_\varphi \end{vmatrix} \propto \cos m\varphi$$

As a consequence of these combined azimutal variations, the meridian planes $\varphi=k\pi/m$ (k = 1,2,...2m) could be considered as Perfect Magnetic Conductor (PMC) surfaces. Conversely, the planes $\varphi=(2k+1)\pi/2m$ behave as Perfect Electric Conductor surfaces. It is obvious that, thin metallic lines deposited on the top of the sapphire disk along the directions $\varphi=(2k+1)\pi/2m$ should have negligible influence on the $WGH_{m,0,0}$ mode. On the other hand, the modes with a different field configuration will be strongly affected by the presence of the metallic lines.

We deposited on each sapphires 14 thin metallic lines in order to select the $WGH_{7,0,0}$ mode. The mode selector is shown in figure 5.

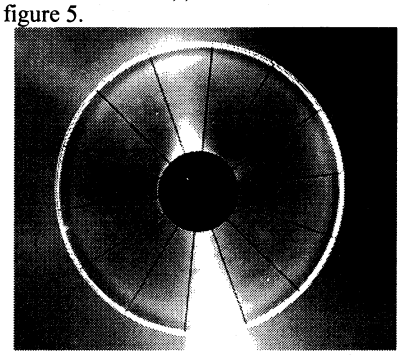


Figure 5. WGMR with the mode selector

The following figures 6 and 7, obtained with the well oriented sapphire, i.e. sapphire C, demonstrate the efficiency of our mode selection technique. It allows the use of the WGMR in an oscillating loop without the need of an additional filter.

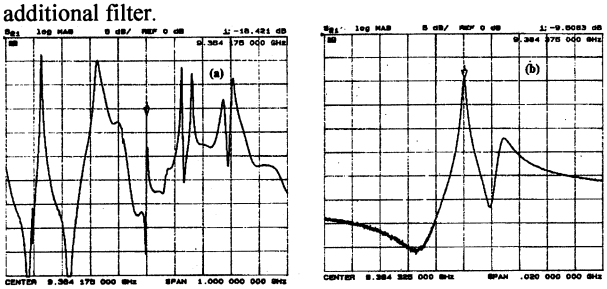


Figure 6. $WGH_{7,0,0}$ sapphire C without mode selector span: a)1GHz b)20MHz

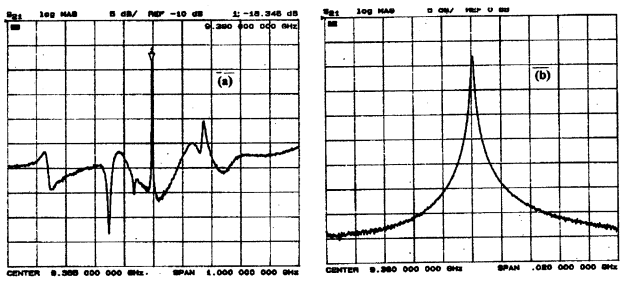


Figure 7. $WGH_{7,0,0}$ sapphire C with the mode selector span: a)1GHz b)20MHz

The figure 8 shows the phase noise spectrum obtained with an oscillator constructed with the WGMR-C resonator. It is constituted by the resonator, one commercial large band (6-12 GHz) loop amplifier and a mechanical phase shifter.

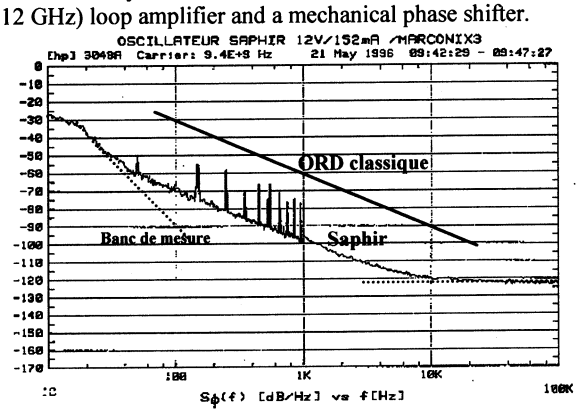


Figure 8. Phase noise spectrum $S_\Phi(f)$ of the Sapphire C oscillator

The phase noise performances, S_φ(1kHz) = -100 dB/Hz, is comparable to the performances published by other groups. It demonstrats that it is possible to obtain good phase noise performances without the use of high quality expensive sapphire crystal. The cost of our crystal is of the order of 2000 FF.

However, if the crystal is not well oriented, The electromagnetic fields configuration should not longer follow the theorical configuration assumed to explain the mode selection principle.

We compare the unloaded Q-factors of the WGMR A,B and C equiped with the same mode selector :

A	B	C
10,000	40,000	140,000

For the well oriented crystal, the mode selector don't affect the Q-factor of the WGMR as espected. However, for B and A diskes the effect of the metallic lines on the performances are significant. We loose here a factor 10 on the Q-factor.

CONCLUSION

In conclusion, we can say that if we are not interested to obtain ultra-high performances, sapphire crystal of medium quality are suffisant. Q-factor higher than 100.000 at 300K and of the order to 1.10^6 at 77K are currently obtained. It is interesting to note that Q-factors of 1.10^6 are obtained with sapphire-HTc superconductors thin films resonators. In that case, the use of expensive and delicate HTc thin films is totaly unjustified. Nevertheless, if the implementation of a mode selector is necessary, we have to take care of the crystal orientation of the sapphire disk to maintain high Q. Misalignement has to be lower than 1° to prevent extra losses. In any case, if we search ultra high performances, high quality, well aligned and optical polished crystals are required.

REFERENCES

[1] *"Approximat approach to the design of shielded dielectric disk resonators with Whispering-Gallery modes."* E.N. Ivanov, D.G. Blair and V.I. Kalinichev. IEEE Trans. on MTT, Vol 41, n°4, April 1993, pp 632-638.

[2] J. Krupka, R.. Geyer, M. Kuhn and J. H. Hinken. *Dielectric Properties of Single Crystals of Al_2O_3, $LaAlO_3$, $NdGaO_3$, $SrTiO_3$ and MgO at Cryogenic Temperatures.* IEEE Transactions of Microwave Theory and Techniques, Vol. 42, N°10, Oct 1994, pp1886-1889.

[3] O. Di Monaco, P. Blot, Y. Gruson, V. Giordano: *Mesure des propriétés diélectriques à basse température.* Actes des 4eme journées de Caractérisation Micro-onde et Matériaux. Chambéry, Avril 1996, pp 45-48.

[4]B.A. Auld. *Acoustic fields and waves in solids.* vol.1, chp.4. ,John Wiley editor, ISBN 0 471 037001

[5] V.Giordano, W. Daniau, G. Martin, F. ladret-Vieudrin, M. Olivier, O. Di Monaco, M. Chaubet. Original Mode Selection For Whispering Gallery Mode Resonator. *Proceedings of the 9th European Time and Frequency Forum,* (1995), pp 297-300

HIGH OVERTONE BULK ACOUSTIC WAVE COMPOSITE RESONATORS - REVIEW OF MODERN STATE AND NEW POSSIBILITIES

G.D. MANSFELD

Institute of Radioengineering and Electronics of Russian Academy of Sciences
11, Mokhovaya str., MOSCOW, 103907 - RUSSIA *

The recent original results of the author together with the review of the data of numerous publications give insight to the problems connected with the theory, design and applications of High overtone BAW composite Resonator structures (HBAR). In this paper following topics are included :

1. Theory and unified expressions for electrical impedances of the resonators consisting of few layers (low loss substrate S, transducer piezoelectric layers P, metallic contact films M) taking into account wave progagation properties of all layers. In particular the impedance, frequency spectrum and quality factors of such new structures as M-P-S-M and M-P-S-P-M are analysed as a function of layer properties and thicknesses and frequency. The expressions for traditional P-P-M-S and M-P-M-S-M-P-S structures are verified by comparison with experiments, in particular with data for the original 1.8 GHz langasite resonator.

2. The simplified approach to the problem of energy trapping in multilayer structures is given. It explains a very good trapping in most existing composite resonators on microwaves. The relatively simple expressions for the evaluation of losses connected with the energy escape from layers of the structure are given. Perfect energy trapping opened the possibility to suggest and develop HBAR method of acoustic losses measurements in thin films and layers.

3. The possibilities of electric tuning and the influence of electric load on characteristics of two port resonators/filters (M-P-M-P-M and M-P-M-S-M-P-M) are discussed. The effect of tuning is based on the fact that the phase of the reflection coefficient of acoustic waves from interface between layers depends on the electric load which is connected to one of the layers. As a result the resonant frequencies of the structure may be tuned by the variation of the load. For gigahertz frequency band it is possible to change the resonator frewquency rather effectively -up to few MHz. The tunability is drastically suppressed by the resistive components in the load and it explains the fact of high stability of traditional resonator (relatively high "parasitic" resistivity of output circuit).

4. Effects in HBAR connected with external control of BAW propagation characteristics in layers. For the case of substrate made of ferrite (YIG) the analysis of possibility of frequency tuning on microwaves by applied magnetic field is given. The best agreement with experiment is achieved with taking into account "exchange" magnetic modes. If the substrate is made of dielectrics and piezodielectrics the propagation of acoustic waves may be controled by applied electric field via the effects of electrostriction, nonlinear piezoeffect and nonlinearity of dielectric constant. The possibility of parametric amplification and generation of acoustic waves in alternating electric field of properly chosen frequency and amplitude is discussed together with the effects of mode interaction due to lattice and electric nonlinearities in HBAR.

5. Results of the use of HBAR structures (Al-ZnO-Al-YAG, AL-ZnO-Al-LGS, Al-ZnO-quartz(AT)) in various microwave oscillators.

---

* : Part of presented original results were obtained in LCEP/ENSMM, Besançon, France.

CESIUM STANDARDS

Chairman: *J. S. Boulanger*

EXPERIMENTAL MEASUREMENT OF THE SHIFT OF CS HYPERFINE SPLITTINGS DUE TO A STATIC ELECTRIC FIELD

E. Simon[1], P. Laurent[1], C.Mandache[2] and A.Clairon[1]
(1) BNM-LPTF, Observatoire de Paris, 61 Avenue de l'Observatoire, 75014 Paris, France.
(2) Institute for Atomic Physics, Nuclear Fusion Laboratory, R76900 Bucharest, Margurele, P.O Box MG-7, Romania.

Abstract

We have used the cold atom frequency standard LPTF FO.1 to perform an evaluation of the frequency shift induced by a static electric field on the different $\Delta M_F=0$ transitions of the cesium $6\,^2S_{1/2}$ hyperfine splittings. In particular, the frequency shift measured on the clock transition is $\delta\nu_0$=-2.271(4) x 10^{-10} $\mathcal{E}^2$ Hz/(V/m)^2. The interest of those results for a future evaluation of the frequency shift induced by the black body radiation in cesium frequency standards is discussed .

1. INTRODUCTION

The black body radiation induces a frequency shift of hyperfine splittings of $6\,^2S_{1/2}$ states of cesium atoms. It is shown that this shift can be deduced from the D.C hyperfine Stark constant and it was previously estimated at 1.76 x 10^{-14} with an accuracy of a few percents for a temperature of 300K (Ref. 1). At the present time, the uncertainty of this estimation is a predominant term in the error budget of the best cesium frequency standards (see Annex). We took advantage of the high stability and of the pulsed operation of the atomic fountain FO.1 (Ref. 2) to measure the D.C hyperfine stark shift 10 times more precisely than previously done (ref 5,6). This experiment will enable a more accurate estimation of the black body shift and should be of a great interest for a 10^{-16} range evaluation of the accuracy of FO.1 and of the future frequency standards using cold Cs atoms such as the PHARAO project (Ref. 3). We will first recall to which extend it is possible to deduce the black body shift from the knowledge of the frequency shift induce by a static electric field. Then, we will describe an experimental measurement of the D.C hyperfine Stark shift for Cs $6\,^2S_{1/2}$ (F=3, M_F=i)—(F=4, M_F=i) transition with |i|=0,1,2,3.

2. THEORETICAL CONSIDERATIONS

2.1 Effect of a static electric field

The application of a static electric field on cesium atoms shifts the energy levels of the clock transition by the Stark effect. The electric field induces a dipolar electric coupling between the $6\,^2S_{1/2}$ F=3 and F=4 levels and the $|n, P_J, F>$ states. From the second order perturbation theory, the frequency shifts of the $6\,^2S_{1/2}$ states (F=3 M_F=0) and (F=4 M_F=0) in presence of a static electric field $\mathcal{E}$ are (Ref. 4):

$$\Delta E_{(F=3,M_F=0)} = -\frac{e^2\mathcal{E}^2}{4\pi\varepsilon_0}\frac{\sum_{n,J,F}\left|\langle 6S_{1/2}3|z|nP_JF\rangle\right|^2}{(W_{nP_JF}-W_{6S_{1/2}3})} \quad (1)$$

$$\Delta E_{(F=4,M_F=0)} = -\frac{e^2\mathcal{E}^2}{4\pi\varepsilon_0}\frac{\sum_{n,J,F}\left|\langle 6S_{1/2}4|z|nP_JF\rangle\right|^2}{(W_{nP_JF}-W_{6S_{1/2}4})} \quad (2)$$

e is the electron charge, ε_0 is the vacuum dielectric constant and W_x is the energy of the x atomic state. The shift of the clock transition is given by the difference of (2) and (1) :

$$\delta\nu_0 = \Delta E_{(F=4,M_F=0)} - \Delta E_{(F=3,M_F=0)} \quad (3)$$

$$\delta\nu_0 = k_0 \times \mathcal{E}^2 \quad (4)$$

The constant k_0 of this differential Stark effect was previously measured (Ref. 5,6) with an accuracy of a few percents.

2.2 Effect of the black body radiation

According to the Planck radiation law, the average electric field radiated by a black body at a temperature T is given by :

$<E^2>$ = (831.9 V/m)^2 x [T(K)/300]^4 (5)

This electric field perturbs the clock frequency of cesium frequency standards (Ref. 7). We can neglect the shift induced by the magnetic field, which is 1000 times smaller at room temperature (Ref. 1). In the case of the atomic fountain FO.1, the thermal radiation seen by the atoms during the Ramsey microwave interaction is nearly the same as the radiation of a perfect black body except that a small correction of about 1% must be applied so as to take into account the emissivity and the geometry of the vacuum tube.

To deduce accurately the black body shift from the measurement of the frequency shift induced by a static electric field, we must take into account the frequency distribution of the electric field generated by a black body and its orientation which is a random function of time. Such a calculation was performed with an uncertainty of a few percents (Ref.1), and gives for a black body at a temperature T :

$$\delta\nu_0 = k_0 \times \varphi \times \left(\frac{T}{300}\right)^4 \times \left(1+\varepsilon\left(\frac{T}{300}\right)^4\right) \quad (6)$$

(k_0 = -2.271 10^{-10} Hz/(V/m)^2, ϕ=832 (V/m)^2 and ε=1.4 x 10^{-2})

This calculation is roughly similar to (1), (2), (3), (4) except that the formula for the A.C Stark shift is used. At a level of a few percent accuracy, only the first excited P state is assumed to make a significant contribution, and the orientation of the

electric field doesn't need to be taken into account. The uncertainty on this estimation is responsible for a $1x10^{-15}$ uncertainty on a Cs frequency standard operating at room temperature. In order to improve the theoretical estimation of the black body shift of one order of magnitude, it is necessary to perform a measurement of the frequency shift induced by a static electric field on the clock transition (F=3, M_F=0)—(F=4, M_F=0) with an accuracy of a few 10^{-3}. The same kind of measurement for the (F=3, M_F=i)—(F=4, M_F=i) transitions with |i|=1,2,3 would also be of a great interest so as to include the orientation of the electric field in the calculation. We have performed those measurements in the atomic fountain FO.1.

3. EXPERIMENT

3.1 Description

The atomic fountain FO.1 was previously described in details (Ref. 2). We recall here briefly the operation of FO.1 during the experiment on the differential Stark shift. A few 10^7 Cs atoms are cooled down in a magneto optical trap at a temperature of a few micro-Kelvin. They are launched in the vertical direction by the moving molasses method. They pass a first time in a TE_{011} microwave cavity which is 30cm above the cooling region. They follow a ballistic flight up to 45 cm above the microwave cavity, then they fall down and they experience the second part of the Ramsey microwave interaction. The population of both hyperfine levels is measured by fluorescence in the bottom of the experiment. A vertical homogeneous magnetic field B_0 of magnitude $1.63x10^{-7}$ T resolves the field sensitive transitions. Two copper plates separated by 2 cm long ceramic interstices have been hanged 17 cm above the microwave cavity. The dimensions of the plates are 10x30x1,5 cm^3. They have been manufactured with a planeity of 10µm. An opposite voltage can be applied on the plates in order to generate an horizontal electric field from 0 to 3000 V/cm. A computation of the electric field was performed by the L.G.E.P* . This computation gives a map of the electric field along the pass of the atoms and it demonstrates that the homogeneity of the field intensity is better than 10^{-3} in a 6x26x1 cm^3 parallelepiped centered in between the two plates.

3.2 Evaluation of k_0

We first consider the clock transition (F=3, M_F=0)—(F=4, M_F=0). We have measured the frequency shift $\delta\nu$ of the atomic resonance fringe when an electric field $\mathcal{E}$ is applied between the copper plates over a span from 0V/cm to 3000V/cm. It gives a maximum shift $\delta\nu$ of about 20Hz. A hydrogen maser has been used as a comparison oscillator. When an electric field is applied, the average atomic frequency between the copper plates is different from the atomic frequency inside of the cavity. In that case, the resonance fringe is not symmetrical and a small correction α of about 1% must be applied to the frequency of the center of the fringe so as to deduce the average atomic frequency between the two microwave interactions (Ref. 8). According to this consideration and to the equation (4), we obtain :

$$\delta\nu_0 = k_0 \times \frac{\int_{t_1}^{t_2} \mathcal{E}^2 dt}{t_2 - t_1} \qquad (7)$$

$$\text{with } \delta\nu_0 = \delta\nu \times (1+\alpha) \qquad (8)$$

The frequency shift due to the electric field is averaged between the time t_1, when the atoms have just experienced the first microwave interaction and t_2, when the atoms are starting to experience the second microwave interaction. We have measured $\delta\nu$ with an accuracy better than 10^{-4}. We estimate $\delta\nu_0$ from (8) with an accuracy of 10^{-3}, which is mainly limited by the experimental uncertainty of the microwave power which α depends on.

The evaluation of k_0 depends on the accuracy of the calculation of the integral term in (7). The geometrical uncertainties on the relative position of the microwave cavity and of the copper plates limits the accuracy on the evaluation of the integral term. We have used two different experimental processes to reduce this uncertainty at a level of 10^{-4}. We have first performed measurements of the frequency shift induced by a fixed value of the electric field when the atoms are launched at different heights above the microwave cavity. An analysis of those results using the computation of the electric field enables to reject the influence of the relative position of the microwave cavity and the copper plates. The second experimental process consists in switching on the voltage on the copper plates during a time ΔT when the atoms are passing in the highly homogeneous electric field region centered in between the two plates. In this case, expression (7) becomes :

$$\delta\nu_0 = k_0 \times \mathcal{E}^2 \frac{\Delta T}{t_2 - t_1} \qquad (9)$$

The switching times of the voltage are recorded on a digital oscilloscope, and a small correction of their parasitic effect ($2x10^{-3}$) is applied on the calculation (9). Those two previous measurements are in good agreement at a level of $5x10^{-4}$, the main remaining source of the geometric uncertainty comes from the relative position of the cooling region of the Cs atoms and of the microwave cavity, this effect causes a 10^{-3} uncertainty on the calculation of k_0.

We have increased the value of the vertical magnetic field B_0 in FO.1 by a factor of 2 so as to see if there were any dependence of $\delta\nu$ on the value of B_0. We observed no variations at a level of a few 10^{-4}. The influence of the motional magnetic field seen by the cesium atoms moving in a static electric field was calculated to be negligible for a 10^{-3} accurate estimation of k.
The main uncertainties on the measurement of k_0 are listed in the following table :

* Laboratoire de Génie Electrique de Paris, Plateau du Moulon, Gif sur Yvette

Physical origin	Uncertainty x 10^3
fringe asymmetry	1
relative position of the cooling region and the microwave cavity	1
voltage applied on the plates	0.3
distance between the copper plates	0.5
velocity distribution of the atoms	0.3
frequency stability of FO.1	<0.1
computation of the electric field	0.5
Total uncertainty (1σ)	**2**

3.3 Zeeman sublevels

The same kind of measurements of the frequency shift induced by an electric field were performed on the (F=3, M_F=i)—(F=4, M_F=i) transitions with i=-3,-2,-1,1,2,3. We have accumulated about ten measurements of the frequency shift induced for each magnetic quantum number M_F, and we have observed long term fluctuations of up to 2% for given values of M_F. We assumed that those fluctuations were due to a magnetic field generated by parasitic currents induced when a voltage was applied on the plates. Such fluctuations should be of opposite sign for the (F=3, M_F=i)—(F=4, M_F=i) and the (F=3, M_F= -i)—(F=4, M_F= -i) transitions. We have performed a simultaneous measurement of the frequency shift induced by an electric field $\mathcal{E}$ on each pair of transitions (|i|=1,2,3). The average shift for each pair of transitions didn't fluctuate more than 0,4%. We give those experimental results in the next part with the following abbreviations :

$$\delta\nu_{(F=3,M_F)-(F=4,M_F)} = k_{M_F} \times \mathcal{E}^2 \qquad (10)$$

$$k_{|M_F|} = \frac{k_{M_F} + k_{-M_F}}{2} \qquad (11)$$

4. RESULTS

In the case of the clock transition, the proportionality of the frequency shift with the square of the electric field given by (4) has been verified with an accuracy of 5×10^{-4} for 0V/cm < $\mathcal{E}$ < 3000V/cm. We have found:

k_0= -2.271(4) $\cdot 10^{-10}$ Hz/(V/m)^2

This result is in agreement with the previous measurements :

k_0= -2.29(3) $\cdot 10^{-10}$ Hz/(V/m)^2 (Ref. 5)
k_0= -2.25(5) $\cdot 10^{-10}$ Hz/(V/m)^2 (Ref. 6)

The different Zeeman sublevels of the $6\,^2S_{1/2}$ hyperfine splittings were resolved with a vertical magnetic field B_0 of magnitude 1.63×10^{-7}T. Using the convention of § 3.3, we have obtained:

magnetic quantum number $\|M_F\|$	$k_{\|MF\|}/k_0$
0	1
1	1.002(3)
2	1.006(3)
3	1.010(4)

5. CONLUSION AND ACKNOWLEDGMENTS

We have performed an evaluation of the shift of the cesium clock transition due to an electric field. Those experimental data should enable a future theoretical evaluation of the black body shift in Cs frequency standards with an accuracy of 10^{-16}. In the future, we will replace the copper plates by a graphite tube heated from the room temperature to 450K so as to perform a direct measurement of the black body shift. That kind of measurement was recently performed in a cesium beam frequency standard (Ref. 7). It gave a result in good agreement with the theoretical calculation (Ref. 1) with an uncertainty of 10%.

We acknowledge the assistance of A.Gerard, P.Aynié, M.Lours, M.Dequin, L.Volodimer, all from the BNM-LPTF. We wish to thank M.L.Mathieu , O.Mathieu and L.Pichon for the helpful computations performed at the L.G.E.P*.

6. ANNEX

We give the last accuracy evaluation of the atomic fountain of the BNM-LPTF :

Physical correction origin	Correction x10^{15}	Uncertainty x10^{15}
Doppler effect	0	0.5
Black body radiation shift	**17.6**	**1**
Quadratic Zeeman effect	-133	0.1
Electronics	0	<1
Cold collisions	1	0.5
Ramsey, Rabi and Majorana pulling	0	0.5
Spectral impurities shift	0	0.2
μ-W leakage	0	0.5
Background gas collisions	0	0.5
Total Uncertainty (1σ)		**2**

7. REFERENCES

1. W.M. Itano, L.L. Lewis, and D.J. Wineland, Phys Rev. **A25**, 1233 (1982).

2. A. Clairon, S. Ghezali, G. Santarelli, Ph. Laurent, S.N. Lea, E. Simon, S. Weyers, and K. Szymaniec, in Procedings of the fifth Symposium on Frequency Standards and Metrology 1995, (J.C. Bergquist, ed., World Scientific Publ., Singapore, 1996), p.49.

3. P. Lemonde, P. Laurent, E. Simon, G. Santarelli, A. Clairon, C. Salomon, N. Dimarcq, P. Petit, C. Audoin, F. Jamin Changeart, F. Gonzalez, PHARAO : A Compact Clock Using Laser Cooled Atoms, Proc. Of the 11th EFTF, Neuchâtel (Switzerland).

4. J.D. Feichtner, M.E. Hoover, and M. Mitsushima, Phys. Rev. **137**, A702 (1965).

5. R.D. Haun and Zacharias, Phys. Rev. **107**, 107 (1957).

6. J.R. Mowat, Phys. Rev. **A5**, 1059 (1972).

7. A. Bauch, R. Schröder, Phys. Rev. Lett. **78**, p.662 (1997).

8. J. Vanier and C. Audoin, The Quantum Physics of atomic Frequency Standards, (Adam Hilger, Bristol and Philadelphia, 1989), p802-804.

SHIFT OF THE CAESIUM HYPERFINE TRANSITION FREQUENCY DUE TO BLACKBODY RADIATION: EXPERIMENTAL VERIFICATION AND RELATED STUDIES

Rolf Augustin, Andreas Bauch, Roland Schröder

Physikalisch-Technische Bundesanstalt, Time-Unit Laboratory
Bundesallee 100, D-38116 Braunschweig, Germany

ABSTRACT

The dynamic Stark shift of the ground-state hyperfine transition frequency in caesium (≈9.2 GHz) due to the electric field of blackbody radiation was measured as a function of the temperature of heated surfaces surrounding the atomic beam of a frequency standard. From our experimental data, a relative frequency shift of $-17{,}9(1{,}6)\cdot10^{-15}$ at room temperature is derived, in good agreement with the theoretical prediction $-16{,}9\cdot10^{-15}$.

Keywords: caesium atomic clocks, blackbody radiation

1. INTRODUCTION

The interaction between atoms and the electric fields of blackbody radiation (BBR) produces two effects: Firstly, it drives atomic transitions and thereby shortens the lifetime of atomic states, an effect which has been clearly observed for Rydberg states (Ref. 1) and also for metastable states (Ref. 2), which both are strongly susceptible to this kind of perturbation (Ref. 3). Secondly, a dynamic Stark shift of the atomic energy levels is induced which can be as large as about 2,5 kHz for Rydberg states (principal quantum number n > 30) (Ref. 3,4). Here we address the minute effect of BBR on atomic ground state levels. Historically, in 1982, Itano, Lewis and Wineland pointed out that the electric field of isotropic blackbody radiation emitted from the surroundings of the atomic beam in a caesium atomic clock leads to a shift of the clock frequency in the same order of magnitude as the standard uncertainty of primary clocks of that time (Ref. 5). Despite of the elemental interest into the matter, for about 14 years the frequency shift has not been verified experimentally, probably because of the technical difficulties encountered with such an experiment. The interest has been revitalized recently, when the first fountain frequency standard employing cold caesium atoms, the FO1 (Ref. 6) became operational: The predicted frequency shift (Ref. 5) amounts to about 6 times the standard uncertainty of the FO1.

In this contribution we report on the final evaluation of our experimental data. Compared to a previous analysis (Ref. 7) the number of data points is enlarged. The experimental procedures are outlined in more detail, and emphasis is laid on the use of the time scales UTC(USNO) and UTC(SU) as long-term stable reference scales, alternately to the use of UTC(PTB). We also give an interpretation of the observed impact of the cavity dimensions on the cavity end-to-end phase difference.

2. THEORETICAL BACKGROUND

According to Planck's radiation law, at a given temperature T, the time averaged quadratic electric field strength of BBR is

$$\langle E^2(t)\rangle = (832\ \mathrm{V/m})^2 \cdot \{T/(300\ \mathrm{K})\}^4 \qquad (1)$$

and the associated magnetic flux density is

$$\langle B^2(t)\rangle = (2{,}8\ \mu\mathrm{T})^2 \cdot \{T/(300\ \mathrm{K})\}^4 . \qquad (2)$$

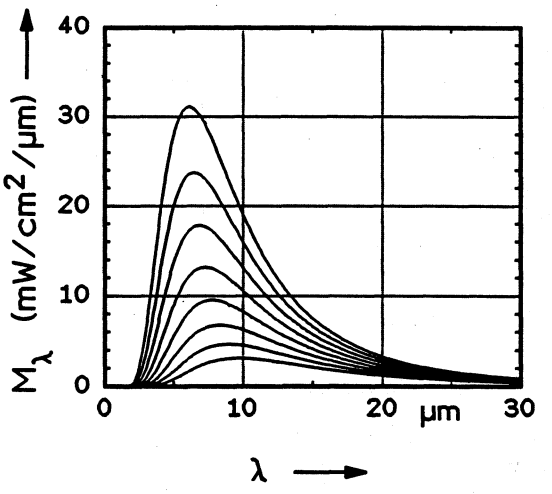


Fig. 1: Spectral radiant excitance M_λ (emitted power per surface element per wavelength interval) of a perfect black surface versus wavelength λ as a function of temperature T, T = 300 K for the lowest curve, T = 475 K for the upper curve, stepsize 25 K.

At room temperature, BBR has its peak spectral density around a wavelength of 9 μm. In Fig. 1 the spectral radiant excitance M_λ of a perfect Planckian radiator (black body) is depicted as a function of the temperature T of its surface. It has been known for some time that the hyperfine splitting interval in the ground state of caesium (^{133}Cs) is reduced in the presence of a static (DC) electric field. In the theoretical treatment (Ref. 8-10) which followed the experiments made by Haun and Zacharias (Ref. 11) and Mowat (Ref. 12) the frequency shift is attributed to the differential polarizability of the two ground state hyperfine levels. Mowat's results on the DC Stark shift had a relative uncertainty of ≈ 0,5 % and the theoretical results, in particular those by Lee et al. (Ref. 10), agreed with Mowat's results within 1 %. The DC Stark shift appears in this case as a third-order

perturbation, and the relevant terms (eq. (15) - (18) in Ref. 8) describe the admixture of excited S and P states to the ground state and include the frequencies of the electric dipole transitions connecting the ground state to excited states as resonance denominators. These frequencies are much higher than the bulk spectral distribution of BBR and thus at room temperature BBR represents a *slowly varying perturbation* (Ref. 3) to the ground state hyperfine levels. The frequency shift in the dynamic case can thus be calculated using the rms. value of the electric field strength of BBR. Including a corrective term χ which accounts for the separation in frequency between the BBR spectrum and the transition frequency of the D1 and D2 line in ^{133}Cs, one expects the following frequency shift (Ref. 5):

$$y_{BBR}: = [\nu(T) - \nu_0] / \nu_0 = \eta \cdot \chi \cdot \{T/(300\ \mathrm{K})\}^4 \quad (3)$$

with $\eta = -16{,}9 \cdot 10^{-15}$ and $\chi = 1 + 0{,}014 \cdot \{T/(300\ \mathrm{K})\}^2$.

Here ν_0 is the unperturbed hyperfine transition frequency of ^{133}Cs, ν_0 = 9 192 631 770 Hz, $\nu(T)$ is the clock frequency of atoms which are subjected to radiation of a blackbody at temperature T, and y_{BBR} is the predicted frequency shift expressed as a relative quantity. The numerical factor η is based on eq. (1) and on the experimental results on the DC Stark effect (Ref. 12). It is assumed that the perturbing BBR is isotropic and unpolarized. The effect of the BBR magnetic field (2) through the dynamic Zeeman effect leads only to a frequency of a few parts in 10^{17} and is thus at present not of particular interest (Ref. 5).

3. EXPERIMENTAL PROCEDURE

3.1 Apparatus, Source of BBR

PTB's experimental atomic beam frequency standard CSX was used to observe the frequency shifting effect of BBR (Ref. 13,14). Briefly, the CSX is a frequency standard with magnetic state selection in hexapole and quadrupole magnetic lenses. The atomic beam, 3 mm in diameter, shows axial symmetry. The mean velocity of the atoms contributing to the hyperfine resonance signal is 405 m/s. Due to the 0,79 m interaction length the transition linewidth amounts to 253 Hz. During our experiments, the atomic flux was chosen such that the short-term frequency instability, expressed by the Allan standard deviation $\sigma_y(\tau)$, was $3{,}5 \cdot 10^{-12} \cdot (\tau / \mathrm{s})^{-1/2}$.
Equation (3) predicts a relative frequency shift of only about 10^{-13}, even if the BBR source temperature is increased by as much as 200 K above room temperature. It seemed impossible to observe the effect at temperatures close to room temperature. Thus we partially surrounded the atomic beam path within the interaction region of the CSX with tubes which could be heated up to 200 K above room temperature. Figure 2 shows schematically the experimental set-up with some details of the heated tubes and the temperature measurement. For technical reasons the heated region was splitted into two parts, each 0,25 m in length. The

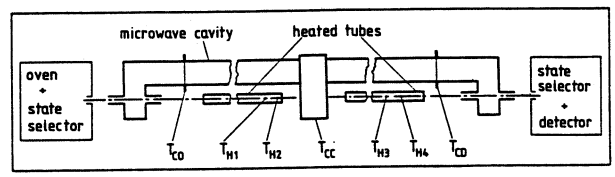


Fig. 2: Sketch of the experimental set-up. The standard constituents of the CSX are shown very schematically, whereas the microwave cavity and the heated tubes are shown in larger detail. T_{H1} - T_{H4} and T_{CO}, T_{CC} and T_{CD} indicate the positions of the thermocouples attached to the heated tubes and the cavity, respectively.

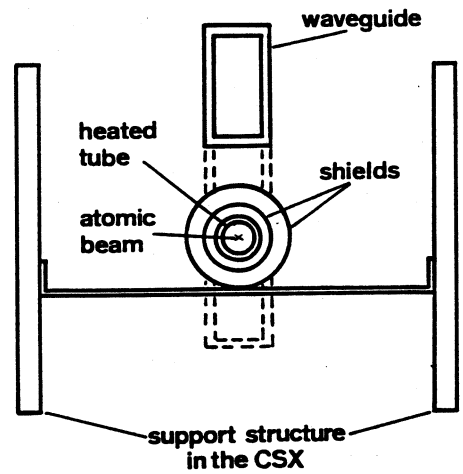


Fig. 3: Axial view onto one of the heaters in the CSX. Thin ceramic spaces separate the central tube from the two reflective shields. All metallic parts are electrically insulated from one another to avoid the flow of thermocurrents.

tubes were surrounded by two coaxial reflective shields and were suspended with low thermal conductance from the support structure inside the CSX as shown in Fig. 3. This was done in order to reduce the heat load on the microwave cavity which finally warmed up by only 12 K when the heated tubes were at 485 K, the maximum temperature used. Heating was performed by passing DC currents through four wires along the tubes. This had to be done without destroying the homogeneity of the weak quantization field which is applied parallel to the atomic beam. We chose a cycling mode, separating 2 or 3 minutes of heating from 9 minutes of taking frequency data. During the latter interval the DC current was switched off and the temperature of the tubes dropped by a few Kelvin exponentially with a time constant of 80 minutes. Nine-minutes time averages of the temperature were recorded using the four thermocouples shown in Fig. 2 from which the spatial temperature distribution along the heated tubes was obtained, as shown in Fig. 4.

The inner surface of the tubes was painted with 3 M Nextel Black Paint (Ref. 15) and served as the source of thermal radiation. The emissivity of the surface, averaged over the BBR spectrum, is estimated from

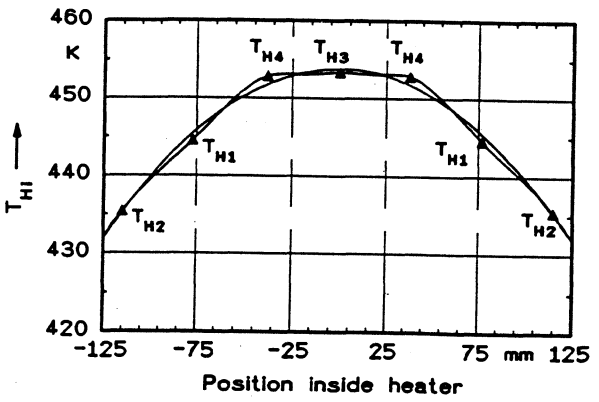


Fig. 4: Typical spatial temperature distribution inside each of the tubes, estimated from the measurements made with the four thermocouples T_{H1} - T_{H4} which were distributed along both tubes. The temperature decrease towards the ends is attributed to radiative loss through the open shields. The solid lines represent a second-order polynomial and a spline-function, respectively, fitted to the measurement points. The spectral emittance obtained by integrating along both approximations of the temperature distribution agreed to better than 2% in all cases.

data on the hemispherical spectral reflectance of the paint at room temperature (Ref. 16) to be 0,961(5). The atomic beam passes along the axis of the tubes. The on-axis intensity of BBR was calculated by integrating over the spectral emittance of the surface of the tubes. The emittance was determined from the spatial temperature distribution (T^4) and the assumption of BBR emission according to Lambert's law of cosines. This assumption is reasonably well fulfilled for the paint (Ref. 16). BBR is concentrated to the region inside the two heated tubes as illustrated in Figure 5. The time-averaged temperature in the central section of a tube is named T_{BBR} furtheron. Based on the BBR intensity profile (Fig. 5), on the temperature T_{BBR} and on eq. (3), we obtained prediction values of the frequency shift,

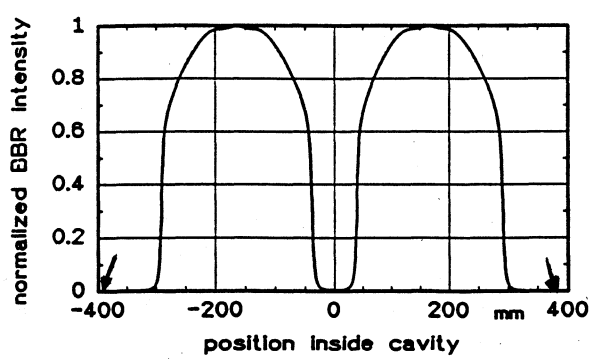


Fig. 5: Normalized intensity of blackbody radiation which is emitted from the inner surface of the heated tubes in the interaction region of the CSX. Outside the tubes, the BBR intensity is assumed to correspond to the temperature ofthe CSX vacuum tank. The two arrows indicate the beginning and the end of the drift region of the cavity (see Fig. 2).

e.g. $y_{BBR} = -69 \cdot 10^{-15}$ at the highest temperature used during the experiments. The systematic uncertainty of the prediction is estimated to about 5 %, relatively, based on the uncertainty of the temperature measurement (absolute value and profile), of the emissivity of the surface of the tubes and of η in eq. (3) (Ref. 5,12).

3.2 Frequency Reference

Frequency measurements were made using predominantly PTB's primary clock CS2 as a reference (Ref. 17), and the relative frequency difference

$$y_{XR}: = [\nu(CSX) - \nu(CS2)] / \nu(CS2) \qquad (4)$$

was recorded. Here ν(CS2) and ν(CSX) are the clock transition frequencies of the CS2 and CSX, respectively, corrected for all systematic frequency shifts (Ref. 17,18) except for the CSX cavity phase difference (see below). The combined frequency instability was $\sigma_y(\tau$ = 1d) = $18 \cdot 10^{-15}$ and $\sigma_y(\tau$ = 6d) = $7 \cdot 10^{-15}$ during normal CSX operation (data points at $T_{BBR} \approx$ 300 K in Fig. 8) and thus at the level which is estimated from the known noise sources in the clocks.

It has been shown that time scales which are based on a huge number of caesium clocks, like UTC(USNO), or on an ensemble of active hydrogen masers, like UTC(SU), are in general more stable than UTC(PTB) which is derived directly from an excellent, yet one single clock, the CS2. This holds for averaging times up to several days. As our basic measurement interval was typically 6 days we felt inspired to refer our frequency measurements alternately to these particularly stable frequency references. GPS observation data from United States Naval Observatory, Washington, and VNIIFTRI, Mendeleevo, Russia, which were recorded following the BIPM common view tracking schedules were used for this purpose. To give an example, a rate difference between UTC(USNO) and UTC(PTB) of $11 \cdot 10^{-15}$ was obtained from 36 6-day intervals with a standard deviation of $8 \cdot 10^{-15}$. Similar results were obtained against UTC(SU), but unfortunately just during our most productive period of data taking the GPS receiving equipment at VNIIFTRI failed and some data points have been lost. The mean rate differences can be ignored for our purpose. They are e.g. buried in the mean frequency difference at T = 300K in Fig. 9 and they show up in offsets between corresponding data points in Fig. 6.

3.3 Accompanying Studies: The Impact of Cavity Dimensions on the CSX Frequency

Observation of the predicted small effect required a careful control of all systematic frequency shifts aside from the BBR effect. In particular we studied the impact of the heat generated inside the apparatus on the clock's performance. It was known beforehand that the realized cavity design would entail a significant sensitivity of the end-to end phase difference ϕ of the cavity on its dimensions - in other words - on its temperature - much larger than the traditionally

considered "cavity pulling" effect (Ref. 18) which is of order 10^{-15} only in our case.
We decided to study separately the impact of both a temperature gradient along the cavity and the mean cavity temperature on the CSX clock frequency. Three thermocouples were attached to the waveguide (see Fig. 2 for the notation). At first, we introduced a large temperature gradient along the cavity by heating only one of the tubes at a time. Thereby we realized equal thermal radiation fields subsequently, and we were able to determine the frequency sensitivity of the CSX on temperature gradients along the cavity without a priori knowledge of the BBR frequency shift. Fig. 6 depicts results obtained for various electrical powers dissipated in the tubes. The procedure of utilizing foreign frequency references was successful in this case. From all data, i.e. the slope of connecting lines in Fig.6, including data with CS2 as a reference, we inferred that the CSX clock frequency ν(CSX) decreased by $276(25)\cdot10^{-15}$ when the temperature difference T_{CO}-T_{CD} was increased by 1 K.

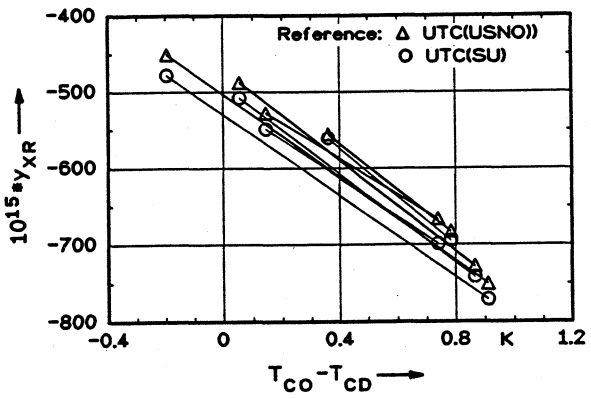


Fig. 6 CSX frequency y_{XR} in dependence on the temperature difference between oven and detector side of the cavity. Data shown here are referenced against UTC(USNO) and UTC(SU). Lines connect data points which were obtained with equal temperature in each of the heated tubes.

At next, we analysed the influence of the mean cavity temperature. Here, we had to separate between the effect of the dimensional sensitivity of ϕ and the effect of BBR under study. We varied the cavity temperature by up to 15 K simply by adjusting the temperature in the CSX laboratory between 288 K and 303 K. As intensity and spectrum of BBR inside the CSX were not significantly altered, the observed frequency shift were attributed to the change of ϕ. The clock frequency was decreased by 6,9 (0,8) $\cdot10^{-15}$ in relative units when T_{CC} was increased by 1 K. In the next section we sketch an explanation of these findings in terms of the cavity parameters, which will be given in more detail in a forthcoming paper.

3.4 Dimensional sensitivity of the end-to-end phase difference of a Ramsey cavity

In order to calculate the phase of the magnetic component of the microwave field at the location of

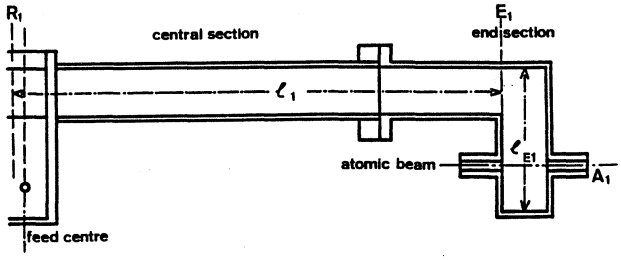


Fig.7 Schematical representation of one arm of the CSX Ramsey cavity. It includes the planes of reference R_1, E_1 and A_1 where the atomic beam passes the cavity at a place where the longitudinal field component is maximum.

interaction with the atomic beam we calculate the phasors at various points inside the cavity. Fig. 7 may serve for illustration of our procedures which were initiated by earlier work of De Marchi et al. (Ref. 19). Two reference planes R_1 and R_2 exist where the field components are equal and in anti-phase. They are symmetrically located with respect to the central H plane feed. The plane E_1 defines the beginning of the downwards section of arm 1. During manufacture the dimensions have been carefully adjusted so that E_1 coincides with a plane of maximum field strength. For the phase in plane E_1, with respect to R_1, we can write

$$\phi(E_1) - \phi(R_1) = \arctan\{ \operatorname{tg} \xi_1 \cdot (1 - | \Gamma_1 |) /2 \} \approx \tan\xi_1 \cdot (1 - |\Gamma_1|) / 2,$$

with the following notations

- electrical length, $\xi_1 = \beta \ell_1 + \varphi_1$
- propagation constant, $\beta = 2 \pi / \lambda$
- wavelength of the 9,2 GHz field in the standard X-band waveguide, λ
- argument of the complex reflection factor of end section 1 as seen in plane E_1, $2 \varphi_1$
- reflection factor as seen in plane R_1, $\Gamma_1 = |\Gamma_1| \exp \{ -j 2 \xi_1\}$, with $|\Gamma_1| = \exp \{-2 \alpha_1 L_1\} \approx 1 - 2 \alpha_1 L_1$ and the total length $L_1 = \ell_1 + \ell_{E1}$ (see Fig.7),
- damping constant along arm 1 of the cavity, α_1.

In addition we have to allow for a phase difference

$$\phi(A_1) - \phi(E_1) : = \phi_{AE1} .$$

Equivalent relations hold for arm 2 of the cavity. With the above relations, and taking into account the physical rotation by $\pi/2$ of the magnetic fields in each of the two waveguide corners of the cavity, the end-to-end phase difference $\phi(A_1) - \phi(A_2) : = \phi_{12}$ can be calculated. Including reasonable approximations we obtain

$$\phi_{12} = \alpha_1 L_1 \tan\xi_1 - \alpha_2 L_2 \tan\xi_2 + \phi_{AE1} - \phi_{AE2} . \quad (5)$$

As an approximation to our experimental procedure, we consider at first the case that the temperature along arm 1 is increased by 1K with reference to arm 2. The additional phase difference $d\phi_{12} = \alpha_1 L_1 d\xi_1 / (\cos^2 \xi_1)$

occurs, with $d\xi_1 = d(\beta\ell) + d\varphi_1$. Now, and in the following, we assume that $d\varphi_1$ can be included in the increase of the total length L_1 and obtain
$d\xi_1 = \beta L_1 \cdot [(d\ell / \ell) + (d\beta / \beta)]$. As the width of the waveguide is close to $\lambda/2$ in our case,
$(d\ell / \ell) \approx (d\beta / \beta)$ is valid, and we can write for the additional phase difference

$$d\phi_{12} = [2 \cdot \alpha_1 \cdot \beta \cdot (L_1)^2] (d\ell / \ell) / (\cos^2 \xi_1) .$$

We have the numerical values $L_1 = 19{,}5 \cdot \lambda/2$, $\lambda =$ 46,53 mm, $\cos^2 \xi_1 = 0{,}5$, the conductivity and thermal length expansion coefficient for the waveguide material copper, $\alpha_1 = 1{,}5 \cdot 10^{-5}$ /m and $(d\ell / \ell) = 1{,}7 \cdot 10^{-5}$, respectively, and finally obtain $d\phi_{12} = 2{,}84 \cdot 10^{-5}$ rad, corresponding to a relative CSX frequency shift of $2{,}65 \cdot 10^{-13}$. We are thus able to explain the order of magnitude of the observed temperature-gradient effect mentioned above.

At next we consider a homogeneous heating of the whole cavity structure. We rewrite (5) to

$$\phi_{12} = (2 \overline{\alpha L}\, \tilde{\xi}) / \cos^2 \bar{\xi} + 2 \tilde{\alpha L} \tan \bar{\xi} + \phi_{AE1} - \phi_{AE2}$$
$$:= \phi_\xi + \phi_\alpha + \phi_{AE1} - \phi_{AE2} \qquad (6)$$

with $\overline{\alpha L} = (\alpha_1 L_1 + \alpha_2 L_2) / 2$, $\bar{\xi} = (\xi_1 + \xi_2) / 2$
and $\tilde{\alpha L} = (\alpha_1 L_1 - \alpha_2 L_2) / 2$, $\tilde{\xi} = (\xi_1 - \xi_2) / 2$.

We notice now that a temperature influence on ϕ_{12} comes in essentially through the terms which contain the long dimensions of the cavity, in particular through $\bar{\xi}$. For the infinitesimal increment $d\phi_{12}$ we obtain

$$d\phi_{12} = 2\, \overline{\alpha L}\, \tilde{\xi}\, d(1/\cos^2 \bar{\xi}) + 2\, \tilde{\alpha L}\, d(\tan \bar{\xi}),$$

and, with $\tan \bar{\xi} = 1$, $\cos^2 \bar{\xi} = 1/2$,

$$d\phi_{12} = (\phi_\xi + \phi_\alpha)\, 2\, d\xi = (\phi_\xi + \phi_\alpha)\, 4\, \beta L_1 (d\ell / \ell). \qquad (7)$$

If we assume that the mean frequency difference observed between the CSX and the CS2 (data points around $T = 300$K in Fig. 8) is entirely due to the sum $\phi_\xi + \phi_\alpha$ and ignore the term $\phi_{AE1} - \phi_{AE2}$ which describes the effect of the end sections of the cavity, (7) predicts a relative CSX frequency decrease of $2{,}7 \cdot 10^{-15}$ when the cavity temperature is increased by 1K. This is smaller than the observed effect by a factor of 2,5. A possible explanation would be that the first two summands in (6) are in reality larger and are partially compensated by the difference $\phi_{AE1} - \phi_{AE2}$ of opposite sign. During the next months we will try to verify this assumption.

4. Final Result

Taking the experimental data from section 3.3, we were able to calculate and apply corrections to all data taken when BBR was admitted and to eliminate the unwanted effects to a large extent. Figure 8 depicts the frequency measurement results y_{XR} as a function of the relevant

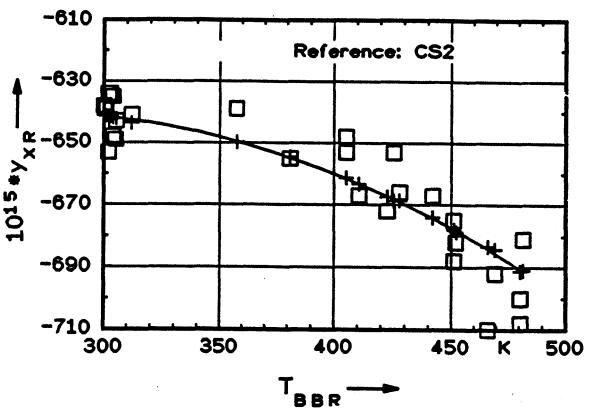


Fig. 8 Experimental frequency data y_{XR} (□), referenced to CS2, as function of temperature T_{BBR} in comparison with the predicted frequency shifts (+) due to BBR. The solid line is a polynomial through the prediction values, facilitating the interpretation of the results

temperature T_{BBR} which were corrected to demonstrate only the effect of BBR. Each point represents a 5- to 7-day average value. The data are shown together with the predicted frequency shifts which were normalized to the experimental results at room temperature. The predicted decrease of the clock frequency with increasing temperature T_{BBR} is clearly confirmed. Not surprising, the same quality of agreement is found in the data referenced to UTC(USNO), which are depicted in Fig.9. A quantitative analysis was made as follows. The difference between experimental data and the predicted frequency shift should be scattered around a line with slope zero. A significant deviation of the slope from zero would indicate that equation (3) or the numerical values involved were incorrect.

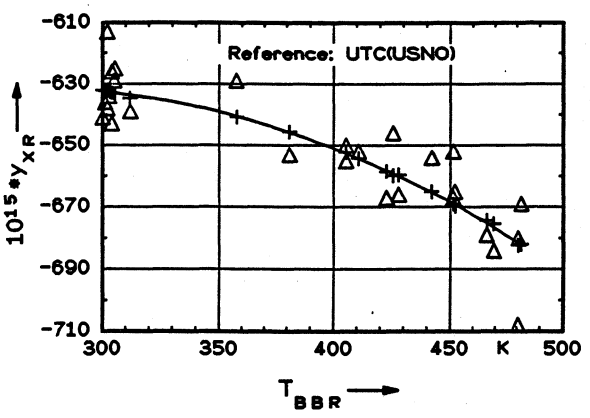


Fig.9 Experimental frequency data y_{XR} (Δ), referenced to UTC(USNO), as a function of temperature T_{BBR} . See Fig. 8 for further explanations.

As there is no reason to question the T^4-dependence of the frequency shift which follows from Planck's radiation law, we interpret the results in terms of η in eq. (3). Thus we determined an experimental value, η_e, to replace η in equation (3) as follows:

η_e = -17,9 (1,6) $\cdot 10^{-15}$ (Ref. CS2)

η_e = -17,2 (2,0)$\cdot 10^{-15}$ (Ref. UTC(USNO))

The rather large uncertainty in η_e is given by the frequency deviation of the individual data points around the regression line which corresponds to a frequency instability of 9,5 $\cdot 10^{-15}$ with CS2 as the reference and 10,8 $\cdot 10^{-15}$ with UTC(USNO) as the reference.

5. Discussion

From a fundamental point of view the realization of the time unit second has to rely on the observation of unperturbed atoms. It is thus indispensable to study experimentally even subtle effects like the one related with blackbody radiation. Within our statistical measurement uncertainty our experimental data confirm the previous theoretical predictions which were based on older DC Stark-shift measurements. We obtained an uncertainty larger than that of the previous DC measurements (Ref. 11, 12) and the difference between our experimental and the previously accepted value of η is insignificant. In our case the uncertainty is dictated by the limited number of data points and the frequency instability of the CSX and the reference clock. The frequency scatter of 9,5$\cdot 10^{-15}$ of the individual data points around the regression line is about 35 % larger than observed in normal operation of the CSX , i.e. without heating the tubes. We have to suspect that the changes of ϕ with temperature could not be estimated accurately enough using the recorded temperatures. The findings sketched in Section 3.4 point to the fact that a more detailed knowledge of the temperature distribution along the cavity would have been desirable. It would be even more advantageous if a beam reversal could be made in the CSX, which is the traditional method to measure ϕ in primary clocks.

Yet, our results justify the current steering process of the scale unit of International Atomic Time TAI. During its 13th meeting in March 1996 the Comité Consultatif pour la Définition de la Seconde adopted Recommendation S2, stating that a frequency correction compensating for the BBR frequency shift shall be applied to the reading of primary clocks. Measures were taken to adjust the scale unit of International Atomic Time TAI to the corrected duration of the SI second in a smooth fashion. The new steering of TAI started in March, 1995, as documented in the Circular T86 of BIPM. The steering in general is described in Ref. 20.

The frequency shift due to the static or the dynamic Stark effect in a caesium clock is independent of the atomic transition line width and thus its determination is a challenging task in view of the potential standard uncertainty of fountain type atomic clocks. As mentioned above, the correction to be applied at room temperature is about 6 times larger than the presently estimated standard uncertainty of the FO1 (Ref. 6). Its determination will have to include a detailed analysis of the temperature and the emissivity of the inner walls of the vacuum chamber of the device. The uncertainty in the numerical factor η is not at all negligible. Results of a new measurement of the frequency shift due to static electric fields have been announced for the 11th EFTF (see this Proceedings). A second measurement of the BBR effect with reduced uncertainty appears desirable.

Acknowledgement

We gratefully acknowledge the support of our colleagues at USNO and VNIIFTRI who made their GPS data available to us. Thomas Heindorff, Harald Brand and Christoph Richter of PTB contributed substantially to the success of our experiment.

References

/1/ T.F. Gallagher and W. E. Cooke, Phys. Rev. Lett. **42**, 835 (1979).

/2/ M. Walhout, U. Sterr, A. Witte, and S. L. Rolston, Optics Letters **20**, 1192 (1995).

/3/ J. W. Farley and W. H. Wing, Phys. Rev. **A23**, 2397 (1981).

/4/ L. Holberg and J. L. Hall, Phys. Rev. Lett. **53**, 230 (1984).

/5/ W. M. Itano, L. L. Lewis, and D. J. Wineland, Phys Rev. **A25**, 1233 (1982).

/6/ A. Clairon, S. Ghezali, G. Santarelli, Ph. Laurent, S. N. Lea, M. Bahoura, E. Simon, S. Weyers, and K. Szymaniec, in *Proceedings of the Fifth Symposium on Frequency Standards and Metrology 1995*, (J. C. Bergquist, ed., World Scientific Publ., Singapore, 1996), p. 49.

/7/ A. Bauch, R. Schröder, Phys. Rev. Lett **78**, 662 (1997).

/8/ J. D. Feichtner, M. E. Hoover, and M. Mizushima, Phys. Rev. **137**, A702 (1965).

/9/ P. G. H. Sandars, Proc. Phys. Soc. **92**, 857 (1967).

/10/ T. Lee, T. P. Das, and R. M. Sternheimer, Phys. Rev. **A11**, 1784 (1975).

/11/ R. D. Haun and J. R. Zacharias, Phys. Rev. **107**, 107 (1957).

/12/ J. R. Mowat, Phys. Rev. **A5**, 1059 (1972).

/13/ T. Heindorff, A. Bauch, and R. Schröder, PTB-Mitteilungen 94, 318 (1984).

/14/ A. Bauch, T. Heindorff, and R. Schröder, IEEE Trans. Instrum. Meas. **IM-34**, 136 (1985).

/15/ 3M Nextel Black Paint, Minnesota Mining and Manufacturing Company, commercial name Black Velvet Coating 2010.

/16/ K. Möstl, in *Absolute Radiometry*, (Academic Press, New York, 1989), p. 118.

/17/ A. Bauch, K. Dorenwendt, B. Fischer, T. Heindorff, E. K. Müller and R. Schröder, IEEE Trans. Instrum. Meas. **IM-36**, 613 (1987).

/18/ J. Vanier and C. Audoin, *The Quantum Physics of Atomic Frequency Standards*, (Adam Hilger, Bristol and Philadelphia, 1989).

/19/ A. De Marchi, O. Francescangeli, and G. P. Bava, IEEE Trans. Instrum. Meas. **42**, 448 (1993).

/20/ C. Thomas, P. Wolf, P. Tavella, *Time Scales*, BIPM, (1994).

THE HIGH C-FIELD CS BEAM FREQUENCY STANDARD: PRESENT STATUS AND FUTURE IMPROVEMENTS

G. A. Costanzo, E. Rubiola, S. Therisod,
M. Siccardi, F. Periale, G. D. Rovera†,
V. Barychev‡, A. De Marchi

Dipartimento di Elettronica, Politecnico di Torino
e-mail: demarchi@polito.it
†LPTF, 61 Av. de l'Observatoire, Paris ‡VNIIFTRI, Mendeleevo, Russia

Abstract

In the past years, preliminary results on the high C-field resonator at the Politecnico di Torino were presented in the EFTF Conference. At this moment the first part of the project is accomplished and the 10^{-14} accuracy goal seems not exceedingly far to be pursued. The largest component in the list of uncertainty contributions affecting the frequency standard arises from disuniformity and stability of the C-field, but measurements on the field dependent transition suggest that it should be feasible to reduce this contribution to a level in the low 10^{-14} range. The latest results on the high field Cs standard will be reported here, with particular emphasis on recent improvements of frequency stability, and projections about the future.[1]

1 Introduction

The main sources of uncertainty in thermal beam Cs standards are related to atomic velocity, to neighboring transitions, and to rf field non-idealities caused by the microwave cavity.

Several innovations have been proposed, during the last decade, for the accuracy improvement of these standards, with the aim to achieve a level of 10^{-14} or better. These include optical selection and detection of atoms, ring terminated Ramsey cavities, and full processing of the information content hidden in the RF response on the tube.

A different approach was chosen at the Politecnico di Torino, based on the high C-field concept as outlined in [1]. In this scheme, the C-field value is set at 82 mT, where the Breit-Rabi law shows a minimum for the $m_F = -1$ hyperfine component. The latter becomes the clock transition, and the closest neighboring lines are some 600 MHz away far from it. Rabi and Ramsey pulling become then unimportant, as well as Majorana transitions, even if the beam passing holes were very large or the spatial variations of the magnetic field were wild.

Therefore, the high value of the C-field leads naturally to a completely different design philosophy compared to other laboratories standards. In the design which was implemented the C-field is excited longitudinally along the beam by a solenoidal magnet [2], and the microwave transition is driven in a cylindrical multilambda Rabi cavity. The latter resonates at the frequency ν_{-1} of the clock transition in the $TE_{01,7}$ mode, and at the frequency ν_0 of the field dependent transition in the $TE_{01,11}$ mode. The $TE_{01,n}$ mode features vanishing fields at the beam passing holes, which enables opening up the holes to host a very wide beam, for S/N maximization. Moreover, it has a high Q factor and a very flat phase map in the transversal plane around the symmetry axis.

As a result, the main source of uncertainty in this Cs beam standard comes from the limited C-field uniformity. The latter should be of the order of 10^{-6} if the corresponding bias must be kept below 10^{-14}. Correction with 10^{-14} uncertainty of the bias effect produced by a less uniform field can also be envisioned as a route to accuracy, provided the stability of the field map be good enough.

In the following, a quick description is given of the main features of the standard in its present configuration. Modifications and improvements are underlined and justified. Recent results are also reported, with particular emphasis on the short term stability, which is not far from 10^{-12} at 1 second. The short term stability of the standard was measured against a high performance selected BVA quartz oscillator, which remains below $2 \cdot 10^{-13}$ up to about 1000 s. For longer term stability measurements the only available reference was an HP 5061B Cesium equipped with a standard tube. No data below 10^{-13} could therefore be obtained.

2 Description of the apparatus

In order to minimize power shift and velocity dependent effects, the frequency standard design was ini-

[1]This work was partially funded by the EC within the Human Capital and Mobility programme (HCM contract #ERBCHRXCT930117), and by CNR, ASI, and the Politecnico di Torino.

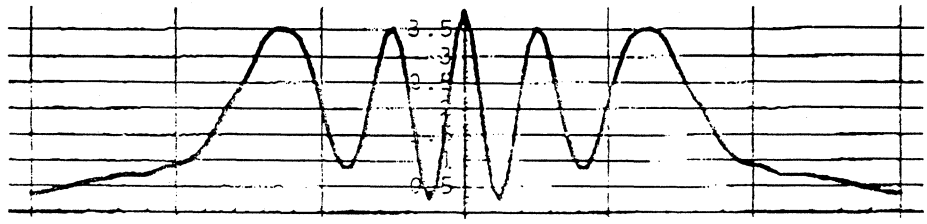


Figure 1: Typical lineshape of the microwave tube response with the $7\lambda/2$ cavity. The power was 6 dB below optimum. The FWHM of the central fringe was 350 Hz

tially started with the idea to adopt gradient magnets at least for state preparation, in order to exploit their velocity selection effect. Bipolar magnets with a wide useful area were chosen, in order to recover from the atomic beam flux reduction caused by the velocity selection. This was allowed by the possibility to safely widen up the beam offered by the combination of the high field with the used cylindrical cavity mode. A beam diameter of 8 mm was adopted. Unfortunately, the gradient magnet which was realized did not manage to grant a sufficient beam separation at the detector [3, 4], and was therefore at least provisionally substituted with a pumping laser beam.

In the present configuration the standard is an optically pumped device. The optical scheme includes a single 150 mW extended cavity laser, locked to the 4-5 component of the D_2 Cesium line with the saturated absorption technique. The laser output is directly used for detection, while is shifted for pumping to the 4-4 transition in a single pass through an AO modulator driven at 251 MHz. Both pumping and detection are performed with a standing wave, obtained by retroreflecting the laser beams with mirrors in both regions. The detection features a high efficiency light collector system [5] with two spherical mirrors and a light pipe.

A simple oven with a "zero lenght" 2 mm diameter diffusing aperture is used. In order to reduce the beam divergence and, consequently, stray atoms, iris diaphragms are placed not far from the source. This effusive source is probably suffering from excessive scattering at the aperture, which reduces the number of slow atoms in the beam, but has been very convenient in the preliminary research phase, during which the vacuum vessel was opened many times. A better source is available, kindly provided by Hewlett Packard Co., and will be used as soon as some residual problems are solved.

The C-field bobbin is wound directly on the Copper tube which constitutes the vacuum chamber, and a single mumetal magnetic shield is also the flux return of the magnet. The latter was designed to drive the mumetal to its maximum permeability working point. In this way the 2 mm thick mumetal sheet acquires the same shielding efficiency of a 10 mm thick mumetal shield at low field. All the experiments were carried out with a relative spatial uniformity, over the interaction zone, of 10^{-4}. This is the level of uniformity obtained in the magnet without corrections,

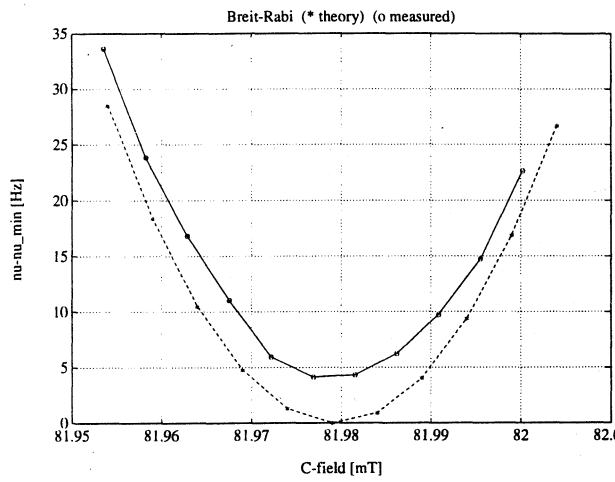


Figure 2: Comparison of measured ν_{-1} values with the calculated Zeeman effect in proximity of 82 mT

and produces a 4 Hz positive shift of the observed ν_{-1} transition. Corrections of the field uniformity during operation are possible, as illustrated below, and are expected to yield a bias effect in the low 10^{-13}, with uncertainty in the low 10^{-14}.

Fig. 1 shows the lineshape obtained at the detector with a microwave power 6dB below optimum. This shape pretty much agrees with theoretical calculations if it is assumed that scattering at the oven aperture penalizes the slow atom content in the velocity distribution [6].

The signal to noise in 1 s is of the order of ten thousand, most likely limited at this moment by stray light in the detection zone. This S/N ratio is still far from the shot noise limited value of roughly $5{\cdot}10^4$ which can be calculated from the theoretical atomic flux of 10^{11} atoms per second expected with the oven at 100 °C.

The servo system is based on slow square wave frequency modulation and a digital servo system, implemented in a personal computer in the same way as in other primary frequency standards [5, 7].

3 Zeeman effect

In Fig. 2 a number of frequency measurements of the $m_F = -1$ clock transition are reported, as taken

at different C-field values around 81.98 mT, where it is theoretically expected to be minimized according to the Breit Rabi law. The calculated frequency-vs-field curve is also shown, evidencing by comparison the expected 4 Hz positive frequency shift caused by the 10^{-4} field uniformity. No obvious deviation from the Breit-Rabi law were found with the present uncertainty of 0.1 Hz.

Higher accuracy than presently reached is needed in order to investigate anomalies of the Zeeman effect such as those caused by diamagnetism and the "hyperfine assisted shift", which are theoretically estimated to be of the order of 10^{-12} [8].

4 C-field stability

The stability of the C-field in the high field standard also needs to be as high as the uniformity, if not higher. In fact it is alway a good idea to design for a stability level better than the target accuracy. It would then be highly desirable to have a field stability in the low 10^{-7}s, which would guarantee a frequency stability limit, at least from this cause, around 10^{-15}.

Eventually, when the standard will be fully operational, it is anticipated that the magnetic field will be stabilized on the $m_F = 0$ field dependent transition with a digital electronic loop, which should be able to yield the desired level of stability in the long term.

However, that servo loop has not been realized as yet, and at any rate it may not be sufficient to guarantee the necessary short term stability of the field. A prestabilization of a different nature may therefore be necessary. For these reasons, and because the 10^{-7} field stability is already necessary in the prototype development stage (if only in order to evaluate the field uniformity), much effort was spent on this problem.

As for the stability of the exciting current, the 10^{-7} level is reached with active control, based on the voltage drop across a series manganin resistor, as illustrated in detail in [9].

Also necessary appeared to be to stabilize the temperature of the magnet. In fact the coercive field $H = NI/L$ depends not only on the exciting current I and the number of turns N (which remain constant by construction and stabilization respectively), but also on the length L of the magnet, which varies with temperature according to the expansion coefficient of Copper. As a consequence, a negative temperature coefficient of the field was expected, in the measure of $-1.6 \cdot 10^{-5}$/K. This would impose a temperature stability of 6 mK of the magnet.

An evaluation of the actual temperature coefficient was performed by correlating values of magnetic field and temperature measured during a 24 hour period, as is done in the scatterplot of Fig.3. The temperature was measured with a Platinum resistor placed directly on the microwave cavity, and the field was measured through the frequency of the ν_{-1} clock transition at a magnetic field offset by about 20μT on the low side of the minimum of Fig. 2. The coefficient appears to be about $-1 \cdot 10^{-4}$/K, which is much higher

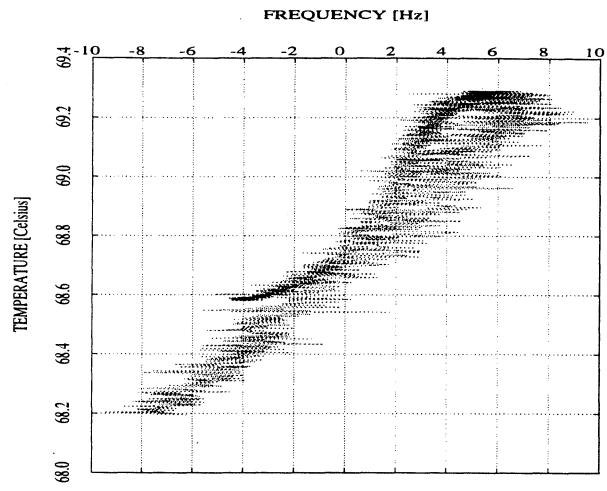


Figure 3: Scatterplot of correlated measured values of temperature and magnetic field in the C-field magnet

than expected. An explanation of this has not been found, but the result is that temperature stabilization to the mK is necessary for the desired magnetic field stability.

This specification is not easy to meet because of the magnet's dimensions and power dissipation (800 W). Nevertheless it was met by the temperature control described below. The magnet is 0.7 m long and 0.24 m in diameter, and weighs about 50 kg. Its average steady state temperature is about 70 °C when energized and water cooled by a flux of 6 liters per minute of water at 15 °C.

A stabilization approach by water flow control was choosen for practical reasons. The temperature inside the magnet, was carefullly studied with the equivalent conductance method [10]. The validity of the model was checked by comparison with available measurements (temperature of the cavity and average temperature of the winding). According to calculations the longitudinal variations of temperature in the winding do not exceed 6 K. Because the C-field will be made uniform by shim coils [11], spatial field variations caused by variable expansion can be tolerated, provided the temperature distribution be constant. Therefore, the thermal coupling between the magnet and the environment must be minimized in order to prevent temperature uniformity variations. This is achieved by water cooling the magnet end caps as long as the cylindrical surface of the winding.

The magnet temperature is evaluated by measuring the resistance of its copper winding. The latter in fact has a thermal coefficient of $3.9 \cdot 10^{-3}$/K.

This approach ensures a good accuracy because the resistance is significant of the average temperature integrated over the entire volume of the copper. Moreover, it is suitable for a tight control loop implementation because the lag time from the actuator (water flow) to the sensor is mainly due to the thermal capacitance of the magnet. Since the current is stabilized

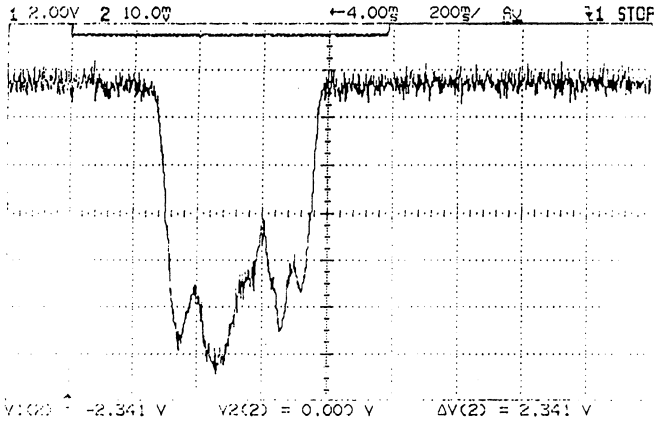


Figure 4: Observed lineshape of the ν_0 field dependent transition which is used to evaluate field uniformity

to some 10^{-7}, its fluctuation is negligible, and the voltage measured across the magnet gives the correct temperature information. The main drawback of the above described approach is the poor gain due to the thermal coefficient of copper. In fact, a 3.9μV measurement precision is required for 1 mK precision.

The actuator is a needle valve positioned right downstream the water intake and driven by a step motor controlled by a PC.

The loop includes a classical PID (proportional, integral, derivative) algorithm implemented in software. The operating parameters are set according to the Zigler-Nichols criterion, well described in [12]. The "ultimate time" of the loop (the oscillation period of the system when it's set at its critical gain) turned out to be 32 s. The integral and derivative time constants were then set to $\tau_I = 16$ s and $\tau_D = 4$ s respectively. The transient response is improved by adopting an anti-windup trap that stops the integration when the actuator is saturated. The noise behavior of the derivative process is significantly reduced by evaluating the derivative from a least square fitting of the most recent data. In this way the fatigue of the valve is significantly reduced. When the pressure of the water supply is changed from 3 bar to 2 bar, it takes the system less than 100 s to recover within a temperature error of a few mK; the peak error during the transient recovery is below 20 mK.

The measured Allan deviation of temperature flickers out below 1 mK, and starts random walking after 1000 s. However, the random walk of these measurements is caused by the voltmeter, whose expected long term stability is 5 mK at 3 months.

5 C-field uniformity

Three independent evaluations of the C-field uniformity along the 0.4 m of the interaction were made. All are consistent in indicating a relative uniformity of 10^{-4}.

The first one comes from direct measurement of the field distribution along the magnet, performed with an NMR probe [2, 3, 11]. The second one can be induced from frequency measurements of ν_{-1} versus C-field, which yield the correct 4 Hz positive shift, as shown in Fig.2.

The third one can be derived from the observation in Fig.4 of the width and shape of the field dependent line. In fact the field dependence of ν_0 is $7{\cdot}10^9$ Hz/T, which translates the observed FWHM of 250 kHz into a peak to peak field variation of about 30 μT, which is again a relative rms uniformity of roughly 10^{-4}.

The lineshape in Fig.4 can clearly be read as the collection of 11 separate Rabi lines. These correspond to the different C-field values experienced by the atoms in the 11 microwave field antinodes of the $TE_{01,11}$ cavity mode which is used to excite the field sensitive $m_F = 0$ transition.

Two different strategies can be devised with the aim of improving the field uniformity.

One calls for acquisition of the field profile before assembly and off-line optimization of the uniformity by the use of shim coils. A full layer of continuous winding, with taps at each turn, was placed for this purpose on the cooling jacket, external to the 24 layer solenoid. A simulated uniformity optimization, with current limitation at 10 A, yielded an rms value of $1.5{\cdot}10^{-5}$ with 18 shims. A higher number of shims would be needed for uniformity in the low 10^{-6} region. The problem with this approach is that the shims are wide (150 mm) and therefore their action range interfere, making it very difficult to correct for short wavelength variations.

The other is suggested by Fig.4, and involves winding 11 shim coils directly on the microwave cavity, inside the vacuum vessel, at the locations of the antinodes of the $TE_{01,11}$ mode. The advantage of this strategy is that the smaller shims have a shorter range of action and therefore do not interfere heavily, making it possible to observe directly in the shape of the $m_F = 0$ transition the effect obtained separately on the 11 resonances. On the other hand, the drawback is that there is no way to tell during operation how uniform the field is in between the corrected antinodes.

6 Frequency stability

A very high power sensitivity was measured by intentionally varying the microwave power level with an attenuator. Its evaluation may have to be repeated with better attention to electronics in a loop in which the signal level is varied, but the measured coefficient of the order of 10^{-10}/dB, is certainly worrysome in view of the 10^{-15} target stability. It is assumed here that the observed sensitivity derives from coupling of the multilambda cavity mode with the spatial field variations, and that the coefficient will turn out to be much smaller once the field uniformity is improved to the desired level. Notably, the lack of this kind of coupling may be an unforeseen advantage of the Ramsey cavity.

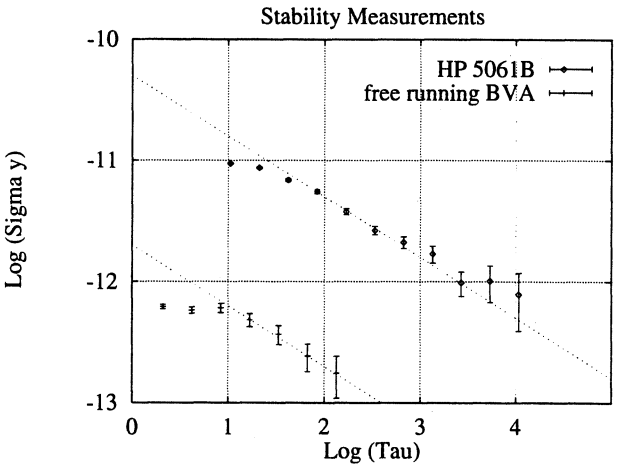


Figure 5: Allan deviation of the frequency comparison between the high field standard and a standard HP5061B tube or a selected BVA oscillator

In Fig.5 the best results obtained for frequency stability are reported. The two plots represent the observed stability of the high field standard measured against the BVA oscillator in the two configurations that have been used for short and long term stability measurements. In order to learn about the short term stability, the BVA was free running. For long term stability instead, it was phase locked to the 5061B standard. In both cases no deviation was observed from the $\tau^{-1/2}$ white frequency noise slope, down to the measurement's flicker floor.

This result proves that the temperature of the magnet was actually stabilized to the mK for the whole measurement duration. The latter was in some runs of the order of one day.

7 Conclusions

In this paper we reported the present development status of the high C-field Cesium frequency standard of the Politecnico di Torino. Its short term stability was measured to be $2{\cdot}10^{-12}/\sqrt{\tau}$ by comparison with a free running BVA oscillator. Three independent ways were followed to evaluate the uniformity of the C- field, and all yielded the level of 10^{-4}. An on-line strategy for the optimization of the uniformity while the standard is operational was devised, which is based on the observation of the shape of the field dependent line. Temperature stabilization of the C-field to the mK magnet was found necessary and realized with a digital control loop.

Acknowledgments

The authors wish to acknowledge useful discussions with many colleagues, in particular R.E. Drullinger and W. Itano of NIST, and D.J. Knight, retired from NPL, and the encouragement of L.S. Cutler, of HP, throughout the development of the standard.

References

[1] A. De Marchi, "The high C-field concept for an accurate Cesium beam resonator", *Proc. 7th EFTF*, 1993.

[2] G. Costanzo, A. De Marchi, M. Nervi, M. Repetto, "High homogeneity solenoidal magnet for Cs frequency standard", *IEEE Trans. on Mag.*, vol. 30 no. 4, july 1994, pp. 2628-2631.

[3] G. Costanzo, E. Rubiola, A. De Marchi, "First results on the high C-field Cesium resonator at the Politecnico di Torino", *Proc. 9th EFTF*, 1995.

[4] G. Costanzo, A. De Marchi, M. Repetto, "State selection bipolar gradient magnet for a high C-field Cesium beam resonator", *Proc. 10th EFTF*, 1996.

[5] R. E. Drullinger, D. J. Glaze, J. Lowe, J. Shirley, "The NIST optically pumped Cesium frequency standard", *IEEE Trans. on Instr. and Meas.*, vol. 40 no. 2, april 1991.

[6] A. De Marchi, P. Tavella, E. Bava, "Multilambda Rabi cavities in atomic beam frequency standards", submitted for publication to Journal of Applied Physics, 1997.

[7] G. D. Rovera, E. de Clercq, and A. Clairon, "An analysis of major frequency shifts in the LPTF optically pumped primary frequency standard," *IEEE Trans. on UFFC*, vol. 41, pp. 457–461, May 1994.

[8] W. Itano, "High magnetic-field correction to Cesium hyperfine structure", *Proc. of the 5th Symposium on Frequency Standard and Metrology*, Woods Hole (1995).

[9] E. Rubiola, "High stability current control in the 10 A range," *IEEE Trans. on I&M* vol. IM-45 n. 5, pp. 865–871, ottobre 1996.

[10] L. Giordana, "Campo termico in un campione di frequenza primario", Tesi di Laurea, Politecnico di Torino, Dec. 1996.

[11] G.A. Costanzo, "Ottimizzazione di un campione atomico di frequenza a fascio di Cesio a selezione magnetica", PhD thesis, Politecnico di Torino, 1996.

[12] G.F.Franklin, J.D.Powell, A.Emami-Naeini, *Feedback control of dynamic systems*, Addison Wesly, USA 1991.

FREQUENCY EVALUATION OF A MINIATURE OPTICALLY PUMPED CESIUM BEAM CLOCK

B. Boussert, P. Cérez, L. Chassagne and G. Théobald

Laboratoire de l'Horloge Atomique
Unité Propre de Recherche CNRS associée à l'Université Paris-Sud
Bâtiment 221 - Université Paris-Sud, 91405 Orsay Cedex, France

ABSTRACT

By taking into account severe requirement on the microwave leakage level of the short cavity implemented in our Cs IV resonator, we have designed a new cavity and run another frequency evaluation of Cs IV. In this experiment, our machine is driven by a standard electronics coming from a HP 5061 B clock. Its main features are :
- a sinusoïdal modulation of the interrogation microwave signal,
- a microwave power stability of about one percent at a temperature of 20 ± 1 °C.

The following results have been obtained :
σ_y(1 day) = 2×10^{-14} : this value holds up to 3 days; accuracy : $\cdot10^{-12}$; repeatability : 3×10^{-13}.

In our opinion, the flicker floor is beginning at 2×10^{-14} and is mainly due to both the free running fluctuations of the microwave power applied to the resonator and the fluctuations of the external static magnetic field. Further long term frequency stability improvements are expected by the use of a most sophisticated electronics. This latter is based on the use of a square-wave frequency modulation and active controls of both the microwave power and the DC magnetic field. Regarding the accuracy, we are mainly limited by the lack of knowledge of the end to end cavity phase shift.

Keywords : Atomic cesium clock; Frequency evaluation

1. INTRODUCTION

In EFTF 94, we reported the first results concerning the frequency evaluation of a miniature optically pumped cesium beam clock called Cs IV. We obtained a good short and medium term frequency stability but the delivered frequency did not match with the definition and suffered from a high sensitivity to the microwave power. As the origin of this drawback was found to be connected with microwave leakage of the Ramsey cavity, we designed a new cavity working properly. We then run another frequency evaluation of Cs IV. In this experiment, this machine was driven by a standard electronics coming from a HP 5061 B clock. After a brief outline of the design and realization of the new cavity, we present the results of comparisons of Cs IV with the Laboratoire de l'Horloge Atomique hydrogen maser and with the new high performance HP 5071 A frequency standard. This allows us to specify the one day frequency stability, the repeatability and the accuracy of the device.
Finally, we shall present ways to improve these first results.

2. CAVITY DESIGN

Presently, it is widely recognized (Refs. 1-2) that magnetic microwave stray fields propagating outside the two interaction regions on the path of the atomic beam generally disturb the atomic states and lead to significant frequency shifts. We demonstrated (Ref. 3) that the main sources of microwave leakages come from the cavity itself. As a matter of fact, there are many places where leakages originate :
- the junctions between the short circuit end caps and the U shaped main waveguide;
- the coaxial feedthrough which supplies the cavity;
- the ensembles hole-waveguide below cut-off ensuring the passage of cesium beam on the one hand and the attenuation of radiations issued from the holes on the other hand.

Consequently, key elements in the design of a good cavity are the following :
a) The use of amagnetic and highly conductive material.
b) An excellent electrical contact between all the constitutive elements of the cavity associated to good surface characteristics.
c) The geometrical features of the ensemble hole-waveguide below cut-off. At the output of the cut-off waveguide, the residual radiated microwave power must be at an acceptable level. This last was empirically fixed by comparison with a so called "well designed cavity". This level was taken equal to –115 dbm when the input excitation is +10 dbm.

The solutions which have been carried are :
- the cavity is made with a copper alloy (Cu-Cr-Zr) material which can be easily machined.
- the U bent cavity body is an assembly of two half shelves carefully welded with an electron gun.

The cavity length is then adjusted in order to get a good electrical symmetry and tuned to the proper resonance frequency. After this setting, the two end caps are welded on the main body by transparency with the same technique as above.
The microwave power radiated outside the cut-off waveguide depends on the size of the hole and the geometrical feature of the cut-off waveguide (section and length) : as the overall dimension of the cesium tube is prescribed, the length of the cut-off waveguide should be L ≤ 15 mm. Taking L = 15 mm, we have empirically found that a –115 dbm microwave level leakage is reached when the hole section and the cut-off waveguide section are respectively 3×5 mm^2 and 4×7 mm^2. Such cut-off waveguides have been finally fastened on the cavity by mean of YAG laser welding techniques. Fig. (1) shows the achieved cavity which has been implemented in the Cs IV resonator.

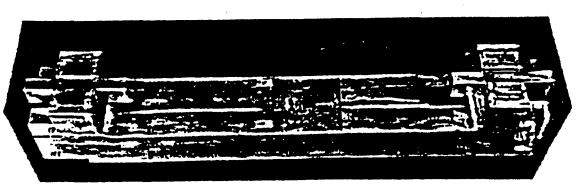


Figure 1 A view of the Ramsey cavity

3. OPERATING CONDITIONS

The Cs IV clock has been in operation since February 1996. The oven temperature was set at 90 °C. The static magnetic field B_c was set at 5.896×10^{-6} T. A measurement of its stability and accuracy was performed until September 96. During this run, the machine has been driven by a standard electronics coming from a HP 5061 B clock. Some modifications were carried on this equipment namely :

- a new AC preamplifier at 137 Hz in order to match the gain before the synchronous detector and avoid saturation.
- a microwave filter at the output of the SRD multiplier in order to improve the spectral purity of the interrogation signal.

The relevant features of the interrogation signal are :

- a sinusoïdal modulation of its frequency at 137 Hz,
- a natural power stability of about 1 % at temperature of 20 ±1 °C.

In the laboratory, two frequency references are available : the first one (H maser) is mainly used for short and medium term frequency stability measurement. The second one (HP 5071 A high performance cesium clock) is essentially devoted to accuracy and long term frequency stability measurements.

4. RESULTS

4.1. Frequency sensitivity to the microwave power

At optimum power (P_{opt} = –12.6 dbm), we measured a sensitivity of -1.5×10^{-13}/db. Compared to the first experimental run with the leaky cavity, an improvement of the frequency sensitivity to the microwave power by a factor 230 has been obtained.

4.2. Frequency stability

The clock was continuously operated for a period of 15 days. Fig. (2) shows the classical law $\sigma_y(\tau) = 5\times10^{-12}\times\tau^{-1/2}$ for τ ranging between 1 s and 5.8×10^4 s. Compared to the previous experiment, an improvement of $\sigma_y(\tau$ = 1 day) by a factor 2 is also observed. Furthermore the beginning of a flicker floor at the 2×10^{-14} level is put into evidence for 1 day < τ < 3 days. With this new cavity, the long term frequency drift which was observed in the first evaluation is now well controlled. In our opinion this flicker floor limitation at 2×10^{-14} level results from the residual static magnetic field fluctuations and from microwave power drifts. Indeed a third magnetic shielding initially designed for this resonator was not implemented during this run. Concerning the microwave power, it is actually temperature dependent. The microwave stability is not better than 1 % during 15 days of data. This induces microwave fluctuations and consequently frequency fluctuations.

Finally, let us mention that the same $\sigma_y(\tau)$ curve has been found again after another measurement run which followed a complete break of the machine due to a failure in the power plant.

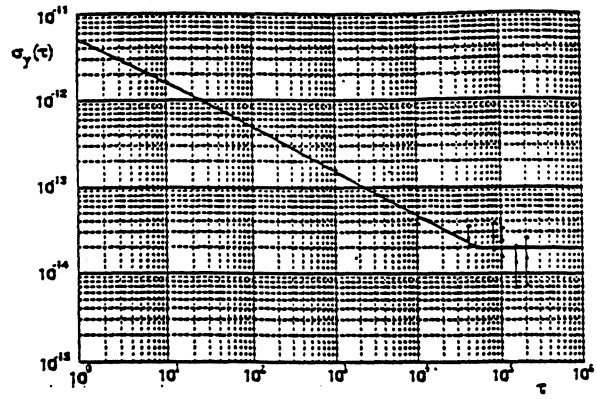


Figure 2 Measured short and medium term frequency stability $\sigma_y(\tau)$ of the resonator Cs IV

4.3. Frequency shift evaluation

4.3.1. Static magnetic field shift

The most important frequency shift results from the quadratic Zeeman effect. It amounts 1616×10^{-13} as the value of the static magnetic field is $B_c = 5.896\times10^{-6}$ T. The errors caused by field measurements and field homogeneity give an uncertainty of 0.7×10^{-13} in our case.

The evaluation of almost the other frequency shifts is based on a numerical calculation of formulas established in the laboratory (Ref. 4) which apply to short tubes excited by a microwave interrogation in phase opposition in each cavity arm and with sine-wave frequency modulation. The relations of interest can be found in Ref. 3 for :

- the quadratic Doppler shift,
- the shifts related to the microwave cavity (cavity pulling and end-to-end phase shift),
- the shifts related to electronics imperfections.

Let us recall that for monokinetic atoms, the above mentioned frequency shifts depend on the interaction time τ (in one arm of the cavity), on the Rabi angular frequency b and on the amplitude ω_m of the modulation. In Cs IV resonator operating at 90 °C, a complete optical pumping leads to a Maxwellian velocity distribution in the beam. Consequently, frequency shifts have to be weighted by the interaction time distribution :

$$f(\tau) = \frac{2}{\tau_0}\left(\frac{\tau_0}{\tau}\right)^5 \exp(-\tau_0^2/\tau^2). \qquad (1)$$

$\tau_0 = 1/\alpha$ is a particular value of τ where α is the most probable velocity in the cesium oven.

In table 1 are gathered the typical values of the parameters needed for our calculation.

oven temperature	t = 90 °C	$\alpha = 213$ m s^{-1}
interaction length in one arm of the cavity	l = 1 cm	$\tau_0 = 46.9$ µs
separation between the two arms	L = 16.2 cm	$T_0 = \frac{L}{l}\tau_0$ = 760 µs
Optimal Rabi pulsation(1)	$b_{opt} = 26\ 500$ s^{-1}	$b_{opt}\tau_0 = 1.24$
Modulation amplitude(2)	$\omega_m = 2\pi\times190$ Hz	$\omega_m T_0 = 0.9$

Table 1

4.3.2. Quadratic Doppler shift

Around the microwave power P_{opt} = −12.6 dbm corresponding to $b_{opt} = 26\ 500$ s^{-1}, the quadratic Doppler shift is -2.5×10^{-13}. If we assume that the interaction time distribution (1) matches perfectly the actual one, the frequency shift uncertainty results from the temperature and microwave power measurement uncertainties. Those ones are $\Delta T = \pm0.5$ K and $\Delta P/P = \pm5/100$ respectively. We finally get a $\pm0.1\times10^{-13}$ uncertainty.

4.3.3. Shifts related to the microwave cavity

- Cavity pulling

The cavity detuning is +3,4 MHz. The loaded quality factor Q is 900 while the atomic Q line is 1.4×10^7. We obtain a frequency shift equal to -1.5×10^{-13}. Its uncertainty, mainly due to the microwave power measurement uncertainty, amounts $\pm0.5\times10^{-13}$.

- End to end phase difference δφ

From experimental tests on the cavity, we notice that the electrical length difference between the two cavity arms is at worst 0.1 mm leading thus to $\delta\varphi = 40$ µrad. This value was used for the computation of the related frequency shift. We obtain -10×10^{-13}.

It appears that $\delta\varphi$ is measured with an important uncertainty and, in order to be cautious, we take $\Delta(\delta\varphi)/\delta\varphi = 1$. It results that the uncertainty due to the end to end phase difference is $\pm10\times10^{-13}$. It is the biggest cause of uncertainty in the total accuracy budget.

4.3.4. Shifts due to electronic imperfections

- Influence of spurious spectral components in the microwave interrogation field

As previously mentioned, the microwave spectrum issued from the HP 5061 B electronic set-up is improved by means of an additive microwave filter. The main spurious line are positioned at ±2.5 MHz, ±5 MHz and ±12.6 MHz from the carrier and their level is −46 dbc; −36 dbc; −40 dbc respectively. The corresponding frequency shifts can be predicted by means of relation 6-94b in (Ref. 5). The most significant frequency shift comes from the 5 MHz line. It is estimated to 3.5×10^{-14} with a negligible uncertainty.

- Influence of even harmonics of the modulation frequency in the microwave interrogation field.

Due to non linearities in the generation of the sine wave frequency modulation, the instantaneous angular frequency of the microwave field contains also a component at frequency $2\omega_M = 2\pi\times274$ Hz. The distortion ratio δ_2 is assumed to be much smaller than 1.
The fractional frequency offset due to this effect is proportional to δ_2 and depends on ω_m. For $\omega_m = 2\pi\times190$ Hz and $\delta_2 = 4\times10^{-6}$, we get a frequency offset equal to -1.5×10^{-13} at our operating optimal power (P_{opt}). It is difficult to evaluate δ_2 which unfortunately increases as ω_m does.

- Influence of an amplitude modulation of the microwave interrogation field

The frequency distortion is usually accompanied by an amplitude distortion. Moreover, the latter always occurs when the cavity is detuned. Consequently an amplitude distortion is present in our case. We consider only the term at $\omega_M = 2\pi\times137$ Hz, with a small amplitude distortion ratio called ζ_1. The frequency offset has been calculated as a function of the microwave power for three modulation amplitudes ($\omega_m = 2\pi\times190$; $2\pi\times260$;$2\pi\times330$) and amplitude distortion ratio ζ_1 respectively equal to 4×10^{-6}; 7.6×10^{-6}; 1.2×10^{-5}. Around the optimal power, the calculated frequency shift is negative and remains $\leq 1\times10^{-13}$.
Finally, it appears that reasonable amplitude and frequency distortions ratios lead to frequency shifts in the range 10^{-13} which are largely dependent on the value of ω_m. This fact is demonstrated by measurement of the overall frequency sensitivity as a function of the microwave power for three values of ω_m Fig. (3). We see that this sensitivity can change its slopes s and we can find a particular value of ω_m which pratically gives $s \simeq 0$. This behaviour underlines the well known limitation of this analogic electronics.

(1) This value optimizes the slope of the error signal for the used modulation amplitude $\omega_m = 2\pi\times190$ Hz.

(2) This value is less than the optimal modulation amplitude which is close to $2\pi\times W/2$, where W = 660 Hz is the linewidth of the clock signal.

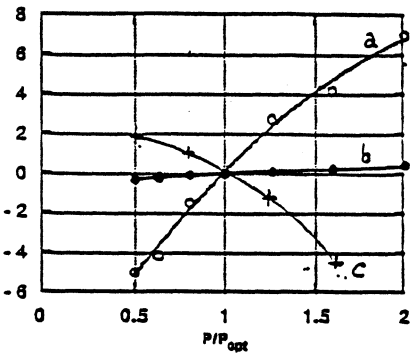


Fig. 3 Sensitivity of the frequency to microwave power (parts in 10^{-13}) : a - $\omega_m = 2\pi\times330$ rad/s; b - $\omega_m = 2\pi\times260$ rad/s; c - $\omega_m = 2\pi\times190$ rad/s; $P_{opt} = -12.6$ dbm

4.3.5. Light shift

A theoretical evaluation of the light shift induced by light scattered by the optics shows that it is less than 10^{-14} if the scattered intensity is about a few $pW.cm^{-2}$. Let us note that in our experiment, we took a lot of precautions to reduce the scattered light in the cavity. As it is impossible to measure it, we run an experiment in which the pumping laser intensity was reduced by a factor 2 while keeping the right operating conditions concerning the clock S/N ratio. No frequency shift greater than 2×10^{-14} was detected. Consequently, we took this value as the uncertainty relative to the light shifts.

4.3.6. Rabi and Ramsey frequency shifts

The microwave spectrum delivered by Cs IV is highly symmetrical. The measured asymmetry is less than 10^{-2}. Consequently, the Rabi pulling frequency shift is less than 5×10^{-14}.

Concerning the Ramsey pulling, it originates in microwave π transitions. In Cs IV, their amplitude is weak because beam holes are small. Their effect on the clock frequency is completely negligible.

4.3.7. Total accuracy evaluation

The different considered frequency offsets are summarized in table 2 with their uncertainty. The final accuracy σ is given by the quadratic sum of the uncertainties. We have $\sigma = 1\times10^{-12}$. This value compares favourably with the accuracy of high performance commercial clocks.

Frequency shifts	Fractional frequency shift $\times 10^{-13}$	Fractional uncertainty $\times 10^{-13}$
C field	1616	± 0.7
Second order Doppler	– 2.5	± 0.1
End to end cavity phase shift	– 10	± 10
Cavity pulling	– 1.5	± 0.5
Spurious components in the microwave	0.35	± 0.035
Harmonic distortions	≤ 5	—
Light shift	0.2	± 0.2
Rabi and Ramsey pulling	0.5	± 0.2
Total uncertainty		±10

Table 2 : Systematic major frequency shifts in Cs IV and uncertainties in parts $\times 10^{-13}$ at the operating optimum power.

4.4. Repeatability

A change in the C-field value followed by a return to the starting value of operation of B_C, all the other parameters being unchanged, did not lead to a frequency shift to within about 2×10^{-13}. We recall that any change in B_C values requires a careful demagnetization of the shielding. In another test, the clock was completely stopped and then started up again. A relative frequency difference equal to 3×10^{-13} was measured.

As a conclusion, we evaluate the Cs IV repeatability to 3×10^{-13}.

5. CONCLUSION

It appears that the cavity change eliminates the drawback related to microwave leakages. The overall frequency performances show that our miniature optically pumped cesium clock compares favourably with the best high performance commercial clocks.

A new digital electronics is presently implemented in which :

1) a square wave frequency modulation is used.
2) the quartz oscillator is locked by means of a main digital loop.
3) an active control of both the microwave power and the static magnetic field is achieved.

Such modifications should improve the long term frequency stability. Regarding the accuracy, we are obviously limited by the lack of knowledge of the end to end cavity phase shifts. Efforts are undertaken to measure it with a method different from the classical beam reversal technique.

ACKNOWLEDGEMENTS

The authors wish to thank C. Audoin for helpful contribution and discussions.
They also thank P. Greiner, A. Richard and J. Schilknecht du Service des Techniques Avancées CEA Saclay for efficient technical assistance in the realization of the cavity.
They gratefully acknowledge the support of STEI (Service Technique Electronique et Informatique de la DGA) and the Pole de Transfert de Technologie Temps-Fréquence.

REFERENCES

(1) Bauch A., Heindorff T., Schröder R. and Fischer B., *Metrologia, 33*, 249-249, 1996.

(2) de Clercq E., Makdissi A., in : *Proc. 5th Symposium on Frequency Standards and Metrology*, 1995. Edited by J.C. Bergquist, Singapore/New Jersey/London/Hong Kong, World Scientific, 409-410, 1996.

(3) Boussert B., *Evaluation des Performances en fréquence d'une horloge atomique miniature pompée optiquement*, Ph. D. Dissertation, Orsay (France), 1996.

(4) Audoin C., Private communication.

(5) Vanier J., Audoin C., *The quantum physics of atomic frequency standards*, Bristol and Philadelphia, Adam Hilger, 1989.

FIRST HIGH PERFORMANCE INDUSTRIAL PROTOTYPE OF OPTICALLY PUMPED CESIUM BEAM FREQUENCY STANDARD, FIRST MEASUREMENTS ON A COMPLETE CLOCK

Michel L. BALDY

TEKELEC Temex
29, Avenue de la Baltique
91953 LES ULIS Cedex FRANCE

ABSTRACT

TEKELEC has developped an industrial prototype of Optically Pumped Cesium-Beam Frequency Standard with the scientific support of the Laboratoire de l'Horloge Atomique (LHA) and the financial support of the French Military Administration (DGA). This work is based on years of experience of the LHA on cesium atomic clocks, especially on short atomic clocks and on the experience of Tekelec in time and frequency for systems (timekeeping, time distribution, synchronisation, etc.) and components (quartz oscillators, rubidium atomic clocks, etc.). The prototype developped by Tekelec is a complete clock (tube and electronic system) designed to be integrated in a 19", 3U packaging.

Keywords: Cesium atomic clock; Industrial prototype

1. INTRODUCTION

The huge growth of synchronization and positioning requirements for military, civilian or space applications will rapidly increase the need for ever more accurate and cheaper frequency standards. We believe that the optically pumped cesium-beam frequency standard is a good candidate for these applications.

The optical pumping technology has improved widely the performances of laboratory clocks and is used in the best National Primary Atomic Clocks (LPTF (Ref. 1), NIST (Ref. 2)) replacing the magnetic deflection technology.

The Laboratoire de l'Horloge Atomique (LHA, Refs. 3,4) has demonstrated the ability of this technology to improve short tube atomic clock performance and therefore optical pumping is of great interest for industrial applications.

Furthermore, the greater simplicity of this technology will lead to lower price.

Some companies have published works on short tube atomic clocks (Refs. 5,6,7,8,9) and announced the ability of this technology to obtain high performance characteristics but, as far as we know, nobody has published high performance results on their prototypes. The best results ever published on short tube atomic clocks have been obtained by the Laboratoire de l'Horloge Atomique (LHA, Ref. 10.) The tube of our prototype is an industrial evolution of the LHA's tube.

2. DESCRIPTION OF THE CLOCK

The schema of figure 1 represents the general operation diagram used in an atomic clock to stabilize the 10 MHz signal generated by a quartz crystal oscillator (OCXO) with the atomic resonator using a synchronous detection.

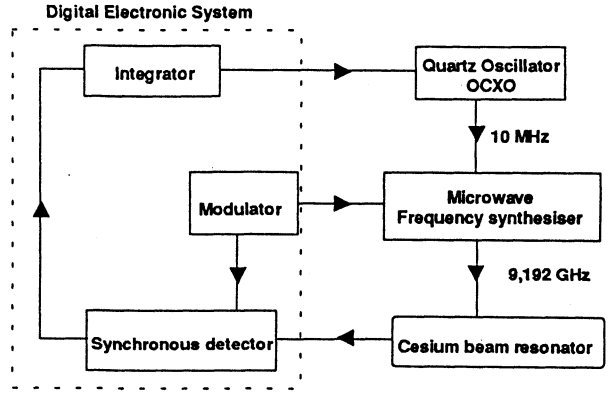


Figure 1 : Block diagram of the operation of an atomic clock

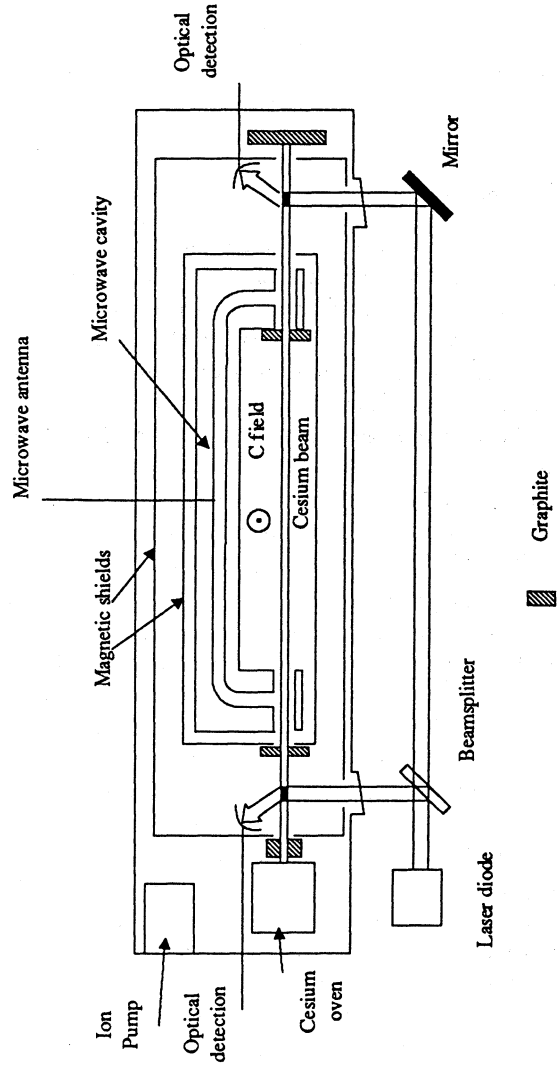


Figure 2 : Optically pumped cesium beam resonator

The main elements of the clock are the cesium-beam resonator, the microwave synthesizer and the digital electronic system.

All the elements of our complete clock have been entirely designed by Tekelec for the future industrial applications.

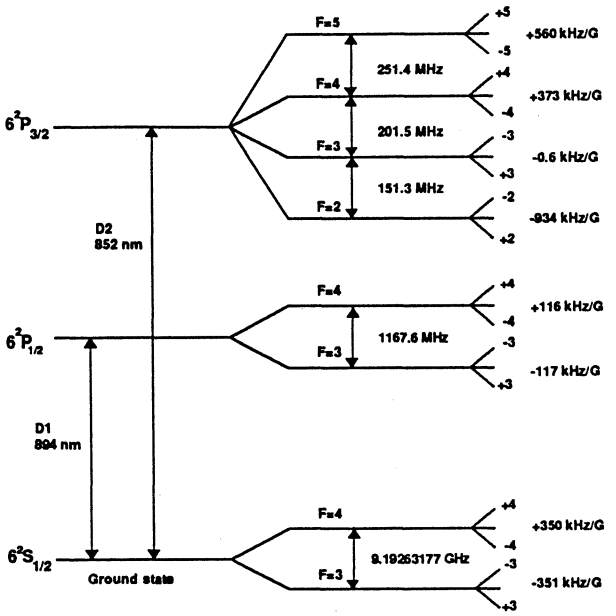


Figure 3 : Energy levels of the cesium 133 atom

The atomic resonator is optically pumped, the figure 2 represents the principle of operation of this tube. In high vacuum conditions maintained by an ion pump, a cesium oven generates a low divergence thermal cesium beam. This beam is shaped by several graphite diaphragms. The atomic preparation and signal detection are obtained by optical pumping using the same laser diode. In each of these zones, a static magnetic field of 40 μT is applied to avoid the Hanle effect. The fluorescence signal in the preparation area is used to lock the laser diode frequency to the 3-3 σ transition of the D2 line ($6^2S_{1/2} \rightarrow 6^2P_{3/2}$) of the cesium frequency diagram (figure 3). A Ramsey microwave cavity excited at 9,19263177... GHz by a coaxial antenna is used for RF interaction. The cavity is magnetically shielded and a low (6 μT) and very uniform static C-field is applied perpendicularly to the atomic beam axis. In the detection area, the fluorescence signal is used as the clock signal. A second magnetic shield protects the optical areas and the cavity. Two prototypes of the tube have been developped by Tekelec. The first one can be dismantled with ultra vacuum flanges and is connected to an external ion pump. The laser diode and the related opticals elements (collimating lens, beamsplitter and mirror) are fixed on an optical table. This prototype is under characterization. The second one, fully sealed, has been designed to be integrated in a 19" 3U packaging with the electronic system. All the elements are identical to the first prototype except the micro ion pump. This prototype is being manufactured and will be tested in a couple of months.

The electronic system developped by Tekelec and used in the prototypes is controlled by microprocessor. The starting sequence is fully automatic and is under the microprocessor control. The microprocessor starts the temperature regulation procedure of the cesium oven and of the microwave cavity. Then, it sweeps the frequency of the laser diode in order to find the six allowed transitions of the D2 lines (3-4, 3-3, 3-2, 4-5, 4-4, 4-3) before locking it to the 3-3 transition. The quartz oscillator frequency control voltage is scanned to find the central part of the microwave spectrum and is locked on the central fringe (clock signal). The C field is locked by checking the neighbouring Zeeman line and the microwave power is also under microprocessor control. The locking loops of most of the electronic elements (laser diode, quartz oscillator, C field, microwave power) are based on the principle of synchronous detection. For instance, the schema of figure 4 shows the diagram of the locking loop of the laser diode. It is a dual-locking loop acting both on the current and the temperature of the laser diode. The digital electronic system (DES) sets the current and temperature setpoints to the loops which control the driving current of the laser diode (LD) and of its Peltier element. A modulation is applied to the driving current of the laser diode. The fluorescence signal of the pumping area, detected by the photodiode (PD), is then demodulated and used to readjust the current and the temperature setpoints. The current loop is used to compensate for short term fluctuations and the temperature loop for long term fluctuations.

Tekelec has developped for its prototype a compact microwave synthesizer designed to obtain a very high spectral purity signal and to be integrated in the 19" 3U package.

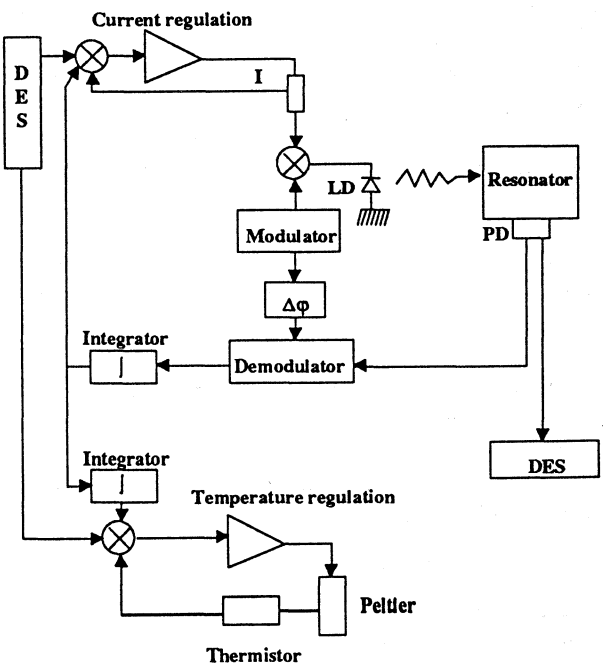


Figure 4 : Diagram of the laser diode locking loop

The complete digital electronic system and the microwave synthesizer have been realized and are used with the two prototypes.

3. MAIN ADVANTAGES

The main advantages of the optical pumping compared to using a magnetic state selection are:

- the atomic preparation is complete, increasing the signal-to-noise ratio,
- all the atoms participate in the clock signal (twice as much as with magnetic deflection), allowing a lower cesium consumption,
- there is no strong magnetic field near the microwave cavity,
- the Rabi-Ramsey spectrum is very symmetrical, increasing the accuracy and allowing a lower C-field, which decreases the magnetic sensitivity of the clock,
- the atom velocity distribution is well known, allowing an increase of the accuracy,
- there is no hot wire detector, which is very fragile, and limits the lifetime of the resonator,
- the technology is much simpler (without deflection), which will lead to lower cost.

Furthermore, we use an electrical coupling and a cavity geometry giving an odd number of longitudinal modes, and then a phase difference of 180° between the two oscillatory fields that the atoms experience in succession. The LHA (Ref. 11) has demonstrated that this configuration improves the frequency stability.

4. MAIN CHARACTERISTICS OF THE CLOCK

The aim of TEKELEC is to produce a state-of-the-art cesium-beam frequency standard for an attractive price.

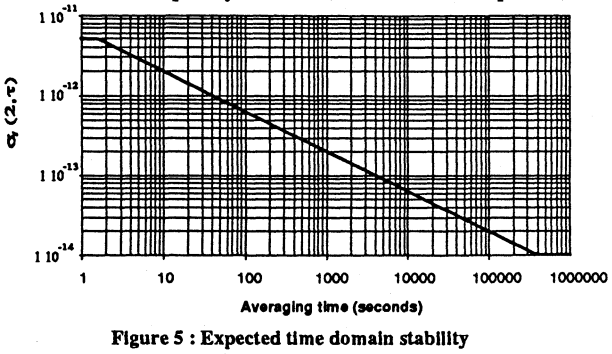


Figure 5 : Expected time domain stability (Allan deviation)

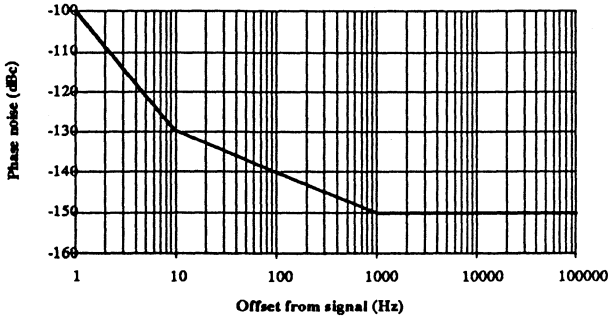


Figure 6 : Expected frequency domain stability

The accuracy will be better than $1 \cdot 10^{-12}$. The expected stability in the time domain follows a curve in $\sigma_y(\tau) = 6.3 \cdot 10^{-12} \tau^{-1/2}$ between 2 and 400 000 seconds, with a flicker floor at $1 \cdot 10^{-14}$ in laboratory conditions (figure 5). The planned stability in the frequency domain is given in the figure 6. As far as we know, these values represent the best characteristics for a commercial standard.

5. FIRST RESULTS

The first prototype has been realized and is under characterization. The full digital electronic system has been integrated and tested.

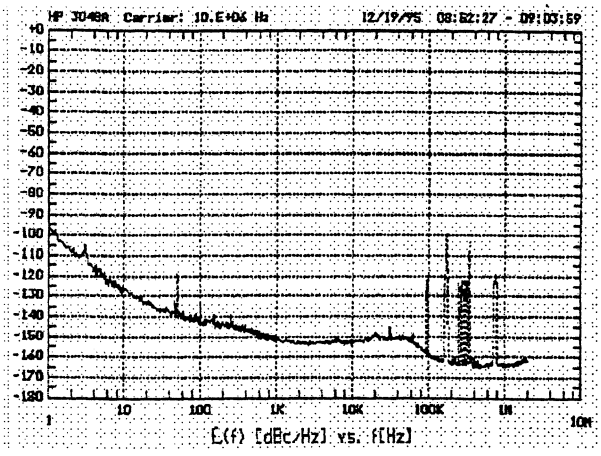


Figure 7 : Phase noise of the complete microwave synthesizer with its quartz oscillator compared to an identical quartz oscillator

The figure 7 gives the phase noise at 10 MHz of the complete microwave synthesizer with its quartz oscillator measured with a second and identical quartz oscillator. The characteristics of the quartz oscillator used are: -100 dBc/Hz at 1 Hz, -130 dBc/Hz at 10 Hz, -145 dBc/Hz at 100 Hz and -155 dBc/Hz at 1 kHz. So, the phase noise of the synthesizer is limited by the quartz oscillator under 1 kHz and the spectral purity fulfills the expected characteristics.

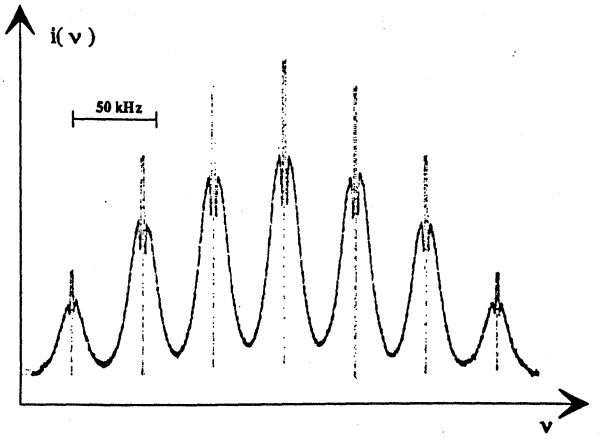


Figure 8: Rabi-Ramsey fringes

The Rabi-Ramsey resonance fringes have been obtained and are represented in figure 8. The difference between two

fringes is 42 kHz, corresponding to a C-field of 6 μT. These fringes present a minimum at the center (black fringe) which is characteristic of an electric coupling of the cavity (antenna). The spectrum is very symmetrical, as expected. The central line, which is used as clock signal, is given in figure 9. The linewidth (FWHM) of the central line, which is 600 Hz, agrees with the theoretical value for this cavity length and the atoms velocity with an oven temperature of 100°C.

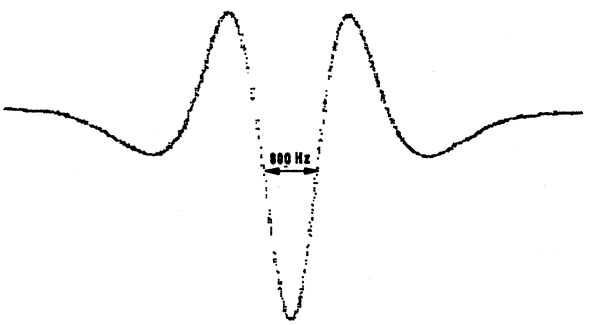


Figure 9: Central line of the Rabi-Ramsey fringes

We were not able yet to characterize it with an hydrogen maser. In order to have an idea of the characteristics of this prototype, we have made some measurements, between 1 and 500 seconds, by comparison with a HP 5071 option 001. The signal-to-noise ratio was not optimised and was of 6 000 in a 1 Hz bandwidth. In these conditions we have obtained $\sigma_y(\tau) = 1.1 \cdot 10^{-11} \tau^{-1/2}$. This value is pessimistic because, in that range, the effects of the lack of stability of the HP 5071A option 001 ($\sigma_y(\tau) = 8.6 \cdot 10^{-12} \tau^{-1/2}$) become important. So we believe that, after optimising the signal-to-noise ratio (>10 000), we will fulfill, and even improve, the expected performance of the clock.

6. FUTURE DEVELOPMENTS

A new prototype of tube is being manufactured. It integrates all the elements of the previous one in a completely sealed case with a micro ion pump. This prototype will be mounted in a conventional 19", 3U package and will be tested in industrial and laboratory conditions with an hydrogen maser. An industrialization is running and will lead to a product available by 1998.

7. CONCLUSION

TEKELEC has developped a prototype of optically pumped cesium beam frequency standard with the scientific support of the LHA. The first results show that the goal of the Tekelec clock, to be the state-of-the-art of commercial frequency standards, is reachable. This belief is confirmed by the very good results obtained by the LHA in an equivalent lab prototype. The industrialization of the product will be done in two versions: one conventional 19", 3U size and one compact version for military applications.

8. ACKNOWLEDGMENTS

The author wishes to thank the french military administration (DGA) for its financial support and especially Mr Klein. The Laboratoire de l'Horloge Atomique (LHA) had a great part in the success of this work thanks to C. Audoin, R. Barillet, P. Cerez and P. Petit for their scientific support. This work is a Tekelec team effort and L. Fialaix, D. Gin, F. Locre, F. Racine and C. Sallot have earned many thanks for their determining participation in this project.

REFERENCES

(1) E. de Clerck, G.D. Rovera and A. Clairon 1994, "*Progress on the LPTF optically pumped primary frequency standard*", Proceedings of the 8th European Frequency and Time Forum (EFTF), March 1994, pp. 497-502.

(2) R.E. Drullinger, J.P. Lowe, D.J. Glaze and J. Shirley 1993, "*NIST-7, the new US primary frequency standard*", Proceedings of the 7th European Frequency and Time Forum (EFTF), March 1993, pp. 549-551.

(3) P. Petit, V. Giordano, N. Dimarcq, P. Cérez, C. Audoin and G. Théobald 1992, "*Miniature optically pumped cesium beam resonator*", Proceedings of the 6th European Frequency and Time Forum (EFTF), March 1992, pp. 83-86.

(4) P. Petit, V. Giordano, P. Cérez, B. Boussert, C. Audoin and N. Dimarcq, "*Performance of a 2 dm3 Optically Pumped Cesium Beam tube: a progress report*", Proceedings of the 8th European Frequency and Time Forum (EFTF), March 1994, pp. 517-522.

(5) T. McClelland, I. Pascaru, J. Zacharski, N.H. Tran and M. Meirs (Frequency Electronics Inc.), "*An optically pumped cesium beam frequency standard for military applications*", Proceedings of the 41st Annual Frequency Control Symposium, 1987, pp. 59-65.

(6) P. Thomann, H. Schweda, G. Busca (Oscilloquartz SA), "*Commercial size optically pumped cesium beam resonator*", Proceedings of the 2nd European Frequency and Time Forum (EFTF), March 1988, pp. 513-520.

(7) P. Thomann, F. Hador (Oscilloquartz SA), "*Short term stability of a commercial size optically pumped Cs tube*", Proceedings of the 3rd European Frequency and Time Forum (EFTF), March 1989, pp. 269-273.

(8) L.L. Lewis (Ball Communications Systems Division), "*Miniature optically pumped cesium standards*", Proceedings of the 45th Annual Frequency Control Symposium, 1991, pp. 521-533.

(9) K. Hisadome, M. Kihara (NTT Transmission Systems Laboratories), "*Prototype of an optically pumped cesium beam frequency standard*", Proceedings of the 45th Annual Frequency Control Symposium, 1991, pp. 513-520.

(10) B. Boussert, P. Cérez, L. Chassagne and G. Théobald, "*Frequency evaluation of a miniature optically pumped cesium beam clock*", Proceedings of the 11th European Frequency and Time Forum (EFTF), March 1997.

(11) C. Audoin, V. Giordano, N. Dimarcq, P. Cérez, P. Petit and G. Théobald, "*Properties of an optically pumped cesium beam frequency standard with $\phi = \pi$ between the two oscillatory fields*", Proceedings of the 7th European Frequency and Time Forum (EFTF), March 1993, pp. 537-540.

TELECOMMUNICATIONS

Chairman: K. Hilty

HOW TO SYNCHRONISE TELECOMMUNICATIONS NETWORKS

Sover Wong

Oscilloquartz S.A., CH-2002 Neuchâtel, Switzerland

ABSTRACT / KEY WORDS

Network synchronisation is very important for switching, transmission and data networks, such as telephony, Plesiochronous Digital Hierarchy (PDH), Synchronous Optical Network (SONET), Synchronous Digital Hierarchy (SDH), and Asynchronous Transfer Mode (ATM) networks. It is essential that the telecommunication systems in these networks are synchronised to meet the Quality of Service (QoS) demanded by the network users. This paper details the object of network synchronisation, the related international specifications, the general synchronisation network design requirements, plus the merits and limits of different synchronisation methods.

Key words: Network synchronisation,
Synchronisation networks

1. INTRODUCTION

1.1 Telecommunication nodes

There are usually many nodes in a telecommunications network. A node can be one system or many systems connected together, eg a radio base station or a telephone switching and data transmission centre. Generally every telecommunication system within a node is synchronised to the node clock (Ref.4).

1.2 Node clock

A node clock is the 'heart' of a node. It can be inside a telecommunications system, eg a Digital Cross-Connect (DCC) system, or a Stand Alone Synchronisation Equipment (SASE). The node clock is also called the Building Integrated Timing Supply (BITS) or Synchronisation Supply Unit (SSU).

2. THE OBJECT OF NETWORK SYNCHRONISATION

If the node clocks in a telecommunication network operate asynchronously then the transmit and receive rates of the telecommunication systems in each node would be different to the other nodes. In this case, the input buffers of the telecommunication systems would frequently overflow or underflow, causing data errors commonly referred to as slips. (Ref.1, Ref.2, Ref.5)

The object of network synchronisation is therefore to avoid and to minimise slips. This can only be achieved by synchronising all the node clocks, and hence all the telecommunication systems, to the same master clock or to a number of pseudo-synchronous (very closely matched, nearly synchronous) master clocks.

In practice, master clocks or Primary Reference Clocks (PRCs) are cesium beam oscillators, and slave node clocks are usually Ovenised Crystal Quartz Oscillators (OCXOs).

The interruption of different services due to slips is shown in table 1. The slip rate between systems can be calculated by equation 1, and slip rates for 8k frames per second signals under various frequency differences are shown in table 2.

Service	Effect of slips
Voice (uncompressed)	Only 5% of slips will lead to audible clicks
Voice (compressed)	A slip will cause a click
Facsimile	A slip can wipe out several lines
Modem	A slip can cause several seconds of drop out
Compressed video	A slip can wipe out several lines. More slips can freeze frames for several seconds
Encrypted data protocol	Slips will reduce transmission throughput

Table 1: Effect of slips on services

Equation 1: $\text{Slips per day} = \text{frequency difference} \times \text{traffic frames per second} \times \text{seconds per day (86400)}$

Frequency difference between systems	Slip rate for 8k frames per second signals
0	0
10^{-11}	1 slip in 4.8 months
10^{-10}	1 slip in 14.5 days
10^{-9}	1 slip in 1.45 days
10^{-8}	6.9 slips per day
10^{-7}	2.9 slips per hour
10^{-6}	28.8 slips per hour
10^{-5}	4.8 slips per minute

Table 2: Frequency differences and slip rates

Apart from frequency difference, wander levels that exceed the input tolerance of telecommunication systems would also cause slips. Wander is slow modulation of the clock or traffic signals from their ideal positions in time (Ref.5) and very low frequency (μHz) wander is impossible to filter out in a synchronisation network.

Contraction and expansion of transmission cables under varying temperatures generate very low frequency wander on the traffic/synchronisation signals. The levels of wander generated by optical fibre and copper cables are shown in equations 2 and 3, respectively.

Equation 2: Wander generated by ≈ 80ps/km/°C optical fibre cable

Equation 3: Wander generated by ≈ 725ps/km/°C copper cable

Node clocks and telecommunication system clocks can generate high amplitude (μs) wander, if they are badly designed or if their OCXOs have drifted in frequency over many years in operation (ageing). Also SONET and SDH networks can generate high amplitude wander on their tributary outputs. (Ref.10 and Ref.11)

3. INTERNATIONAL TELECOMMUNICATION UNION (ITU) RECOMMENDATIONS

3.1 ITU-T G.823 (Ref. 9)

The relative wander tolerance of telecommunication systems in 2 Mbit/s digital networks must be at least 18μs. In practice, however, the relative wander tolerance of many telecommunication systems is only just above 18μs. Therefore a relative wander level above 18μs would generally cause slips and interruption to services.

3.2 ITU-T G.811 (Ref.6)

The minimum frequency accuracy, ie the maximum frequency offset from Co-ordinated Universal Time (UTC) for a PRC is 10^{-11}.

Therefore the maximum frequency difference between any two PRCs is $2x10^{-11}$, and the maximum slip rate between two PRC synchronised (sub)networks is 1 slip in 2.4 months for 8k frames per second signals, eg 64 kbit/s and 2 Mbit/s signals.

3.3 ITU-T G.822 (Ref.8)

For an end-to-end inter-national tandem traffic connection as shown in figure 1, the nominal slip rate is 1 slip in 72/(n-1) days, where n is the number of pseudo-synchronous PRCs along the tandem traffic connection. Note that this equation is also applicable to intra-national tandem traffic connections.

For category (a) traffic performance, the maximum slip rate is 5 slips per day in 24 hours, for greater than 98.9% of time. According to equation 1, the overall maximum frequency difference along a pseudo-synchronous tandem traffic connection is $7.2x10^{-9}$ and only for less than 1.1% of time.

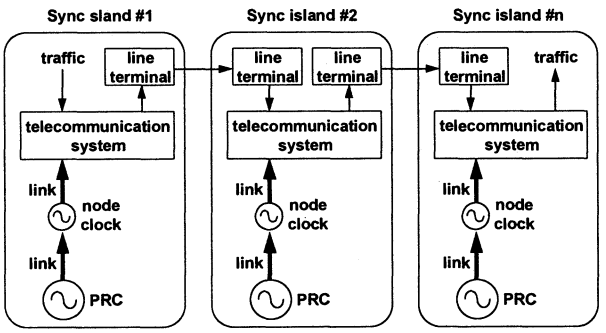


Figure 1: Pseudo-synchronous network model

The relationship between category (a) traffic performance, PRC availability, node clock availability, and link availability can be expressed by equation 4.

Equation 4: $0.989 \approx (PRC_{avail})^x \times (\text{node clock}_{avail})^y \times (link_{avail})^z$

x = Total number of PRCs along the end-to-end tandem traffic connection

y = Total number of node clocks along the end-to-end tandem traffic connection

z = Total number of links between the PRCs and node clocks, and the total number of links between the node clocks and the telecommunication systems along the end-to-end tandem traffic connection

To meet ITU-T G.822 category (a) performance, according to equation 4, the availability of each PRC, node clock and link must be >>0.989 (typical network operator requirement is 0.9995). Non availability of a local and transit node clocks to category (a) performance is when they have lost all their PRC synchronisation network connections, ie when they are in hold-over mode and that their output frequencies has drifted beyond $7.2x10^{-9}$ and $3.6x10^{-9}$ (see section 3.4), respectively.

3.4 ITU-T G.812 (Ref.7)

For transit node clocks, the maximum frequency off-set when entering hold-over mode is $5x10^{-10}$ and the maximum frequency drift whilst in hold-over mode is 10^{-9} per 24 hours. According to equation 1, ITU-T G.822 category (a) traffic performance is violated when a transit node clock remains in the hold-over mode for longer than 3.5 days (halved from equation 1 because slips occur at the transit node in hold-over and at the next PRC synchronised transit or local node). Therefore the repair time of a transit node failure must be less than 3.5 days.

For local node clocks, the maximum frequency off-set when entering hold-over mode is 10^{-8} and the maximum frequency drift whilst in hold-over mode is $2x10^{-8}$ per 24 hours. According to equation 1, ITU-T G.822 category (a) traffic performance is violated when a local node clock remains in the hold-over mode for longer than six hours. Therefore local (and transit) node clocks must be protected against single failures.

4. GENERAL SYNCHRONISATION NETWORK DESIGN REQUIREMENTS

To protect the synchronisation network against single failures, the following redundancies are necessary:

a) The PRC is internally or externally duplicated or triplicated, ie 1+1 or 1+2 protected.

b) The node clocks are internally 1+1 protected.

c) The node clocks have two or more diverse connections to a PRC.

d) The telecommunication systems have two or more connections to their node clock.

Additionally, any autonomous protection switching in the synchronisation network must not cause further network synchronisation problems, especially timing loops (slave clocks synchronising back to themselves). The long term frequency offset caused by a timing loop can be 10^{-7} or higher, which would seriously degrade the quality of many services, eg telephony signalling systems and Global System for Mobile communications (GSM) base stations.

When a timing loop is unknowingly created, it can be very difficult to find and break. This is because the error is often hidden and that the frequency offset is not alarmed by any slave clock or telecommunication system. Also it is impossible to determine the head end of the timing loop since there is no provision of synchronisation trace identifier in transmission signals.

5. METHODS TO SYNCHRONISE TELECOMMUNICATION NETWORKS

5.1 Centralised master clock network synchronisation

The centralised master clock synchronisation network has only one active master clock, as shown in figure 2. The master clock is logically located at the centre of the synchronisation network, and the node clocks are either directly or indirectly connected to it.

Since every node clock and telecommunication system clock operate at the same frequency as the master clock, there is no difference in the transmit and receive traffic rates between nodes. Therefore the end-to-end on-net traffic slip rate is nominally zero, when there is no failure in the network.

To construct a centralised master clock synchronisation network, it is important to ensure that the synchronisation links would not generate high amplitude (μs) wander on the synchronisation signals. Therefore very long overground cables that are subject to wide temperature changes, and SONET or SDH tributary connections should not be used as synchronisation links.

The synchronisation signals should only be transported over SONET or SDH connections as aggregate signals, or over pure PDH connections as tributary signals. Therefore only network operators who own all their SONET, SDH or PDH transmission links could implement centralised master clock network synchronisation.

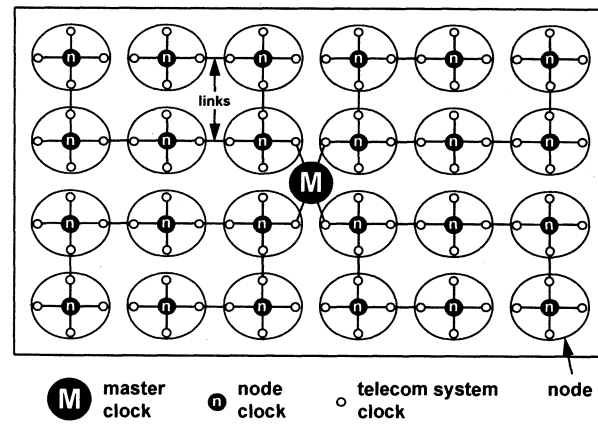


Figure 2: Centralised master clock network synchronisation

If the level of wander in a centralised master clock synchronisation network exceeds 18μs, then it would be necessary to partition it into several centralised master clock synchronisation subnetworks.

5.2 Distributed master clocks network synchronisation

The distributed master clocks synchronisation network has a number of active pseudo-synchronous master clocks. It is actually a collection of small autonomous centralised master clock synchronisation subnetworks (islands) grouped together.

A group of small synchronisation subnetworks is easier to plan and implement than a large synchronisation network. There is less wander in the subnetworks, as the PRC synchronisation subnetwork connections are shorter. Also the chance of a timing loop being accidentally created is significantly reduced, since the synchronisation subnetworks are smaller. However, the nominal end-to-end on-net traffic slip rate is 1 slip in 72/(n-1) days where n is the number of pseudo-synchronous PRCs along a tandem traffic connection.

It is technically feasible to deploy a fully distributed master clocks synchronisation network, as shown in figure 3, but it would be too expensive to deploy a cesium clock in every node. An economical method to generate the required master clock signals is to use the timing from navigation receivers, eg Global Positioning System (GPS), to discipline the node clocks.

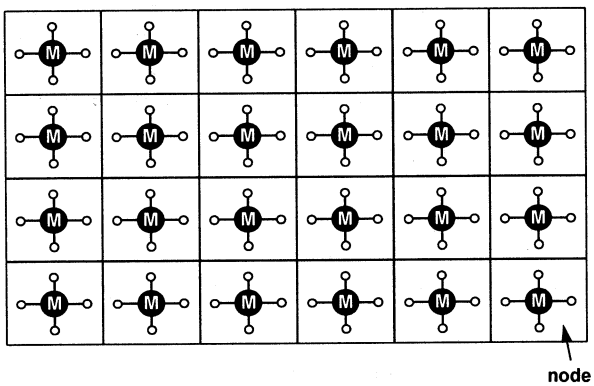


Figure 3: Fully distributed master clocks network synchronisation

Although a basic GPS receiver is small and inexpensive, it must be connected to a relatively large and expensive SASE to obtain PRC performance. Ideally two GPS receivers should be deployed in each node to provide the necessary protection and availability. Therefore it is not economical to deploy fully distributed GPS master clocks network synchronisation for very large networks, with more than 40 nodes as shown in figure 4.

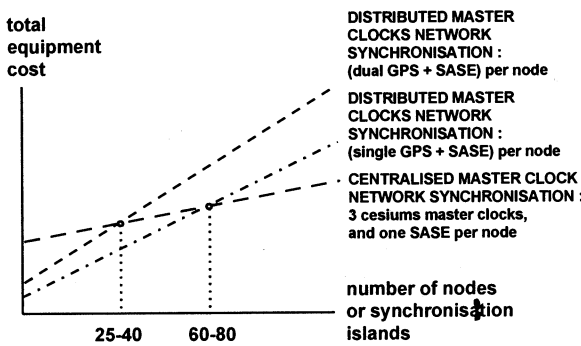


Figure 4: Cost break points of centralised master clock versus distributed master clocks network synchronisation

It is more economical to deploy a partially distributed master clocks synchronisation network as shown in figure 5, where each master clock is responsible for the synchronisation of a regional subnetwork.

Note that the off-air signals from GPS satellites can be intentionally or unintentionally jammed by a local transmitter and they can be blocked by some obstruction. In some areas, local authority planning permission is required to install the unsightly antennas. Furthermore, the active GPS antennas can be damaged by lightning hits, and that the length of the coaxial cable from the antenna is limited (typically 100m without a down-converter).

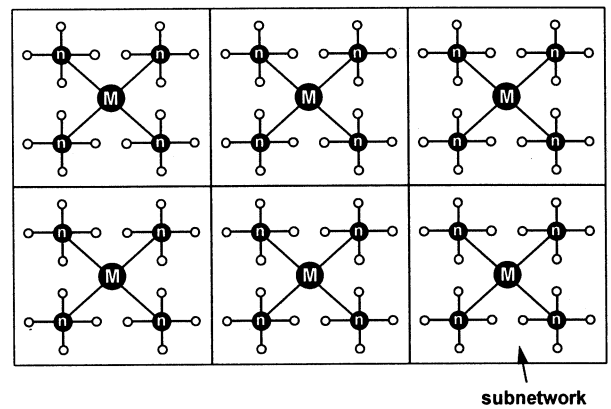


Figure 5: Partially distributed master clocks network synchronisation

Since the availability of the GPS transmission or reception cannot be guaranteed, it is prudent to use as many GPS receivers as necessary and as few as possible, or only use the GPS timing signals as back-ups.

5.3 Clock signals from a co-operating network

If a co-operating (or adjacent) network has master clock signals that are easily accessible then a network planner could, in theory, use them to synchronise his network. In this case his synchronisation network is slaved to the co-operating network, and both networks are operating at the same frequency. Therefore the slip rates for on-net and off-net traffic to the co-operating network are nominally zero. However, any disturbance in the master synchronisation network would also disturb the slave synchronisation network. Hence the quality (eg frequency off-set, wander amplitude and signal availability) of the clock signals from the co-operating network must be fully specified and guaranteed before they are used.

Clock signals from a co-operating network can be received at a few synchronisation gateway nodes only, as shown in figure 6, or at every node as shown in figure 7. Figure 6 is equivalent to centralised master clock network synchronisation, and figure 7 is equivalent to fully distributed master clocks network synchronisation.

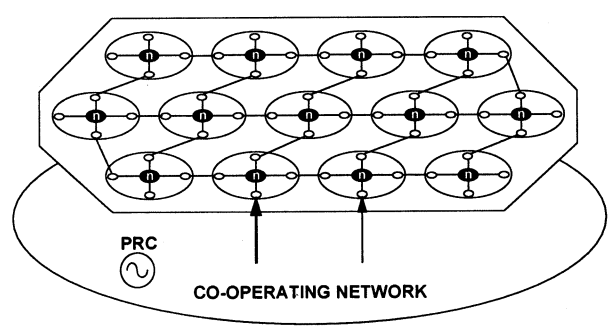


Figure 6: Clocks signals from a co-operating network to two gate way nodes

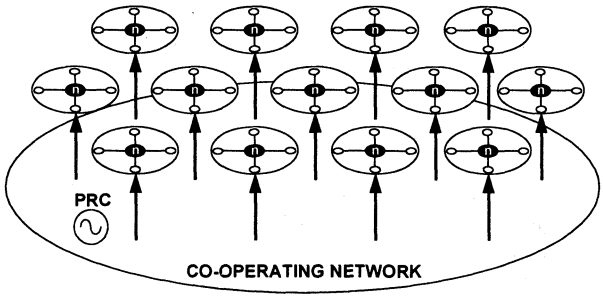


Figure 7: Clocks signals from a co-operating network to every nodes

If a co-operating network operator guarantees the master clock signals, then there is usually a significant user fee on each of them. The cost to lease many clock signals could be very high, and therefore it would make this method uneconomical for large networks with many nodes as shown in figure 8.

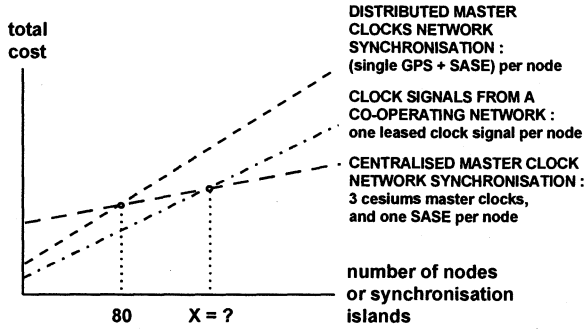


Figure 8: Cost break points of using clock signals from a co-operating network versus centralised and distributed master clocks network synchronisation

A low cost solution is to use a few guaranteed leased master clock signals from a co-operating network to synchronise a small number of subnetworks as shown in figure 9.

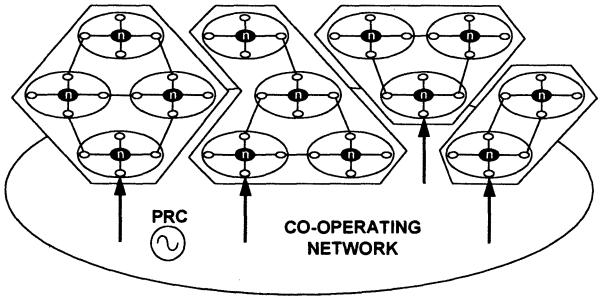


Figure 9: Clocks signals from a co-operating network to four subnetworks

6. WHICH NETWORK SYNCHRONISATION METHOD?

Table 3 summaries the features of different network synchronisation methods. For each synchronisation method, there is a maximum or minimum network size for optimum cost. A network synchronisation solution may be inexpensive when the network is small, but would become very expensive when the network grows beyond a certain size. Therefore, a network planner should weight the technical and financial merits of each network synchronisation method, for the foreseeable growth of his traffic network, before choosing a final solution.

Since each network synchronisation method has different merits and drawbacks, and that the structure of every telecommunications network is different, it is impractical to apply the same network synchronisation method to every network. Therefore a network planner should verify the practicality of each network synchronisation method on a realistic paper or computer model before choosing a final solution. It is possible that an optimum solution would involve all the aforementioned network synchronisation methods.

Figure 10 shows a synchronisation network partitioned into a few synchronisation subnetworks that are individually synchronised by a primary PRC and a secondary clock signal from a co-operating network.

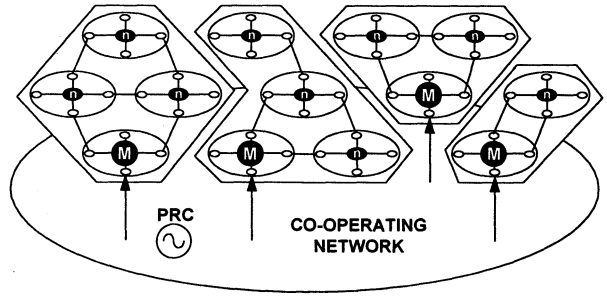


Figure 10: Hybrid network synchronisation

Synchronisation method	Network planning	Chance of creating timing loops	Network wander level	Nominal on-net slip rate	Synchronisation links required between nodes	Size of network for optimum cost
Centralised master clock	Difficult	Low if the sync plan is well designed	High for large networks covering wide areas	Zero, if the wander levels in the network are within limits	Own SONET or SDH aggregates. Own or leased tributaries over pure PDH connections	>40 nodes
Fully distributed master clocks	Very easy	Impossible	Negligible	1 slip in $\frac{72}{(n-1)}$ days n = number of master clocks	None	<40 nodes
Partically distributed master clocks	Easy	Very low for small subnetworks with well designed sync plans	Low for small subnetworks	1 slip in $\frac{72}{(n-1)}$ days n = number of master clocks	Own SONET or SDH aggregates. Own or leased tributaries over pure PDH connections	<40 subnetworks
Fully distributed clock signals from a co-operating network	Very easy	Impossible	Dependent on the co-operating network	Zero for on-net and off-net traffic to the co-operating network	None	Dependent on the cost of leased clock signals
Partially distributed clock signals from a co-operating network	Easy	Very low for small subnetworks with well designed sync plans	Dependent on the co-operating network	Zero for on-net and off-net traffic to the co-operating network	Own SONET or SDH aggregates. Own or leased tributaries over pure PDH connections	Dependent on the cost of leased clock signals

Table 3: Summary of features for various network synchronisation methods

7. CONCLUSIONS AND RECOMMENDATIONS

Each network synchronisation method has its unique merits or limits, and there is no generic solution for every network. However, the following recommendations are generally applicable:

a) If a traffic network is constructed using own PDH, SONET or SDH transmission links, then any aforementioned synchronisation method can be deployed. However, the optimum (cost versus performance) solution for a network with more than 40 nodes is centralised master clock network synchronisation.

b) If a traffic network is constructed using SONET or SDH tributaries (eg leased lines), then only fully distributed master clocks network synchronisation or fully distributed clock signals from a co-operating network can be deployed. Note that the latter method is only recommended if the clock signals are fully specified and guaranteed to have very high availability to ITU-T G.811 performance.

c) If a centralised master clock synchronisation network is found to have excessive levels of wander, ie when the Maximum Time Interval Error (MTIE) and Time Deviation (TDEV) measurements are above the international recommendations, then it is necessary to adopt distributed master clocks network synchronisation.

Also an initial synchronisation solution may evolve to other synchronisation solutions at different phases of the network. eg a network synchronisation strategy could be:

a) Phase 1 - fully distributed clock signals from a co-operating network

b) Phase 2 - partically distributed GPS master clocks and clock signals from a co-operating network as back-ups.

c) Phase 3 - centralised cesium master clock and partically distributed GPS master clocks as back-ups.

8. REFERENCES

1. Smith R. and Millott L.J. "Synchronisation and Slip Performance in a Digital Network", BT Engineering Vol 3 July 1984
2. Kearsey B.N. and McLintock R.W. "Jitter in Digital Telecommunication Networks", BT Engineering Vol 3 July 1984
3. Mulvey M., Kim I.Y. and Reid A.B. "Timing issues of constant bit rate services over ATM", BT Technical J Vol 13 No 3 July 1995
4. ITU-T G.803 "Architectures of transport networks based on the synchronous digital hierarchy (SDH)"
5. ITU-T G.810 "Considerations on timing and synchronisation issues"
6. ITU-T G.811 "Timing requirements at the outputs of primary reference clocks suitable for plesiochronous operation of international digital lines", 1988
7. ITU-T G.812 "Timing requirements at the outputs of slave clocks suitable for plesiochronous operation of international digital links", 1988
8. ITU-T G.822 "Controlled slip rate objectives on an international digital connection"
9. ITU-T G.823 "The control of jitter and wander within digital networks which are based on the 2048 kbit/s hierarchy"
10. Bellcore GR-436-CORE "Digital network synchronization plan"
11. Bellcore SR-NWT-002224 "SONET synchronization planning guidelines"

9. BIOGRAPHY

Mr. Sover Wong obtained his first Degree (Electronics Engineering) in 1979. Over the last 18 years, he has been appointed by GPT, Ascom, Ericsson and Alcatel as Principal Systems Design Engineer, SDH Product Specialist, SDH Systems Expert and SDH Product Manager, respectively. Currently he is the Telecom Technology Manager of Oscilloquartz S.A. Mr. Sover Wong is also a Chartered Engineer, a Member of the Institution of Electrical Engineers (MIEE), an active member of European Telecommunications Standards Institute (ETSI) and International Telecommunication Union (ITU).

A HIGH-PERFORMANCE TELECOMMUNICATIONS CLOCK SUPPLY EMPLOYING A NOVEL PHASE DETECTION TECHNIQUE

Ralph Urbansky, Harald Walter, Albrecht Rommel
Lucent Technologies GmbH Deutschland

Abstract - New transmission standards and services have raised the demand for improved network synchronisation quality which requires the deployment of high performance slave clocks. Two parameters have a significant impact on the performance of the clock: The oscillator stability and the phase detector characteristic. The phase detector should combine two contradictory requirements: A wide detection range and a high resolution to guarantee the required pull-in range and Maximum Time Intervall Error (MTIE) performance. The high MTIE performance in the slave mode results from a novel phase detector approach, which is based on sigma-delta-modulation techniques. The results are a high resolution of less than 30 ps and a large input phase range of 125 µs, which can be extended easily without affecting the resolution.

Keywords: Clock Supply, Maximum Time Intervall Error, Phase Detector, Sigma-delta-modulation, Synchronisation, Synchronous Digital Hierarchy, Synchronisation Supply Unit

1. Introduction

Network synchronisation has gained increasing attention since the introduction of the Synchronous Digital Hierarchy (SDH). The main advantage of SDH networks is the enhanced flexibility offered by the introduction of switching capabilities. The switching capabilities are provided by particular frame formats and bitrates, which are nominally synchronous to the payload signals to be switched. The synchronisation method employed by SDH is based on pointers which indicate the start of the payload frame with respect to the transport frame. The pointer processor aligns the transport frames for switching purposes while minimising the delay of the payload introduced in the pointer processor buffer. Therefore, the SDH pointer processing technique accommodates delay, delay variations (wander) and drift of the payload signal without data loss, thus enabling SDH network elements to be operated plesiochronously.

Pointer adjustments impose jitter onto the payload signal and compensate for wander which becomes apparent at the PDH interfaces. The SDH network synchronisation quality has an impact on the jitter and wander performance of digital connections, which may result in an increased probability of bit errors or frame slips, respectively.

In ITU-T Recommendation G.803 the currently recommended synchronisation technique comprises an SDH-based synchronisation distribution network. The synchronisation signals are carried via the SDH line signal and so-called SDH Equipment Clocks (SECs) in addition to the highly stable slave clocks according to ITU-T Rec. G.812. The timing characteristic of the SDH line signal is not impaired by wander due to buffer fill variations or by jitter associated with occasional bit rate justification events. The increased number of clocks in a chain implies the need for enhanced G.812 slave-clock specifications with respect to the MTIE performance in locked mode.

In [1] a low-cost SDH equipment clock employing a phase detector intended for integration into an ASIC has been proposed. This paper discusses the requirements with respect to highly stable slave-clocks, which are the basis for improved network synchronisation. Two parameters have a significant impact on the phase error generated by the clock: The oscillator stability in holdover mode and the phase detector performance (resolution) in locked mode (slave mode). The phase detector should combine two contradictory requirements: A wide detection range and a high resolution to guarantee the required pull-in range amd MTIE performance. The high MTIE performance in the slave mode results from a novel phase detector approach, which is based on sigma-delta-modulation techniques.

In Section 2 the functional requirements and the architecture of the clock supply will be presented. Section 3 examines some design aspects of phase locked loop (PLL) circuits intended to be applied in a clock supply. In Section 4 the novel phase detector will be presented. Section 5 shows measurement results of the clock supply.

2. Clock Supply Architecture

A synchronisation supply unit provides for the clock signal of SDH equipment or switches. This implies the need for a high degree of reliability and availability, which is covered by a suitable redundancy and protection concept in conjunction with control and monitoring functions provided by a management system.

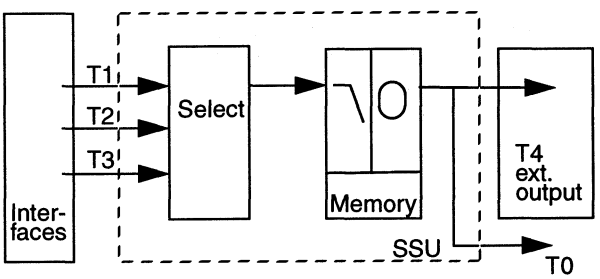


Fig. 1: SSU Clock Function

Fig. 1 shows the block diagram of a synchronisation supply unit (SSU) according to ETSI 3017-2 [7]. The timing signals T0 to T4 are:

- T0: Internal network element timing reference signal
- T1: Timing signal derived from STM-N input
- T2: Timing signal derived from 2 Mbit/s input
- T3: Timing signal derived from 2 MHz synchronisation input
- T4: External timing output

An SSU is a logical function which:

- accepts synchronisation inputs from a number of sources,
- selects one of these inputs,
- filters this sources clock and
- distributes the resultant clock to other elements within a node.

In the event of failure or degradation of all synchronisation reference inputs, the SSU will use an internal timing source. This mode of operation is called holdover.

The physical implementation of this function may be integrated within an SDH network element or may be integrated within a digital exchange of the Public Switched Telephone Network (PSTN). The new approach described in this paper was realised as a stand-alone unit, called Stand-Alone Synchronisation Equipment (SASE).

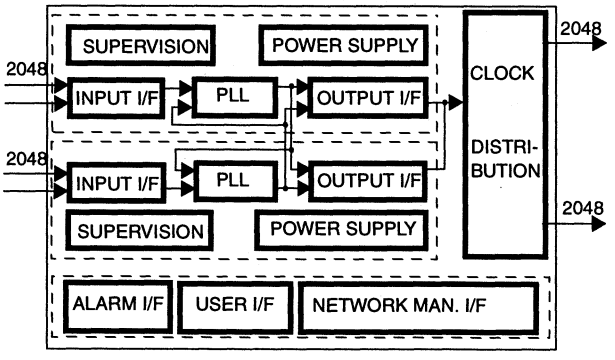


Fig. 2: SASE Architecture

The requirements mentioned before lead to the SASE architecture shown in Fig. 2. At least two synchronisation interfaces of 2048 kHz or 2048 kbit/s are required to provide for the protection of synchronisation links. Interfaces operating at the optical line bit rate (STM-N interfaces, 155.52 Mbit/s or multiples) are required only for SSUs integrated in SDH network elements.

As SDH network elements may be synchronised by T4 signals (2048 kHz input interface) 2048 kHz output interfaces are required at the SASE. For internode distribution at least two outputs are needed, for intra-node distribution even more, depending on the specific application. Therefore a clock distribution function should be integrated. The transport of timing quality information is possible with 2048 kbit/s outputs according to G.704. For 2048 kHz clock outputs a separate information channel for transporting quality information has to be used.

For the operator a user interface is necessary, in addition to an alarm and supervision interface. Network synchronisation management and monitoring requires a network management interface to allow for the integration in modern transport and mangement networks.

The whole system should offer a high degree of flexibility achieved by a modular design to allow for the upgrade to a Primary Reference Clock (PRC) or for the downgrade to simple clock distribution functionality. The hardware and software should be fault tolerant and highly reliable. Therefore all main plug-in units are duplicated and have their own on-board power-supply.

3. PLL Design

The main part of an SSU concerning the requirements of pull-in range, noise generation and transfer is the phase locked loop circuit. An intelligent control loop design with monitoring and supervision functionality for the inputs, outputs and PLL state guarantees high performance.

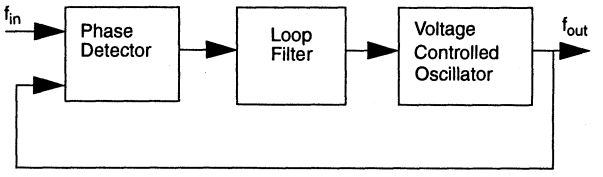


Fig. 3: PLL Block Diagram

Fig. 3 shows a general PLL block diagram. The digital PLL consists of a phase detector, a loop filter and a Voltage Controlled Oscillator (VCO). These components determine the characteristics of the PLL, which in turn determines the SASE performance.

The oscillator generates the output frequency. Its ageing and temperature characteristic is responsible for the holdover performance. Its short term stability determines the noise generation.

The loop filter supplies the VCO with the control signal. The low-pass characteristic of the PLL provides for jitter reduction capabilities. The parameters of the loop filter result in a cut-off frequency of 3 mHz. The filtering is performed by a micro-controller in a real-time system. The controller can handle simultaneously other tasks like frequency supervision, performance monitoring and data transfer to the network management interface.

The phase detector measures the phase difference between the input signal and the VCO clock. It needs a wide detection range for achieving a high pull-in range and input jitter and wander tolerance. For low noise generation and high stability the phase detector resolution should be less than 1 ns to have no significant impact on the overall clock supply phase error of 24 ns (for observation intervals between 0.1s and 9s) as per ETSI 3017-4 [9].

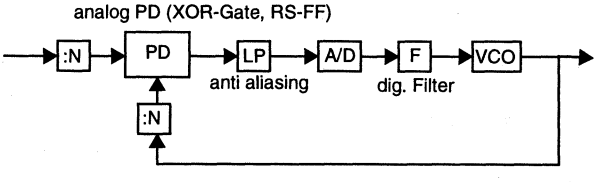


Fig. 4: Phase Detector Solutions (State-of-the-art)

The state of the art solution for phase detectors (PD) are counter detectors and Exclusive-Or (EXOR) or Flip-Flop (FF) detectors, [2] [3].

A counter detector counts the number of clock transitions of one input signal (e. g. input clock) with respect to a number of periods given by the other input signal (e. g. VCO clock). A wide detection range can easily be implemented and the interface to the control loop is digital. However, the phase resolution is determined by the operation frequency. As highly stable Oven

Controlled crystal Oscillators (OCXO) are normally operated in the 5 to 10 MHz range, frequency multipliers are needed to increase the resolution.

An FF or EXOR detector, as shown in Fig. 4, can provide a high resolution with respect to the operation frequency. However, the output signal is an analogue voltage and therefore requires an analogue-to-digital conversion preceding the loop filter. To extend the detection range, frequency dividers may be used before the phase detector at the cost of reduced resolution, which would imply the need for increased resolution of the A/D-converter. Therefore, in terms of a cost optimised solution, a wide pull-in range and a high resolution are contradictory requirements.

The phase detector presented in this paper resolves the contradiction of these requirements. This novel approach is based on sigma-delta-modulation techniques [4].

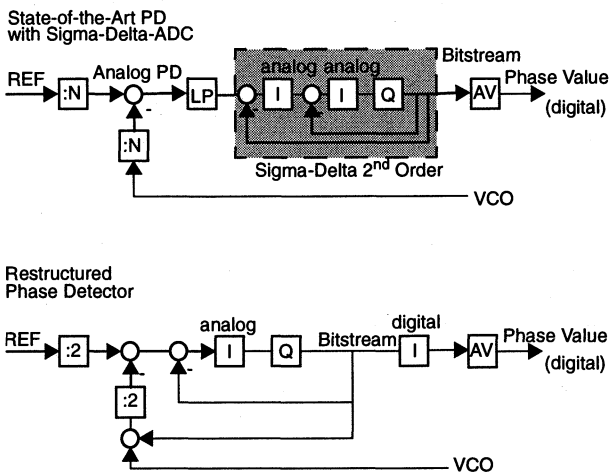


Fig. 5: Novel Phase Detector: Block Diagram

In Fig. 5 the A/D-converter of Fig. 4 has been replaced by a sigma-delta-modulator. As the sigma-delta-modulator may operate at a high clock frequency, the antialiasing-low-pass filter LP may be omitted. The order of the sigma-delta-modulator has a major influence on the resolution thanks to the spectral shaping of the quantisation noise. Therefore, for a certain resolution the clock frequency can be reduced by increasing the order to a second order modulator as shown in Fig. 5.

The frequency dividers preceding the phase detector may be considered as integrators, which translate the input frequency into a phase value. This allows the replacement of the first integrator of the sigma-delta-modulator by the combination of the phase detector with the frequency dividers. The outer feedback loop of the sigma-delta-modulator is implemented by applying the quantised signal as a disable signal at the VCO-related frequency divider.

This modified phase detector structure is complemented by an integrator, as basically frequency differences are compared. The word size of this integrator determines the pull-in range of the phase detector without overflow, which is independent on the division ratio of the frequency dividers. Therefore, the division ratio can be kept small to provide for a high resolution (noise immunity etc.) of the phase detector.

The resulting scheme contains only one analogue integrator. There are still two feedback signals of the sigma-delta-modulator. One of them directly influences the integrator in the VCO path. Therefore the phase difference at the PD doesn't exceed 1 UI (Unit Interval = $2 \cdot \pi$) and long dividers are not necessary. Phase detection is performed at 1024 kHz and the input integrators are only dividers by two.

A similar approach has beeen proposed in [5] for frequency synthesizers. In this application the operation frequency of the sigma-delta-modulator was divided by N, where the ratio depends on the output frequency.

4. A Novel Phase Detector

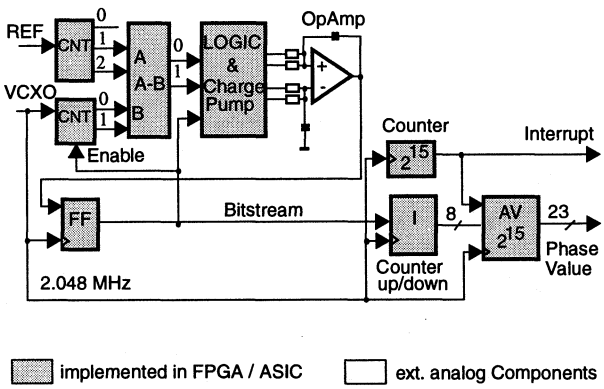


Fig. 6: Novel Phase Detector: Realisation

Fig. 6 shows the resulting implementation of the novel phase detector. The shaded components are part of a Field Programmable Gate Array (FPGA), the operational amplifier and its filter are the only analogue components.

The operational amplifier is a low cost standard type with a bandwidth-gain product of 4 MHz and a good common mode rejection. The digital part consists of 75 Flip Flops and combinatorial logic in a logic cell array.

The input counters and the differential circuit are implemented with two bits only. Because the reference clock is divided by two, the reference counter needs a third bit, but only the two most significant bits are used for the subtractor. The subtractors result is a two bit value. The output value of the subtractor now is combined with the bit stream. In steady-state the bit stream has equal number of high and low clock cycles.

For noise reduction the output of the charge pump is made differential. High impedance output is implemented by the same level at the differential output. In steady-state operation the integrators output voltage is about the FF input threshold voltage.

The resulting bitstream represents the frequency difference of the two input signals. It is then integrated to the phase difference in an 8 bit up-down-counter, that counts up when the bit stream is high and vice versa.

A 15 bit averaging filter (AV) now generates the 23 bit phase value. The resulting range is 256 UI, that means 125 µs with a resolution of 29.8 ps. After 2^{15} clock cycles of 2 MHz a new calculation cycle is started.

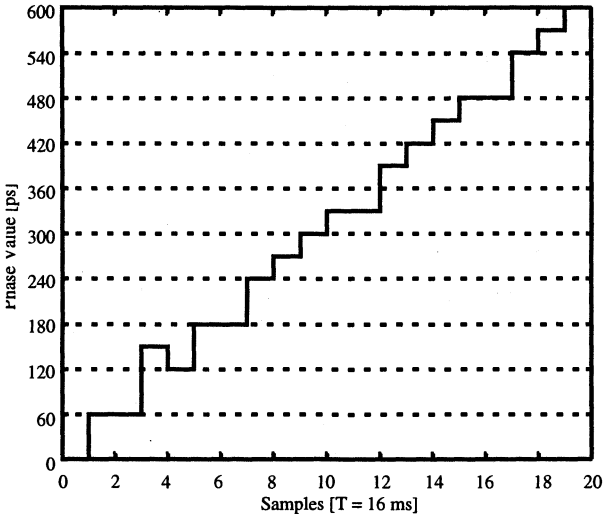


Fig. 7: Novel Phase Detector: Resolution

Fig. 7 shows the phase detector resolution. It is determined by the size of the averaging filter. A 30 ps resolution is expected thanks to the calculation of the mean value over 2^{15} values.

For this measurement two frequency generators at the input have been used with a frequency difference that was adjusted in a way that the phase value increased one bit for every sample, that means that the delta f was about 30 ps in 16 ms, that is about 3.84 mHz for 2048 kHz input. The irregularities in the expected staircase signal in Fig. 7 are due to the oscillator phase noise.

The phase detector characteristic allows for sufficient margin for other implementation imperfections and oscillator phase noise, because the required MTIE values according to ETSI 3017-4 [9] are 24 ns in the observation interval range of 100 ms to 9 s.

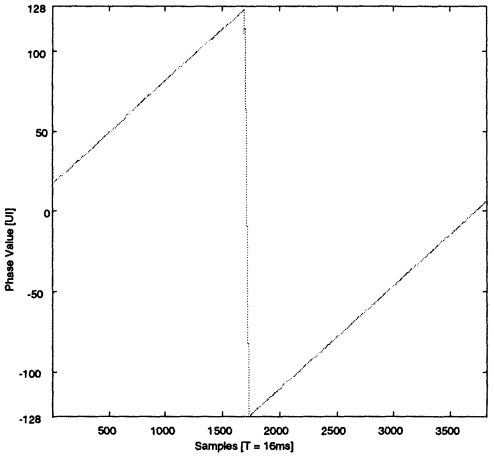


Fig. 8: Novel Phase Detector: Detection Range

Fig. 8 shows the whole detection range of the phase detector. It is +/- 128 UI. The wander range of a G.823 input signal is about 36.9 UI, that can easily be covered by this range.

5. Measurement Results

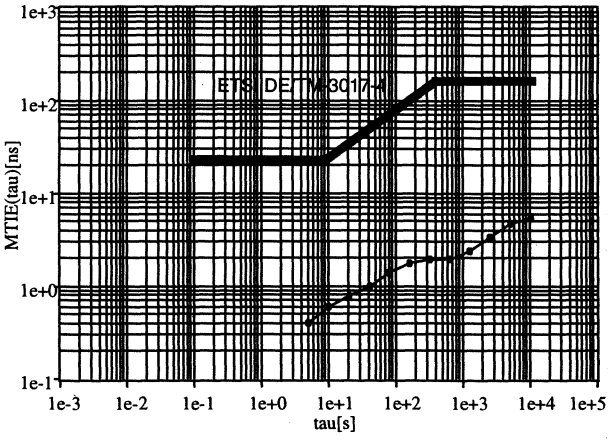


Fig. 9: Measurement Results MTIE

Fig. 9 and Fig. 10 show measurement results of the Maximum Time Intervall Error (MTIE) and the Time DEViation (TDEV), respectively. They were obtained on the basis of a hybrid phase locked loop using the novel phase detector, a digital loop filter, state-of-the-art digital-to-analogue converter and an OCXO.

The phase detector value has 23 bits, the filter calculation is made on a 64 bits basis, the D-A converter has 14 bit and the OCXO gain is about $6.7 \cdot 10^{-11}$ per bit. The last value is improved by software quantisation noise shaping technique at the filter output to $1.7 \cdot 10^{-11}$ per bit.

The oscillators Allan variance is about $5 \cdot 10^{-11}$ per second.

The graph in Fig. 9 shows results that are always 30 to 40 times better than the limits according to ETSI 3017-4 [9] for constant temperature. The sampling interval was 3.3 seconds, according to ETSI DE/TM 3017-3 [8].

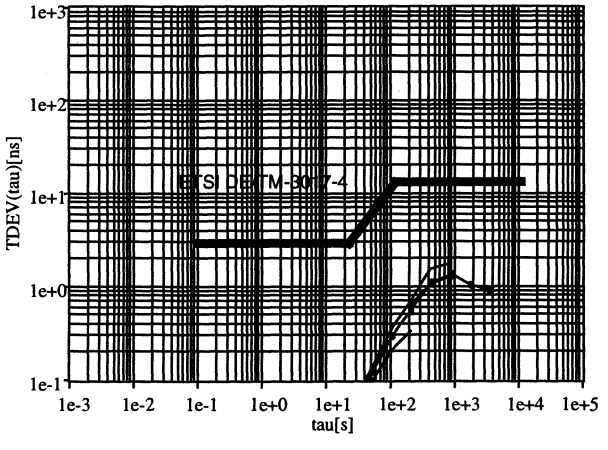


Fig. 10: Measurement Results TDEV

For the calculation of the TDEV results in Fig. 10 the algorithm described in ETSI DE/TM 3017-1 [6] was used. The results are in close agreement with the MTIE measurement results.

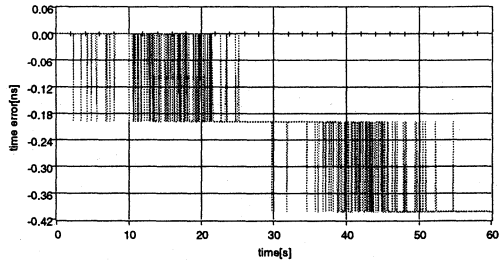


Fig. 11: SASE: Input Reference Switching

Input reference switching effects have been tested by applying two different reference signals with the same frequency but a phase offset. For the start of a new synchronisation process in the PLL the phase detector gets a reset on the up-down-counter from the microcontroller to adjust it in the middle of the total range.

The filter uses the first phase value as the reference value for the regulation in order to minimise the phase drift in case of input reference switching.

The resulting phase drift at the output, for input signals with same frequency but different phase is shown in Fig. 11. The measured phase drift is in the resolution range of the measurement set-up (200 ps). ETSI 3017-4 [9] allows a phase error of 60 ns for the new synchronisation process.

6. Conclusion

The introduction of Synchronous Digital Hierarchy (SDH) based transmission networks requires a new network synchronisation strategy and a new type of high performance clock supplies, called Synchronisation Supply Unit (SSU) or Stand-Alone Synchronisation Equipment (SASE). The logical function of a SSU defines also the SASE architecture.

Tightened requirements for Maximum Time Intervall Error (MTIE) performance and jitter tolerance lead to new types of phase detectors with high resolution and wide detection ranges.

A novel phase detector that is almost fully integrated and has few analogue components leads to cost effective phase locked loop solutions. Simulation and measurement results have demonstrated the feasibility and the performance of this approach.

7. References

[1] R. Urbansky, W. Sturm: "A Novel Slave-Clock Implementation Approach for Telecommunications Network Synchronisation", Proc. 10th EFTF, Brighton, UK, March 1996, pp. 534-539.

[2] D. H. Wolaver: "Phase Locked Loop Circuit Design", Prentice Hall, Englewood Cliffs, New Jersey, 1991.

[3] F.M. Gardner:"Phaselock Techniques", John Wiley & Sons, New York, 1979.

[4] J.C. Candy, G.C. Temes: "Oversampling delta-sigma data converters: Theory, design and simulation", IEEE Press, 1991.

[5] Beards, Copeland: "Oversampling Delta-Sigma Frequency Discriminator", IEEE Transactions on Circuits and Systems II:Analog and Digital Signal Processing, Vol.41, No.1, January 1994.

[6] ETSI DE/TM 3017-1 (prETS 300 462-1) "Definitions and terminology for synchronisation networks", March 1996.

[7] ETSI DE/TM 3017-2 (ETS 300 462-2) "Synchronisation network architecture", September 1996.

[8] ETSI DE/TM 3017-3 (prETS 300 462-3) "The control of jitter and wander within synchronization networks", April 1996.

[9] ETSI DE/TM 3017-4 (prETS 300 462-4) "Timing characteristics of slave clocks suitable for synchronization supply to Synchronous Digital Hierarchy (SDH) and plesiochronous Digital Hierarchy (PDH) equipment", September 1996.

The Applicability of MTIE and TDEV for Assessing the Quality of Telecom Services

Michael Wolf

Alcatel Telecom

Telecom services are affected by phase noise effects in the telecom network. Switched services can be degraded by byte slips and leased line service quality may be degraded by phase transients and temporary frequency offsets. It is up to the designers of telecom equipment and telecom networks to limit such degradations

- **by limiting phase noise generation by the equipment,**
- **by limiting phase noise accumulation in the network and**
- **by designing phase noise tolerance into input ports of the equipment.**

The designer's job is supported by related ITU and ETSI standards which are quite stable since a long time in the jitter domain (phase noise of typical frequency above 20 Hz), but which are under development for the wander domain, i. e. for long term phase fluctuations.

The Maximum Time Interval Error (MTIE) parameter which limits the phase offset magnitude over an observation interval has been established for telecom purposes by CCITT since long. This parameter nicely matches with the requirements for designing buffers and clock filter bandwidth.

The TDEV parameter has been newly introduced in standards as an additional parameter for clock stability specification. Although is allows for analysing the phase noise type of noise sources in terms of phase and frequency modulation it is not clear how equipment design parameters can be derived from TDEV specifications and how service performance can be assessed from data signal TDEV performance.

The paper shows that the TDEV parameter relates to the phase noise accumulation power for accumulation of MTIE over a number of noise sources, but it is not known how the accumulation power can be calculated from given TDEV to MTIE ratios. In the worst case the accumulation is linear with the number of cascaded noise sources and in the best case it is virtually non existent. Knowledge about phase noise accumulation impacts the design aspects of networks and thus impacts the service quality performance and the need to implement phase noise reducing equipment in the network.

1. Impact of Phase Noise on Transmission Performance

All synchronously operating switching and multiplexing telecom equipment needs to synchronise the incoming data signals to the internal system clock. This is performed by a buffer which is written with the incoming data clock and which is read with the system clock. The control of buffer fill in the case of a phase wander of the input signal against the system clock is done by controlled slipping or by a stuffing technique. In the case of a slip a part of the payload signal is corrupted, in the case of a stuffing event the payload data remains unaffected but the phase of the data signal is shifted which normally leads to a saw tooth phase movement over time.

The system clock can either be provided from external by a synchronisation reference input, or it may be derived from the incoming data signal applying appropriate filtering functions (see figure 1).

For appropriate equipment design for input phase noise tolerance the standards specify a minimum peak to peak sinusoidal phase modulation which should not cause any malfunction of the equipment [1]. This modulation amplitude as a function of modulation frequency complies to the Maximum Time Interval Error (MTIE) as a function of the observation in-

terval which is specified for network nodes (see figure 2).

Accommodation of incoming data phase fluctuation is performed by a suitable buffer design when reference clock is supplied from external. When the phase offset exceeds the buffer limit in the long term (days or weeks) controlled slips are performed in order to manage the buffer fill. This is the technique applied by digital switches based on 64 kbit/s. Instead of controlled slips the transmission equipment applies the stuffing technique which allows to increase or to decrease the number of bits transported per unit time by replacing dummy bits in a multiplex signal by data bits or vice versa. While slips [2] corrupt the data a stuffing event introduces a phase hit on the data, but keeps the data unaffected.

In the case of derived clock the incoming phase fluctuations are accommodated by a suitable buffer design for modulation frequencies above the clock filter cut off frequency. A suitable design of filter bandwidth and buffer size guarantees an error free operation of the equipment (see figure 3).

Newly introduced in the standards is the TDEV parameter as an additional specification for output phase noise and input tolerance. Compliance of the TDEV specification to the MTIE specification is ensured by simulation calculations. An direct impact of TDEV specifications to equipment parameters is not seen and it is questionable whether the clock models used in simulations do really reflect the phase noise type of telecom networks. A potential impact of the TDEV parameter on phase noise accumulation calculations is discussed in the course of this paper.

At the destination of the digital transmission the analog service signal is recovered from the received data. The time scale for this process is taken from the clock recovered from the received data. Now, when the data is affected by a phase transient the time scale is distorted by the associated temporary frequency offset. This causes a frequency shift of the recovered analog signal. A service particular sensitive to such distortion is video. The effect may cause the video colour subcarrier to exceed the tolerance and thus cause degradation of the video performance.

A further impact on the quality of telecom services is coming from bit error bursts caused by buffer overflow events or by controlled slips which both may be caused by excessive phase noise.

2. Characteristics of Telecom Phase Noise

Telecom networks employ a lot of equipment which is used for multiplexing and switching of payload data signals and for encoding them for transmission over the transmission medium. Each signal processing element generates phase noise and some of them do also filter it. And even the transmission medium contributes to the resulting phase fluctuations by temperature dependent propagation delay effects.

The essential question is how to meet the phase noise network limit specification. The network is composed of various equipment which can be described by its individual phase noise performance in terms of phase noise transfer characteristic, which is described in a linear approach by a filter bandwidth and the gain peaking of the transfer function, and by the intrinsic phase noise generation.

It is obvious that the payload signal path comprises of a chain of cascaded phase noise sources and some phase noise filters. But there is a second path of cascaded phase noise generating and processing units to be studied in the network. This is the so called synchronisation trail which links the reference clock input of a network element to the master clock of the network. The synchronisation trail is often orthogonal to the payload path (see figures 4 and 5). Therefore we have to study two aspects of phase noise accumulation in order to check whether the output phase noise limit is met in a particular network implementation (for an overview see references [5] and for the related standards see [6]).

Phase noise accumulation has been studied since a long time for Gauss'ian noise sources [3], but the amplitude distribution was given implicitly in these papers by assuming a white phase modulation spectrum of arbitrary phase of the spectral components. However, the power of accumulation depends on the statistics of the phase excursions of the intrinsic phase noise of an equipment. According to the MTIE specification the peak to peak value of these excursions is limited, but there are several probability distributions of phase amplitude which have the same peak to peak values (see figure 6). When the peak value is reached with a very low probability only the accumulation is of less power than in a case when the high phase offset is reached with a higher probability. As the TDEV(τ) parameter gives the phase noise power in a certain frequency band, which is defined by the observation interval τ, it relates to the effective width of the probability distribution.

We introduce three types of phase noise according to their different statistical behaviour, the type A having a Gauss'ian or even a rectangular distribution and which is of a "normal expected" statistical behaviour of noise processes. Then we have two extreme types of statistics, one of them we call the type B with a distribution which concentrates nearly 100% of all the probability close to the mean, and the other we call the

type C which concentrates nearly all the probability at the peak amplitude (see figures 7 and 8).

The type A phase noise is generated by

- Clock recovery from line signal.
- Bit stuffing and high rate byte stuffing.
- Clock frequency dividing by switched counters.
- Intrinsic phase noise of some clock equipment implementations.
- Propagation delay temperature effect.

The type B phase noise is caused by

- Intrinsic phase noise of some clock equipment implementations which are different from that which cause type A noise.
- Low rate byte stuffing which causes infrequent phase transients.

The type C phase noise has never been seen by the author but it is possible to create it by comparably simple digital phase control mechanisms.

The impact on the accumulation power is quite drastic: Although the accumulated phase offset is the total of all cascaded phase excursions and the worst case offset is the total of all particular MTIE's, the probability of this to happen within the observation interval τ is virtually zero, and this reduces the accumulation power seen in practice to approximately the fourth root of the number N of cascaded noise sources in the case of non correlated white phase modulation sources of type A. For the type B the accumulation is of much less power. In contrast to that the accumulation power of type C sources is one, which means linear accumulation proportional to the number N (see figure 9).

As all of these noise sources are assumed to meet the same MTIE specification the peak to peak phase magnitude is the same for all these sources, but the accumulated noise peak magnitude is quite different. The TDEV specification could help here, because the TDEV of these three sources is different, and the lower the TDEV of the source the lower is the accumulation power.

The example shown here is based on the white phase modulation type which applies when the master to slave chain is working correctly. This is valid for the payload path in any case and for the synchronisation trail in normal operation condition.

Non normal operation of the synchronisation trail is caused by failure effects or configuration management impact which causes input reference switching or hold–over operation. Input switching causes phase transients which represent a flicker phase modulation and holdover operation causes temporary frequency offset which is a flicker frequency modulation [4].

The impact of phase noise accumulation on the networks may be quite high. It is required to support a leased line service all over Europe which means that a payload data signal may be routed over hundreds of network elements which are all creating phase noise. The question is whether the signal meets the network jitter limit at the end or whether a jitter reducing filter function needs to be inserted somewhere, which would cause additional cost and effort for a network operator.

3. MTIE and TDEV – What can it be used for?

The MTIE parameter is suitable for specifying the input noise tolerance of equipment which is implemented by a suitable buffer size and suitable PLL parameters.

The output MTIE can be used to assess the payload performance in terms of maximum buffer overflow, i. e. slip, rates and in terms of a maximum temporary frequency offset which is just the quotient of $MTIE(\tau) / \tau$.

We have seen that the TDEV parameter does not directly map to any equipment parameter, but it relates to the phase noise accumulation power. But how could the output TDEV performance be specified for a single equipment in order to limit accumulation to a given power? This requires scientific research, and the author, who is a development engineer of telecom equipment and hence is not dedicated to research, wants to encourage researchers to work on this subject.

4. Conclusion and further relationship

The issue of phase noise accumulation has been discussed on the basis of the peak phase excursions characterised by the MTIE parameter. It has been shown that a noise power descriptor such as TDEV can describe the impact of the noise amplitude distribution on the accumulation law, but that the exact relationship requires further research.

During the study the author has found that the issue of phase noise accumulation relates closely to the ATM congestion problem which comes up with multiplexing a number of variable bit rate connections on a single line. The mean bit rate maps to the mean phase, the cell burst maps to the phase excursion and the output buffer maps to to operation range of a phase comparator. In ATM different parameters are used to characterise the statistics, but the accumulation problem is still unresolved as far as the author knows.

Acronyms

PLL	Phase Locked Loop
ETSI	European Telecommunication Standards Institute
ITU	International Telecommunication Union
CCITT	precursor organisation of ITU
MTIE	Maximum Time Interval Error
TDEV	Time Deviation

References

1 ITU Rec. G.823, The Control of Jitter and Wander within Digital Networks which are Based on the 2048 kbit/s hierarchy.

2 ITU Rec. G.822, Controlled Slip Rate Objectives on an International Digital Connection.

3 Some basic work on jitter accumulation has been performed a long time ago, see e. g.:

C. J. Byrne, B. J. Karafin, D. B. Robinson, Jr, "Systematic Jitter in a Chain of Digital Regenerators", The Bell Technical Journal 42 (November 1963):, pp. 2697–2714.

D. L. Duttweiler, "The Jitter Performance of Phase–Locked Loops Extracting Timing from Baseband Data Waveforms", The Bell Technical Journal 55 (January 1976), pp 37–58.

C. Chamas, "Accumulation of Jitter: A Stochastic Model", AT&T Technical Journal 64 (January 1985), pp 43–76.

B. Powell et al, "Additional SONET Islands Jitter Simulation Results", T1X11.3/92–017, (1992), a contribution to Standards Committee T1–Telecommunications, ANSI.

D. Cislow, G. Zampetti, "Cascaded Clock Measurement and Simulation findings", Time and Frequency Workshop on Synchronization Standards for Telecommunication Systems, NIST, Boulder, March 1994.

4 For the phase and frequency stability see:

IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology, IEEE Std 11 39 – 1988.

D. W. Allan, " Time and Frequency (Time Domain) Characterization, Estimation, and Prediction of Precision Clocks and Oscillators, IEEE Transactions UFFC–34 (November 1987), pp 647–654.

J. A. Barnes et al, "Characterization of Frequency Stability", IEEE Transactions IM–20 (May 1971), pp 105–120.

M. A. Weiss, "Time and Frequency Definitions and Models", Time and Frequency Workshop on Synchronization in Telecommunications Systems, NIST, Boulder, February 1996.

5 J. C. Bellamy, "Digital Network Synchronisation", IEEE Communications Magazine, April 1995, pp 70–83.

6 prETS 300 462 Generic Requirements for Synchronization Networks, ETSI.

Author

Dr.–Ing. Michael Wolf
c/o Alcatel SEL
Lorenzstr. 10
D–70435 Stuttgart
Germany

Tel: 0711 821 44 770
Fax: 0711 821 42 316
Internet: mwolf@lts.sel.alcatel.de

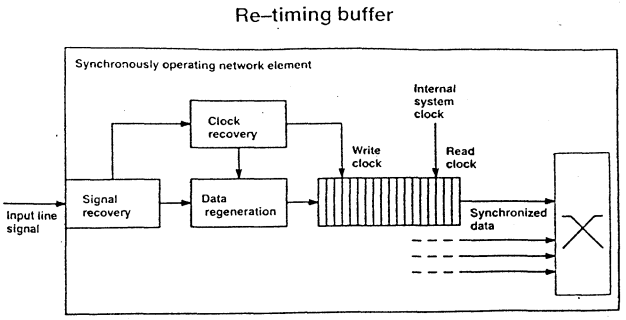


Figure 1: Synchronising the receive data to external reference clock (left) or to a filtered derived clock (right).

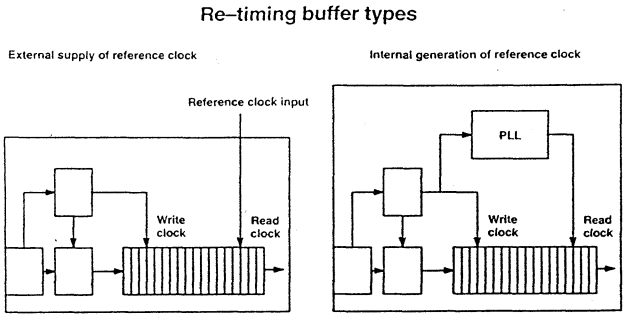


Figure 2: MTIE specification limits for the output phase noise at any network interface (left) and the corresponding tolerance requirement for sinusoidally phase modulated test signal at any equipment input (right).

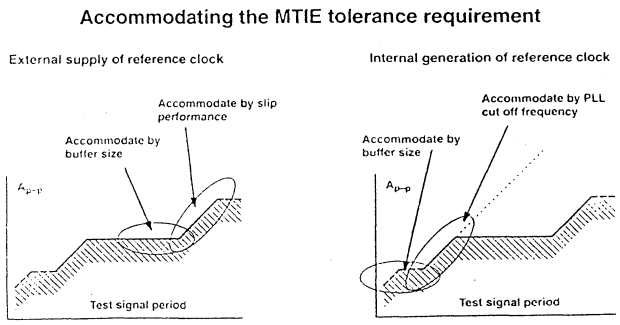


Figure 3: Accommodating short term phase noise by buffer size and the long term noise by slips in the case of external timing (left) and by suitable PLL design for derived clock timing (right).

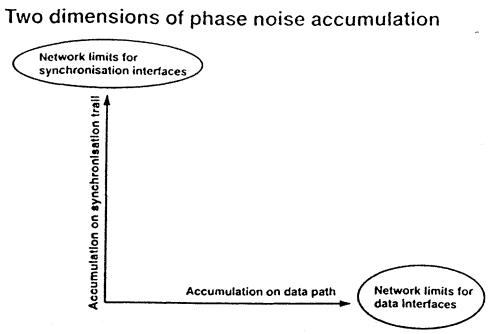


Figure 4: Accumulation of phase noise on the data path (horizontal) and on the synchronisation trail (vertical).

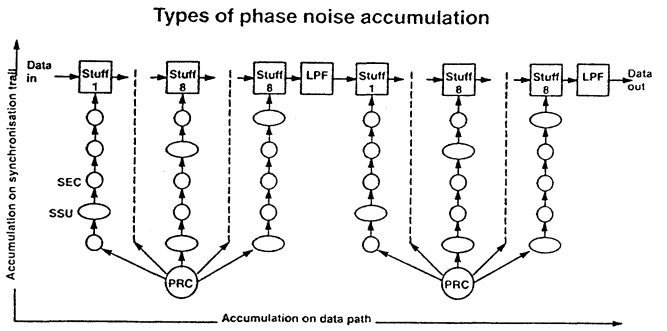


Figure 5: The data path consists of cascaded signal processing equipment (phase accommodation by bit or byte stuffing and low pass filtering of phase) and the sync trail contains cascaded slave clocks.

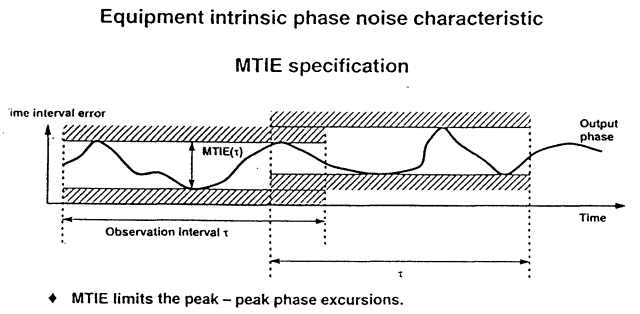


Figure 6: Limiting the equipment intrinsic phase noise by MTIE(τ) specification.

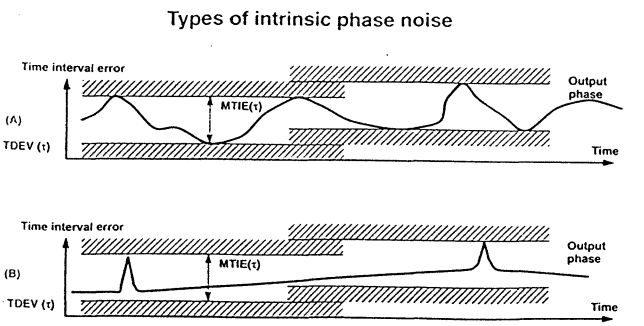


Figure 7: Two examples of phase noise touching the MTIE limit but with different amplitude statistics.

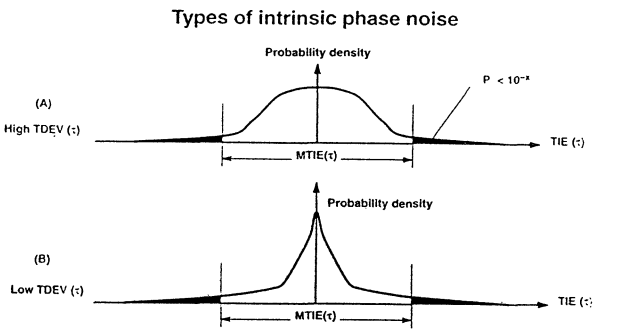


Figure 8: Amplitude statistics related to the signals fo figure 7.

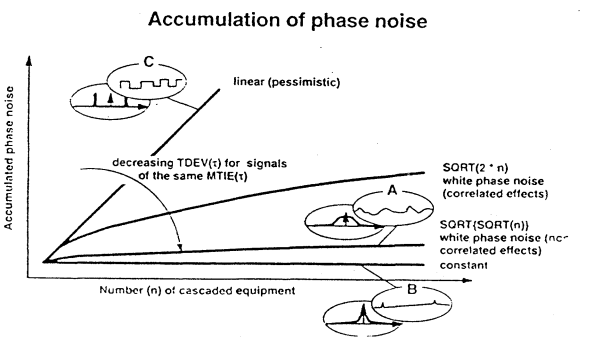


Figure 9: The accumulation power depends strongly on the amplitude statistics.

ROBUST STATISTICS AND DIAGNOSTICS OF TIMING SIGNAL

Andrzej Dobrogowski

Institute of Electronics and Telecommunications
Poznań University of Technology
60-965 Poznań, Piotrowo 3A, Poland
email: dobrog@et.put.poznan.pl

ABSTRACT

The paper explores the techniques of removing the effect of outliers on estimated values of parameters. Least median of squares method was used to achieve the robustness of estimation and for pinpointing outliers. Some preprocessing of the DPCM type is suggested to clean the timing signal from impulsive noise.

Keywords: *robustness, diagnostics, outlier, robust estimator, breakdown point, influence function, autoregressive model, signal cleaning*

1. INTRODUCTION

Timing signal of digital communication network belongs to the category of the most important signals of the network. It is the signal of global consequence for the digital network. The outstanding feature of the digital network timing signal is its omnipresence within the network. The timing signal provides all digital processes running in the network with time scale. A digital network timing signal originates in primary reference clock (PRC) or in primary reference source (PRS). PRC and PRS are very high quality clocks. However, in a course of transmission the timing signal undergoes pollution which has to be suppressed in synchronization supply unit (SSU) synchronized, hopefully, to unpolluted part of incoming timing signal. On the other hand pollution of timing signal, properly interpreted, provides information about transmission path. Therefore, omnipresent timing signal can be an agent for digital network monitoring system. First of all it provides information on synchronization process to a management system of synchronization layer. Timing signal may be used for time transfer throughout a digital communication network. Timing signal as a reference frequency signal and/or time signal may be article of trade between common or private carries.

The aforementioned lends a perspective to the subject of specifying timing signal. Proper identification of commonly accepted model of the timing signal (clock signal) should be done. Applied estimators should extract from received signal original features of timing signal despite of masking effects imposed on the signal during its transfer. Signal processing procedures ought to point out outliers, that is, deviations from the model followed by majority of data. Wanted signal processing algorithms come from robust statistics and diagnostics.

In the paper some comments about robustness and diagnostics are given. Observation outliers are discussed. Breakdown point and influence function are considered. Several robust estimators are presented. Least median of squares (LMS) regression and reweighted least squares (RLS) based on the LMS are used for finding frequency drift. Autoregressive model of timing signal is presented and its robust estimation is discussed. Processing of timing signal removing impulse noise is suggested.

2. ROBUST STATISTICS AND DIAGNOSTICS

A set of raw data resulting from an experiment can be roughly divided int two subsets. In a properly designed experiment the majority of raw data reflects mechanism which is responsible for investigated phenomenon. The rest of data (outliers) contaminates results which would have been provided by the majority of raw data itself. Thus the proper (pre)processing of raw data is needed and some suitable methods of statistics has to be incolved. In this case, methods of robust statistics and diagnostic are the right choice.

Robust statistics and diagnostics complement each other in a process of raw data analysis (Ref.1). Working with the whole set of raw data robust statistics provide us with information characterizing the "good" majority of data and pinpoint outliers ("bad" data). The study of outliers may lead to interesting conclusions, e.g., to explanation for origin of outliers or it may suggest introducing some change into assumed model. Diagnostics do the job in the opossite order. Using diagnostics tools we try to find outliers and delete (or correct) them first. As a result we get the "good" majority of raw data which we processed using means of nonrobust statistics.

It is worthwhile to quote Huber's claim (Ref.2): "The purpose of robustness is to safequard against deviations from the assumptions. The purpose of diagnostics is to

find and identify deviations from the assumptions. Outlier detection/rejection is diagnostics, not robustness. Robustness is making a procedure insensitive to outliers".

3. OUTLIERS

Outliers or gross errors are wrong (bad) values of measurement. "Badness" of raw data may result from many reasons. Let us enumerate only a few.

1. Assumed model is only an approximation of real world. Ocassionally remarkable deviations from the model are expected to appear because of very nature of investigated phenomenon.
2. An experiment supplies an investigator with quantized data.
3. When the quality of timing signal is investigated, rather long-lasting experiment carried over rather common conditions suffers from environmental changes.
4. Some wrong values of data result from copying or transmission errors.

When timing signal is investigated the most critical contaminations resulting in gross errors are sudden or very fast changes of phase of the signal. Even one such error may have a strong effect on final results of estimation. When the quality of synchronization of digital communication network is considered we can observe phase jumps because of several reasons. Among them are:

- fading of synchronization signal,
- temporary "out of synchronism" condition of different pieces of equipment of synchronization chain along which the synchronization signal is transmitted,
- switching events affecting synchronization chain.

When some signal is under study then very often outliers can be interpreted as an impulse noise. Studying timing signal we are trying to reveal some trends characterizing signal behaviour in time. Regression analysis helps us to do the job. In some situation simple linear regression method should recover a trend. Let us remind the simple regression model

$$y_i = \Theta_1 x_i + \Theta_2 + e_i \quad \text{for i} = 1,...,\text{N} \qquad (1)$$

in which the parameters Θ_1 (the slope) and Θ_2 (the intercept) are to be estimated. Variable y_i is called the response variable, x_i is called the explanatory variable (carrier), e_i is the error term, i is the case number and N is a total number of cases. We can explain regression outliers as follows. If there is an observation (x_i, y_i) shifted considerably up or down from the expected position (taking into account the suggestion of linear regression line), this point is called an outlier in the y-direction. When an observation (x_i, y_i) is shifted considerably to the left or the right from the expected position, this point is called an outlier in the x-direction. The last outlier is also called a leverage point bacause of the possible effect it has on least squares (LS) estimator of regression line.

4. BREAKDOWN POINT AND INFLUENCE FUNCTION

The breakdown point is a global measure of reliability of the estimator while the influence function (IF) is a local characteristics of the estimator (local robustness).

Introducing the concept of breakdown point we will follow explanation given by Rousseeuw and Leroy (Ref.1). Let us take any specific sample of N observations Z and let

$$T(Z) = \hat{\Theta} \qquad (2)$$

be an estimator of the vector of parameters. Then consider all possible corrupted samples Z' that are obtain by replacing any m of the original observations by arbitrary values. Bias(m; T,Z) that the corruption can caused is defined as

$$bias(m; T, Z) = \sup_{Z'} ||T(Z') - T(Z)|| \qquad (3)$$

where the supremum is over all possible Z'. The breakdown point $\varepsilon_N^*(T,Z)$ of the estimator T at the sample Z is

$$\varepsilon_N^*(T,Z) = \min\left\{\frac{m}{N}; bias(m; T, Z) \to \infty\right\} \qquad (4)$$

Thus the breakdown point is the smallest fraction of corruption that can forced the estimator T take on values infinitely far from T(Z).

The LS method gives regression estimator having the breakdown point equals

$$\varepsilon_N^*(T,Z) = \frac{1}{N} \qquad (5)$$

It means that one outlier can carry regression estimates over all bounds. We say that LS has the breakdown point of 0% because for increasing N right side of (5) goes to zero.

The meaning of the influence function is well understood from the relation

$$T(y_1,...,y_{N-1},y) \approx T(y_1,...,y_N) + \frac{1}{N} IF(y; T, f) \qquad (6)$$

In a "good" sample (y_1,...,y_N) we replace one observation by an outlier at y. IF indicates the influence of the outlier y on the value of the estimator T when the "good" majority of observations undergoes distribution f.

An important characteristic of robustness of the estimator T based on influence function is the gross-error sensitivity γ^*, defined as (Ref.1)

$$\gamma^* = \sup_y |IF(y; T, f)| \qquad (7)$$

Gross-error sensitivity gives the maximal effect induced by an outlier on the estimate.

5. ROBUST ESTIMATORS OF REGRESSION

The often used regresion estimator enables to find estimate $\hat{\Theta}$ of parameters Θ from condition

$$\underset{\hat{\Theta}}{Minimize} \sum_{i=1}^{N} r_i^2 \tag{8}$$

where the residual r_i is the difference between actual observation y_i and estimate $\hat{y}_i$ of observation

$$r_i = y_i - \hat{y}_i \tag{9}$$

The least squares (LS) estimator (8) behaves excellent for gaussian error e_i (in fact for this kind of error LS estimator is optimal) but very poor for error term with outliers. As we mentioned above the breakdown point of LS estimator is 0%.

Replacing the squared residuals by another function of the residuals yields the condition

$$\underset{\hat{\Theta}}{Minimize} \sum_{i=1}^{N} \rho(r_i) \tag{10}$$

where ρ is the even function of its argument with a unique minimum at zero. Estimators coming from the condition (10) are called M-estimators. Differentiating expression (10) with respect to the regression coefficients $\hat{\Theta}_j$ results in

$$\sum_{i=1}^{N} \psi(r_i)\mathbf{x}_i = \mathbf{0} \tag{11}$$

The solution of (11) is not equivariant with respect to a magnification of the y-axis (functions $\rho(\cdot)$ and $\psi(\cdot)$ do not "breathe" along the y-axis), and we have to standardize the residuals and solve rather

$$\sum_{i=1}^{N} \psi(r_i / \hat{\sigma})\mathbf{x}_i = \mathbf{0} \tag{12}$$

instead of (11). Solving (12) $\hat{\sigma}$ must estimated simultaneously. For $\psi(\cdot)$ Huber's proposal (Ref.1)

$$\psi(u) = \min(c, \max(u, -c)) \tag{13}$$

is often used. M-estimators are robust with respect to outliers in the y-direction but leverage pooints make them breaking down. It should be pointed out, however, that there are situations excluding possibilities of appearance of leverage points. For example, in time series analysis the values of explanatory variable (instant of time) are fixed in advance very often.

The least median of squares (LMS) estimator is defined by fulfilling the condition

$$\underset{\hat{\Theta}}{Minimize} \ \underset{i}{med} \ r_i^2 \tag{14}$$

LMS estimator is very robust with respect to outliers in both directions (in y as well as in x) (Ref.1). Its breakdown point is 50%, which is the highest possible value. The LMS regression estimate $\hat{\Theta}$ results in the narrowest vertical width of the strip surrounding LMS regression line and containing 50% of all residuals.

Weighted least squares (WLS) estimation method is defined by

$$\underset{\hat{\Theta}}{Minimize} \sum_{i=1}^{N} w_i r_i^2 \tag{15}$$

Knowledge of LMS residuals is used to establish values of weights. The possible choice of weights could be:

$$w_i = \begin{cases} 1 & if \ |r_i / \hat{\sigma}| \le 2.5 \\ 0 & if \ |r_i / \hat{\sigma}| > 2.5 \end{cases} \tag{16}$$

This is a policy of hard rejection of outliers. It can be applied soft or smooth rejection of outliers. WLS estimator possesses the high breakdown point being statistically more efficient then LMS estimator. Let us notice that the concept of WLS estimation reflects philosophy of diagnostics.

6. AR MODEL OF TIMING SIGNAL AND ITS ROBUST IDENTIFICATION

Results of measurement of timing signal usually take the form of times series, e.g., time error signal resulting from measurement. A time series is a sequence of N consecutive observation $y_1, y_2, \ldots, y_N$ measured at regular intervals. Often the observations obey the rule

$$y_i = \mu + x_i \ , \ i = 1, \ldots, N \tag{17}$$

where μ is a location parameter, and x_i are fitted by a zero-mean autoregressive moving average (ARMA) model. In what follows we consider autoregressive AR(p) model for x_i. According to that model

$$x_i = \alpha_1 x_{i-1} + \ldots + \alpha_p x_{i-p} + e_i \tag{18}$$

The larger value of p is used the better approximation of ARMA is given by AR. From (18) and (17) for measured (observed) value we have

$$y_i = \alpha_1 y_{i-1} + \ldots + \alpha_p y_{i-p} + \gamma + e_i \tag{19}$$

where the intercept γ equals $\mu(1-\alpha_1-\ldots-\alpha_p)$. As i ranges from p+1 to N there are N-p sets $(y_i, y_{i-1}, \ldots, y_{i-p})$ which can be used to estimate $\alpha_1, \ldots, \alpha_p, \gamma$ in (19) by means of some regression estimator. So we have matrix equation

$$\begin{bmatrix} y_{p+1} \\ y_{p+2} \\ \cdot \\ \cdot \\ \cdot \\ y_N \end{bmatrix} = \begin{bmatrix} y_p & y_{p-1} & \cdots & y_1 & 1 \\ y_{p+1} & y_p & \cdots & y_2 & 1 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ y_{N-1} & y_{N-2} & \cdots & y_{N-p} & 1 \end{bmatrix} \times \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \gamma \end{bmatrix} + \begin{bmatrix} e_{p+1} \\ e_{p+2} \\ \cdot \\ \cdot \\ \cdot \\ e_N \end{bmatrix} \tag{20}$$

Now we will use further a very simple AR (1) model

$$y_i = \alpha_1 y_{i-1} + \gamma + e_i \tag{21}$$

Stationarity of $\{y_i\}$ needs the absolute value of the slope α_1 to be less than 1. Plotting the observed values in pairs (y_{i-1}, y_i) in a diagram we get a scatterplot of y_i (ordinate) versus y_{i-1} (abscissa) which can be used to estimate the slope α_1 and the intercept γ of AR(1) process. In time series analysis two types of outliers pose the problems (Ref.1), viz, innovation outliers and

additive outliers (Ref.1). The innovation outlier results in on outlier in y direction (outlier in the response variable) and some number of "good" leverage point, which even can improve the accuracy of estimation. The additive outlier produce outlaying response variable and bad leverage point (in case of AR(1)). They say that additive outliers occur more often then innovation outliers. The possibility of appearance of outliers calls for robust procedure of parameters' estimation for time series (robust identification of time series model). The LMS or RWS linear regression methods can be applied.

7. IMPULSE NOISE PREPROCESSOR

Impulse noise disturbs information time series conveys very severely. E.g., it can seriously change AVAR and TDEV of timing signal, therefore the impulse noise should be suppressed before further processing of the signal. The DPCM (differential pulse code modulation) modulator (Ref.3) can be adopted for the impulse noise suppression. The arrangement is shown in Fig. 1. The specific solutions applied in (pre)processor in Fig. 1 and the effectiveness of the impulse noise supression are considered in literature (Ref. 4 and 5). Let us notice, that the Huber's $\psi(\cdot)$ function (13) can be used in the cleaning signal block of the preprocessor and the results of robust identification of the signal maybe exploited in prediction.

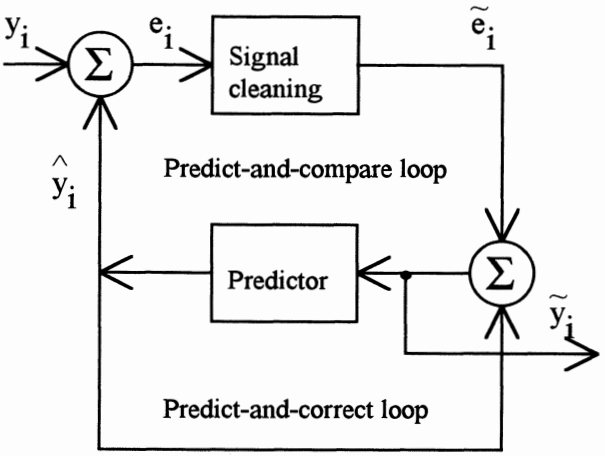


Fig. 1. Impulse noise removing preprocessor

8. THE EXPERIMENT

In the experiment time error between two timing sugnal was measured. Analog version of time error is given in Fig. 2. Then the first difference z_i and the second difference q_i processes of time series of time error were computed. Their analog versions are shown in Fig. 3 and Fig. 4 correspondingly. Using the LS method and then the robust LMS and WLS the existence of the first and second order driff in the observed (measured or computed) time error series was investigated. Program package of robust statistics, named PROGRESS (Ref. 1), was used for estimation.

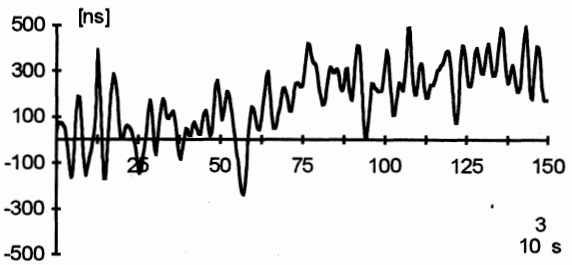


Fig. 2. Time error

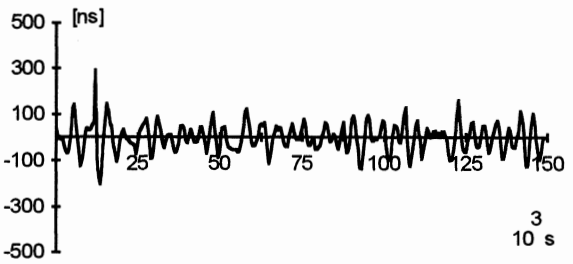


Fig. 3. First difference of time error

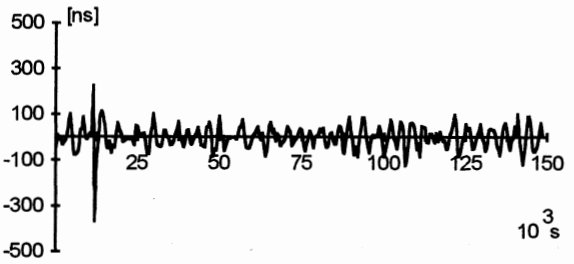


Fig. 4. Second difference of time error

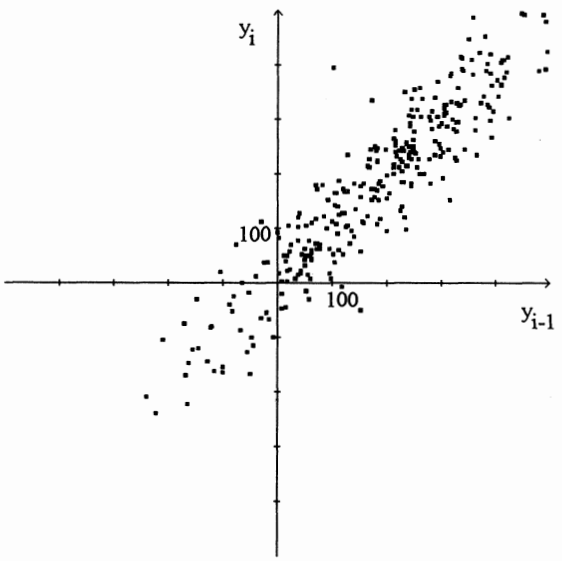


Fig. 5. Scatterplot $y_i=f(y_{i-1})$ of time error

Then the identification of the AR(1) model of the measured time error process was carried out. The data presented in the form of scatterplot in Fig. 5 were exploited for that purpose.

The first difference process z_i (the scatterplot given in Fig. 6) and the second difference process (the scatterplot given in Fig. 7) are also identified in therms of AR(1) model.

The results of experiment show that robust methods of statistics and diagnostics enable to make statistical inference free from the bias introduced by disturbance even if the relatively high quality signal is investigated.

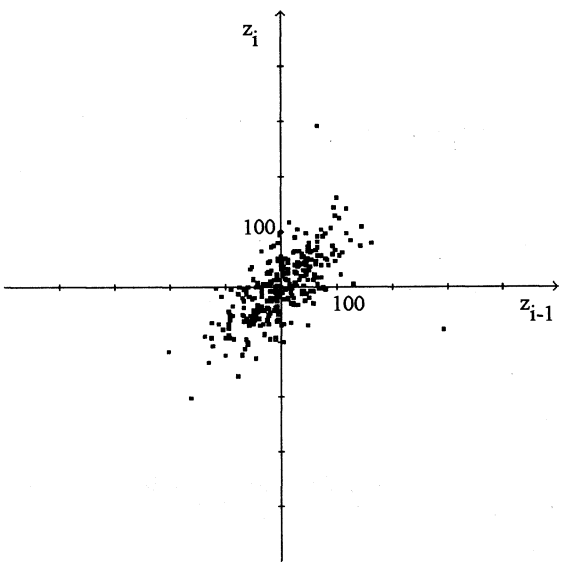


Fig. 6. Scatterplot $z_i=f(z_{i-1})$ of first difference of time error

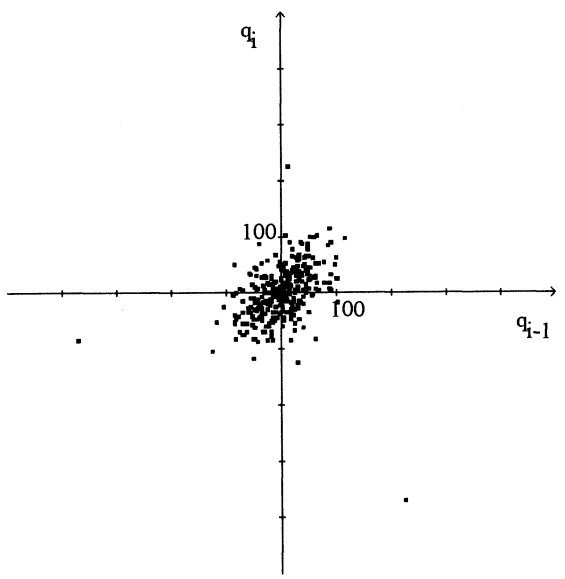


Fig. 7. Scatterplot $q_i=f(q_{i-1})$ of second difference of time error

9. REFERENCES

1. P.Rousseeuw and A.Leroy, Robust Regression and Outlier Detection, Wiley, New York, 1987.
2. P.J.Huber, Between Robustness and Diagnostic, in Directions in Robust Statistics and Diagnostics Part I, W.Stahel and S.Weisberg (eds.), The IMA Volume in Mathematics and its Applications, Springer-Verlag, New York 1991, pp. 121--130.
3. B.Sklar, Digital Communications Fundamentals and Applications, Prentice Hall, New Jersey, 1988.
4. S.R.Kim and A.Efron, Adaptive Robust Impulse Noise Filtering, IEEE Trans. on Signal Processing, vol. 43, No 8, August 1995, pp. 1855-1866.
5. J.Sawicki, Eliminating Impulsive Noise Using Adaptive Filtering Algorithm, XIXth National Conference Circuit Theory and Electronic Networks, Kraków - Krynica, Poland, October 23th - 26th, 1996, Proceedings, pp. II/417 - 422.

This work was performed in the frame of project TB-44-485/97-DS.

METHODS OF IMPROVING THE ESTIMATION OF LONG-TERM FREQUENCY VARIANCE

D. A. Howe

Time & Frequency Div., National Institute of Standards & Technology
Boulder, Colorado 80303

1. Abstract

I discuss a novel technique for periodically extending by reflection an ordered series of phase-difference measurements $\{x_{k'}\}$ to yield estimation of new frequency and time variances having improved long-term confidence and the same mean as the Allan variance. In addition, I describe a correction to a negative bias in the sample variance based on the actual number of observations in the measurements at long term.

Keywords: *two-sample frequency variance; time variance; Allan variance; TOTAL variance*

2. Background

This paper assumes some familiarity with the Allan variance and five common noise models (white phase modulation, WHPM; flicker of phase modulation, FLPM; white frequency modulation, WHFM; flicker of frequency modulation, FLFM; random walk of frequency modulation, RWFM) present in frequency standards, clocks, and synchronization systems [Refs. 3,21]. It is widely recognized that a trend (given by a slope in log coordinates) in the autocorrelation function and hence its Fourier transform $S_{\bar{y}}(f) \propto f^{\alpha}$ has a correspondence to a trend (given by μ) in the two-sample frequency variance [Refs. 3,10,21]. This power-law correspondence between α and μ means that, in general, procedures for better estimation of the power-law type in the f-domain have a parallel in the τ-domain [Ref. 2]. For those in the business of operating extremely precise clocks and oscillators, a principal task is characterizing the frequency stability of the devices relative to other devices and for comparison with the device's own history. Thus pertinent characterizations often refer to changes over relatively long segments of time.

One of the recommended ways to estimate the stability has been the two-sample frequency variance known as the Allan variance and corresponding statistic given by ("< >" denote infinite time average)

$$\sigma_{\bar{y}}^2(\tau) = \frac{1}{2}\langle(\bar{y}_{i+1} - \bar{y}_i)^2\rangle ; \hat{\sigma}_{\bar{y}}^2 = \frac{1}{2(M-1)} \sum_{k=1}^{M-1} (\bar{y}_{k+1} - \bar{y}_k)^2 \qquad (1)$$

where if $\{\bar{y}_{k'}\}$, k' = 1,2,3,...,N-1 are fractional frequency differences averaged over τ_0 derived from N consecutive phase differences $\{x_{k'}\}$("primed" indexes means τ_0-spacing); then $\{y_k\}$, k = 1,2,3,...,M are fractional frequency differences averaged over interval $m\tau_0 \equiv \tau$. Hence $\hat{\sigma}_{\bar{y}}^2$ is implicitly dependent on dimensionless quantity m, a scale parameter which for efficiency can be limited to rational powers of 2, i.e., 2^i=m, i=0,1,2,3,...; (see for example, [Ref. 1]).

A record of residual fractional frequency fluctuations implies that we have administered some form of trend removal (detrending). The removal of a trend such as drift is done electronically (as a voltage steering correction in a clock or oscillator synchronizing servo, for example), computationally (as a regression to an internal estimate of a continuous polynomial), or even mechanically (as a thermally-compensated cavity, for example). The response of the two-sample Allan variance eq (1) at long term is highly variable with forms of detrending and exhibits a negative bias [Refs. 4,8,13,17,23,24]. In cases where the last point is obviously false as judged by the rest of $\sigma_y(\tau)$ plot, it is thrown out. Worse, however, if it seems plausible and is judged somehow as "okay," it is retained and can be used to conform to some expected or suitable long-term behavior.

Realizing this, I introduce a new variance which has the same mean as the Allan variance eq (1) but re-expresses deviates in terms of an averaged combination of "in-phase" and "phase-shifted" sampling functions [Ref. 11] The new variance takes the traditional Allan variance and its single 2 sample function and combines it with an orthogonal 3 sample function proportioned so that its averaging time τ is taken out of the 2τ sampling-interval's middle (not just the first half and second half). The comparison is shown in fig. 1a and 1b. This increases the effective number of independent observations from one to two in the sampling interval thus increasing the number of degrees of freedom which I'll discuss in a moment. Even more crucial however is that sensitivity to removing drift from a segment of data in process (detrending) is eliminated in this variance. The new variance is given by

$$\sigma_{\bar{y}*}^2(\tau) = \frac{1}{4}\left[\left\langle\left[\frac{\bar{y}_{k+3}+\bar{y}_{k+2}}{2} - \frac{\bar{y}_{k+1}+\bar{y}_k}{2}\right]^2\right\rangle + \left\langle\left[\bar{y}_{k+3} - \frac{\bar{y}_{k+2}+\bar{y}_{k+1}}{2} + \bar{y}_k\right]^2\right\rangle\right], \qquad (2)$$

where the average frequency is taken over $\tau/2$ rather than τ as in eq (1). All possible τ/τ_0 = m time shifts (i.e., meaning deviates in the statistic are maximally overlapped) in addition to two shifts (separated by $\tau/2$

Contribution of the U.S. Government, not subject to copyright.

only as in eq (2)) result in smoothed estimates of the new variance at long term. The sampling functions are shown in fig. 1c. The smoothed estimate to eq (2), called TOTALVAR, uses a novel data manipulation which simplifies the procedure and which maximally uses data at hand.

To begin, an estimate to eq (2) has an equivalent description in signal processing as an "in-phase" and "phase-quadrature" discrete functional component separation. The variance component having the 3 sample function is computed by shifting the process's observation window by $\tau/2(\propto m/2)$, for the "phase-quadrature" variance and adding this to the Allan component or "in-phase" variance [Ref. 11]. This yields a combination of sample statistics given by

$$\sigma^2_{\bar{y}}(m\tau_o)=\frac{1}{4(M-1)}\left[\sum_{k=1}^{M-1}(\nabla\bar{y}_{k+1,m})^2+\sum_{k=1}^{M-1}(\nabla\bar{y}^*_{k+1,m})^2\right], \qquad (3)$$

where, in terms of phase data $\{x_{k'}\}$ spaced by τ_o, hence the "primed-k" index (t is the running time in seconds)

$$\bar{y}_{k,m}=\frac{x_{k'-m}(t_{k'}-m\tau_o)-x_{k'}(t_{k'})}{m\tau_o},$$

and

$$\bar{y}^*_{k,m}=\frac{x_{k'-m/2}(t_{k'}-\frac{m\tau_o}{2})-x_{k'+m/2}(t_{k'}+\frac{m\tau_o}{2})}{m\tau_o},$$

$x_1, x_2, \ldots x_N$ is wrapped such that $x_\varsigma = x_{\varsigma \mathrm{mod} N}$ for $\varsigma<1$ and $\varsigma>N$, i.e., $x_1=x_{N+1}$ which reindexes to x_N. It is important to point out that shifting the data and using a wrap (circularizing $\{x_{k'}\}$) is equivalent to changing the sampling function and merely simplifies the form of the sample variance eq (3) corresponding to eq (2) [Ref. 6]. There is no assumed extension of the original observed data; the wrap (or implied extension) is a computational procedure.

The treatment of changes in systematics (i.e., non-stationary second increments of $\{x_{k'}\}$) has been addressed [Refs. 12,13]. This has resulted in a wrap of series $\{x_{k'}\}$ which reflects about the last value of the series rather than simply wraps the data about an axis. This pictorially is shown in fig. 2 and is discussed in **4** below. The final procedure is as follows with notation change $x_{k'} = x(k')$ to distinguish a different (extended) set $\{x_o\}$:

For $x(1),\ldots,x(N)$, remove a slope and constant (endmatching procedure) to produce $x_0(1),\ldots, x_0(N)$, where $x_0(1) = x_0(N) = 0$. Adjoin $x_0(N+1),\ldots,x_0(2N-1)$, where $x_0(N+j) = -x_0(N-j)$, $j=1$ to $N-1$. For $\tau=m\tau_0$, TOTALVAR (with square-root as TOTALDEV) is given by

TOTAL FREQUENCY VARIANCE:

$$\sigma^2_{TOTAL}(\tau)_{\bar{y}}=\left(\frac{1}{\tau^2}\frac{1}{2(N-m-1)}\right)\times \sum_{n=1}^{N-m-1}[x_0(n)-2x_0(n+m)+x_0(n+2m)]^2. \qquad (4)$$

Notice that n+m stops just short of N; that avoids an identically zero second difference.

The evolution of a time error over τ (TIME VARIANCE) between clocks is defined as the frequency variance times (τ^2·constant) and is calculated using the same procedure as eq (4) with a slight modification which distinguishes levels of phase coherence, namely WHPM and FLPM [Ref. 21]. Determining phase coherence is important particularly in synchronous networks which need to characterize time coherence in the long term in addition to the more relaxed but less meaningful requirement of frequency coherence. Phase coherence, in a statistical sense, is inferred by the trend in the variability (convergence property) as a function of τ-averaged phase. Thus we construct TOTAL-TVAR (with square-root as TOTAL-TDEV) as

TOTAL TIME VARIANCE:

$$\sigma^2_{TOTAL}(\tau)_{\bar{x}}=\left(\frac{1}{3}\frac{1}{2(N-m-1)}\right)\times \sum_{n=1}^{N-m-1}[\bar{x}_0(n)-2\bar{x}_0(n+m)+\bar{x}_0(n+2m)]^2. \qquad (5)$$

A summary of methods of improved estimation applied to eqs (4) and (5) are presented next.

3. Endpoint Matching

Often we remove at least a regressed linear slope in an ordered set of time-difference measurements given by $\{x_{k'}\}$. In practice this removes an overall frequency difference but instead of removing a regressed linear slope, there is a case for removing a linear slope designed to match the endpoints of time series $\{x_{k'}\}$. This has other advantages in the context of the data extension procedure of eqs (4) and (5).

A series $\{x_{k'}\}$ represents the finite observation of ordered random variables. All observations are made through an observation window function which (unless otherwise noted) is rectangular in shape. Treating the series as a function over a finite interval, one commonly determines rate and drift using regression analysis of x(t) or its derivative y(t). Subsequent removal of a chosen model of trends can significantly alter an essential characteristic of the original series of measurements particularly at long-term for non-white noise types [Refs. 4,5]. By using a circular representation of a finite time series, one can arbitrarily time shift any chosen observation window or sampling function such as shown in fig.1 [Refs. 6,11,14]. In the moved-window case, the beginning

and endpoints of the series $\{x_{k'}\}$ are somewhat matched given by the trend removal, correlation, finite physical boundaries, and measurement-system bandwidth of the measured series. However, the turn-on/turn-off transient of the window, becomes an artifact which is not representative of the functional "roughness" or "smoothness" of the series and as such should be eliminated by a removal of linear slope which matches the endpoints prior to calculating a measure of such quantities as variance. If the incremental differences are stationary this methodology does not change the linear combinations of incremental differences contained in eqs (4) and (5).

4. Reflected Time Series

The wrap procedure implies an extension of the data by the original series which of course has identically distributed statistics, that is, the circularized series has the same variance as the original series. To reduce endpoint turn-off transients, a source of what is called in signal processing "leakage" [Refs. 15,18], we apply an endpoint match. We also desire smooth derivatives in the extension to properly estimate low-frequency noise components in whatever statistic we choose. This is particularly important in handling RWFM and drift. Frequency variances are not affected by reversing the direction of the series of measurements. Therefore, reflecting and wrapping the time series about the first and last points (both made to equal 0 as shown in 3 above) implies that we have made the most prudent assumptions and practical extensions of the time series in order to reduce transients, and hence reduce leakage. Figure 3 which shows series $\{x_{k'}\}$ reflected about the last point, represents simulated RWPM (or equivalently WHFM). We construct a new sequence of numbers which reflect about the last point (zero) however again I clarify that this construction is not a real extension; it is a convenience in calculating a maximally-overlapped estimate of the new 3-sample/2-sample variance suggested here.

5. Normalization of Sample Variance and Effective Degrees of Freedom (edf)

5.1 Illustration: We start with the sample mean of a set of discrete variables indexed by k′ as given by

$$\bar{x} = \frac{1}{M}\sum_{k'=1}^{M} x_{k'}, \qquad k'=1,2,3,\ldots M. \tag{6}$$

The mathematical form of the standard sample variance of the mean looks very much like the sample mean, an average, but in fact it is not an average, i.e., a sum of numbers divided by the total number of numbers as in eq (6). The sample variance is sometimes described as a sum of deviates divided by (or "scaled" by) the assumed number of degrees of freedom, often believed to be M, the total number of deviates. But the individual deviates are always some "difference" quantity derived from the grand mean of the set of random variables (recall the basic measure is "variance") and there can never be any more than M-1 differences that are "independent observations" (M-1 degrees of freedom). Calculating the variance implies that M (the number of samples) is at least two. Random variables may be differenced relative to a non-zero mean (or an assumed or actual zero mean), relative to a starting, ending, preceding, or following random variable, or for that matter any quantity derived from the set of M variables. Starting with independent random variables, any variance which properly preserves independence should have a scaling or normalization factor given by M-1. Division by this scaling factor yields a sample variance often called an "unbiased estimate," and the scaling factor is the number of independent observations.

Therefore, if $\bar{x} = 0$, the normalized variance (or variation) about a zero mean is given by the mean-square deviates as:

$$\sigma^2_{\bar{x}=0} = \frac{1}{M-1}\sum_{k'=1}^{M} (x_{k'})^2. \tag{7}$$

Recall that the actual number of observations is M in our quest to find some "true" mean which we don't know but which the finite grand or sample mean estimates. Therefore we can assume M independent observations (M degrees of freedom) to the extent that the sample mean estimates the true mean. For a white (WH) process, the sample mean whose variability decreases as M is the optimum estimate of the true mean, thus M degrees of freedom yields an unbiased estimate of its variance. For a random walk (RW) process, there is no true mean, nevertheless there is a sample (moving) mean, the bias on its variance is readily calculable and turns out to be small (an error by a worst case factor of 1.5 at $\tau = T/2$ for the two-sample frequency variance). If we can judge the noise type, hence correct for bias at long term, then we have essentially M degrees of freedom. Noise type can be determined from empirical data by its normalized autocorrelation function discussed next.

5.2 Effective Number of Independent Observations (edf): Barnes [Refs. 3,10,21] did extensive work on estimation of the two-sample frequency variance, and introduced bias functions and fractional degrees of freedom. Small-sample statistics in which $\tau \to T/2$ was not treated as a special case is the main subject at hand. In order to treat this case, I very briefly introduce concepts of self-similarity, or the autocorrelation function. Where $\bar{y}_t$ are frequency deviations averaged over τ, we can relate an autocorrelation to the two sample variance [Ref. 25]. We have in simplified form

$$\Sigma(\bar{y}_t - \bar{y}_{t+\tau})^2 = \Sigma(\bar{y}_t^2 + \bar{y}_{t+\tau}^2 - 2\bar{y}_t\bar{y}_{t+\tau}) = \tag{8}$$

$$\Sigma\bar{y}_t^2 + \Sigma(\bar{y}_{t+\tau}^2 - 2\bar{y}_t\bar{y}_{t+\tau}) = \Sigma\bar{y}_t^2 + (\Sigma\bar{y}_{t+\tau}^2)(1 - \frac{r_\tau}{2})$$

$$\textit{where } r_\tau = \frac{\Sigma\bar{y}_t\bar{y}_{t+\tau}}{\Sigma\bar{y}_{t+\tau}^2}.$$

The coefficient r_τ is a good approximation to the normalized sample autocovariance (autocorrelation) given by

$$r_L = \Sigma x_i x_{i+L} / \Sigma x_i^2 \qquad (9)$$

where the x_i are deviations from the mean of the series considered, and L denotes the lag between the values of the product. It has often been found adequate to assume that the correlation between more distant values arises solely from that between the directly neighboring values; if that correlation is R, the correlation coefficient r_L of values L time units apart becomes equal to R^L [Ref. 16].

The assumption for WHFM is that the normalized autocorrelation of average frequency deviations yields $r_\tau=0$ everywhere except at $\tau=0$ where $r_0=1$. A non-zero autocorrelation in a series "reddens" its spectrum of deviations $S_y(f)=\sum h_\alpha f^\alpha$, giving a greater share of the total variability to longer periods and a smaller share to shorter ones [Ref. 7]. Processes where $\alpha<0$ contain memory in the sense that correlation between long time intervals arises from values at the shortest interval denoted as τ_0 [Ref. 5]. As a consequence, the variability of sample averages of frequency deviations with memory (approaching RWFM) decreases more slowly with increasing sample size than does that of averages from a white series without memory [Ref.22]. Thus the number of independent observations is expressed as an *effective* number of degrees of freedom *(edf)* smaller than the actual number used in the summand of the sample variance [Ref. 20]. This effective number of independent observations (*edf*) is essentially the number of equivalent degrees of freedom based on the autocorrelation function in the samples themselves. We have

$$(edf) = (M - \epsilon_\gamma(M)) \qquad (10)$$

where ϵ depends on M and autocorrelation properties, subscript γ designates the type of sample frequency variance, and *(edf)* can be fractional, i.e., non-integer. To be concise, I limit the discussion to M=2 which turns out to be the most interesting case (longest term) representing $\tau=T/2$.

Properties of the distribution of the usual Allan-deviation estimate have been studied using fractional degrees of freedom on the confidence interval [Refs. 9,10]. Similar studies can also be applied to the actual time series, since there is a correspondence between these distribution properties and the underlying noise process of the data. Keep in mind that the two-sample frequency variance is a time-averaged, standard-sample variance against a previously measured mean. Its square root defines a relative uncertainty on this τ-averaged mean. Thus the sample Allan deviation is an uncertainty of a sample mean frequency decomposed by τ averaging times. The underlying noise type defines the trend in this uncertainty. Moreover the number of degrees of freedom in the underlying noise type has a correspondence to the degrees of freedom in the uncertainty. In this form for the Allan variance, $\epsilon_{AVAR}(2)$ can only range from 1 to 1.5 corresponding to WHFM to RWFM respectively since RWFM as an integral of WHFM. This is because the means of segments of data will converge to the grand mean only half as fast for RWFM vs. WHFM.

5.3 Actual Number of Observations (h): The sample Allan variance is useful as a power-law (octave-band) spectral estimator but is time-shift (phase) sensitive and depends on where we start the calculation with respect to data in process. For large data sets and small scale values of m, the odd and even values of index k overlap and average together in eq (1) for a fairly accurate estimation of a broadband spectral distribution of variance of first differences of average frequency. The division by 2(M-1) which corresponds to 2(N-m-1) in eqs (4) and (5) is arguably due to overlapping two sets of deviates and has constant "2" only for WHPM, FLPM, and WHFM but ought to approach M-1 (or N-m-1) for an accurate estimate in the statistic as $m\tau_0 \to T/2$, since the first and last deviates do not overlap. However, at the largest scale, the estimate is negatively biased for non-WHFM (i.e., FLFM and RWFM) because there is only 1 (not 2) "two-sample" sample. This reason causes an estimation error at long term in virtually all cases in which the noise is no longer WHFM but is FLFM or RWFM even though the estimation is supposedly unbiased even at these large τ values (or equivalently large values of m).

For the statistic TOTALVAR given in eq (4), there is considerable overlap in the summation. It is widely known that this effectively smooths the estimate [Refs. 10-12]. But the number h of actual observations (hence the scaling factor) has not changed and remains a straightforward calculation giving

$$h = \frac{[2(\log_2(\frac{N-1}{m})) - 1]}{(\log_2 \frac{N-1}{m})}(N-m-1). \qquad (11)$$

We find in simulation studies that eq (11) should be applied for RWFM, and not at all for WHFM [Refs. 12,19]. This suggests that the division by 2(M-1) ought to approach M-1 because of the effects of correlation and not necessarily because of a connection with the actual number of observations as discussed above. Nevertheless, if h is the actual number of observations, note that for large M values, *(edf)* ≐ h, but that for small M, *(edf)* becomes increasingly sensitive to the value of ϵ. A table of values of M, and corresponding ϵ_{AVAR} for various noise types is planned for future work. For the purpose here, we find that in the presence of RWFM, substituting 2(N-m-1) in eqs (4) and (5) by h in eq (11) above removes a negative bias in estimates of sample frequency variance at long averaging times.

6. Conclusion

I have introduced new statistics of frequency and time variances which yield improved estimation of both frequency stability and noise type, particularly at long averaging times. An initial procedure involves regressing to global basis functions such as orthogonal

polynomials. This procedure of detrending assumes the residuals are white phase noise. Unfortunately frequency variations usually exhibit systematic effects that eventually change and are interpreted as arising from a divergent noise type hence non-stationary sample frequency variance. Hence in the case of the two-sample Allan variance the effective number of observations (equivalent degrees of freedom) is slightly reduced by factor ϵ which ranges from 1 in the presence of WHFM to 1.5 for RWFM for the largest τ-value at T/2.

7. References

[1] Allan, D.W. and Barnes, J.A., A modified "Allan Variance" with increased oscillator characterization ability, Proc. 35th Freq. Control Symp., 470-476 (1981).

[2] Allan, D.W., Weiss, M.A., and Jespersen, J.L., A frequency-domain view of time-domain characterization of clocks and time and frequency distribution systems, Proc. 45th Freq. Control Symp., 667-678 (1991).

[3] Barnes, J.A., Chi, A.R., Cutler, L.S., Healy, D.J., Leeson, D.B., McGunigal, T.E., Mullen Jr., J.A., Smith, W.L., Sydnor, R.L., Vessot, R.F.C., and Winkler, G.M.R., Characterization of frequency stability, IEEE Trans. Instrum. Meas., **IM-20**, 105-120 (1971).

[4] Barnes, J.A., The measurement of linear frequency drift in oscillators, Proc. 15th Ann. PTTI Meeting, (1983). Article is also contained in NIST Technical Note 1337 (1990).

[5] Beran, J., Statistical methods for data with long-range dependence, Statistical Science, **7**, 404-427 (1992).

[6] Bloomfield, P., Fourier analysis of time series: an introduction, Wiley, New York, 1976. Pages 55-60 are a useful discussion of the "circular definition" of a noise (or roughness) coefficient, namely the variance, using "wrapped" data.

[7] Gilman, D.L., Fuglister, F.J., and Mitchell, J.M., Jr., On the Power Spectrum of "Red Noise", J. Atm. Sci. **20**, 182-184 (1963).

[8] Greenhall, C.A., A structure function representation theorem with applications to frequency stability estimation, IEEE Trans. Instrum. Meas., **IM-32**, pp. 364-370 (1983).

[9] Greenhall, C.A., Recipes for degrees of freedom of frequency stability estimators, IEEE Trans. Instrum. Meas., **40**, pp. 994-999 (1991).

[10] Howe, D.A., Allan, D.W., and Barnes, J.A., Properties of signal sources and measurement methods, Proc. 35th Ann. Symp. Freq. Control, A25-A35 (1981). Article is also contained in Natl. Inst. Stand. Technol. TN-1337 (1990).

[11] Howe, D.A., An extension of the Allan variance with increased confidence at long term, Proc. 1995 IEEE Int. Freq. Cont. Symp., pp. 321-329.

[12] Howe, D.A. and Lainson, K.J., Simulation study using a new type of sample variance, Proc. 1995 PTTI Meeting, November, pp. 279-290 (1995).

[13] Howe, D.A. and Lainson, K.J., Effect of drift on TOTALDEV, Proc. 1996 IEEE Int. Freq. Control Symp., pp. 883-889.

[14] Howe, D.A., Circular representation of infinitely extended sequences, Proc. 1995 IEEE Int. Freq. Cont. Symp., pp. 337-345 (1995).

[15] Howe, D.A. and Percival, D.B., Wavelet Variance, Allan Variance, and Leakage, IEEE Trans. Instrum. Meas., IM-44, 94-97 (1995). Also, Howe, D.A. and Percival, D.B., Wavelet Analysis for Synchronization and Timekeeping, Proc. IEEE Int. Freq. Control Symp., pp. 791-794 (1994).

[16] Kendall, M.G. and Stuart, A., *The Advanced Theory of Statistics*, **3**, Griffin, London (1966).

[17] Legachev, V.A. and Pashev, G.P., Estimation of linear frequency drift coefficient of frequency standards, Proc. 1996 IEEE Int. Freq. Control Symp., pp.

[18] Percival, D.B. and Walden, A., Spectral analysis for physical applications, Cambridge University Press, New York, 1992.

[19] Riley, W.J., Addendum to a test suite for the calculation of time domain stability, Proc. 1996 IEEE Int. Freq. Control Symp., pp. 890-893.

[20] Radok, U., Assessing climate persistence from climatic noise, *Climatic Change* (1992).

[21] Sullivan, D.B., Allan, D.W., Howe, D.A., and Walls, F.L., Characterization of Clocks and Oscillators, NIST Technical Note 1337 (1990).

[22] Vernot, F., and Vincent, M., Estimation of the uncertainty of a mean frequency measurement, Proc. 11th European Freq. and Time Forum, 1997 (to be published these proceedings).

[23] Weiss, M.A., Allan, D.W., Howe, D.A., Confidence on the second difference estimation of frequency drift, Proc. IEEE 46th Int. Freq. Control Symp., pp. 300-305 (1992).

[24] Weiss, M.A. and Hackman, C., Confidence on the three-point estimator of frequency drift, Proc. 24th Ann. PTTI Meeting, pp. 451-460 (1992).

[25] Yoshimura, K., Characterization of frequency stability: Uncertainty due to the autocorrelation of the frequency fluctuations, IEEE Trans. Instrum. Meas., **IM-27**, 1-7 (1978).

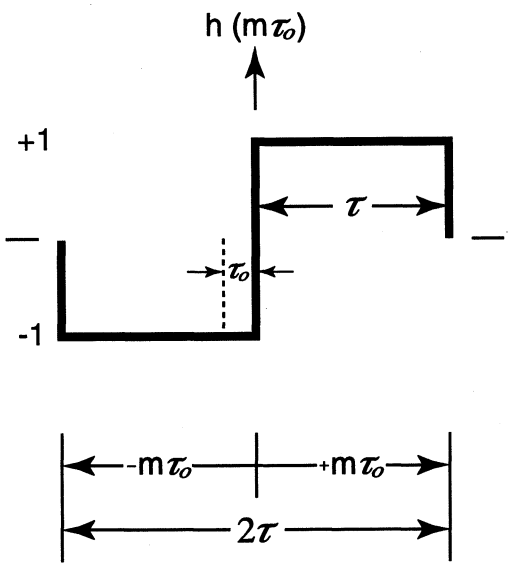


Fig. 1a: The Allan variance (2 sample) function has one degree of freedom per 2τ interval and is sensitive to the method of drift removal (detrending) over large τ values.

FIG. 1a

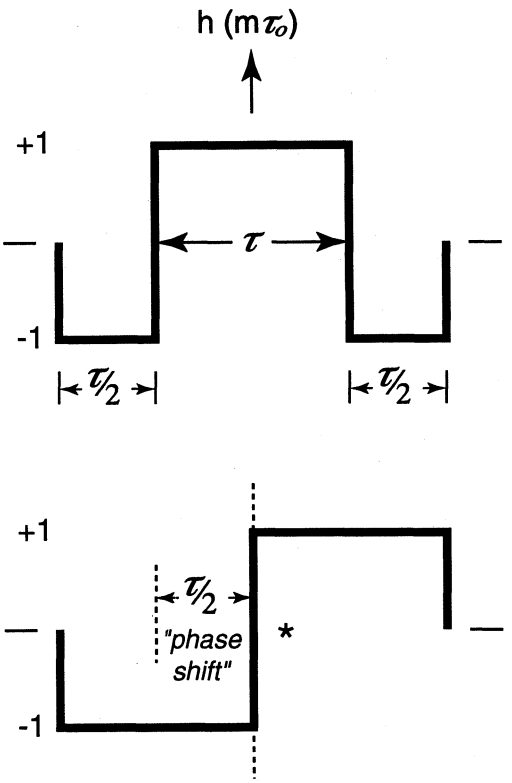


Fig. 1b: The important advantages to the new 3-sample/2-sample variance is that it retains all of the desirable convergence properties of the Allan variance, has the same mean, has at least one (depends on noise type) additional degree of freedom for any segment of the data, and is not sensitive to removal of drift [Ref. 11]. The additional degrees of freedom are especially crucial at long averaging times where the number of degrees of freedom is limited to about only 1 using even the max-overlap sample Allan variance.

FIG. 1b

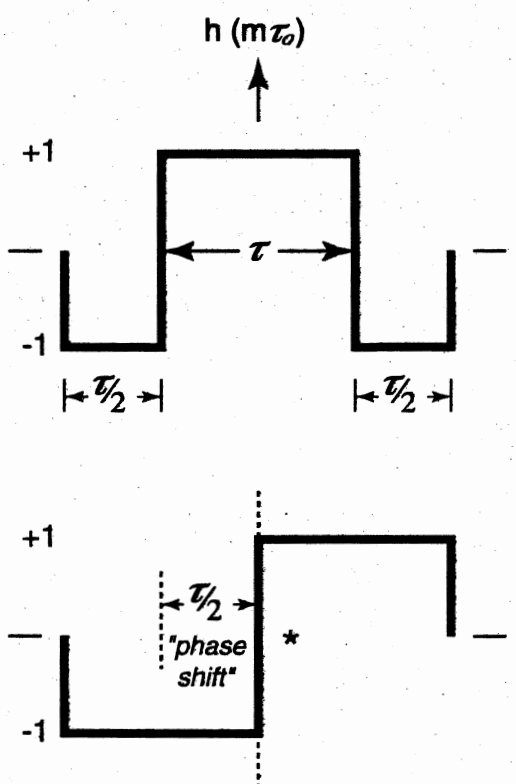


Fig. 1c: Maximum-overlap sampling smooths the estimate of the 3-sample/2-sample frequency variance over 2τ by shifting the τ-sample by τ_o and averaging all possible frequency variances. This averages all possible phase shifts of τ-sample relative to the 2τ interval [Refs. 11-14,19].

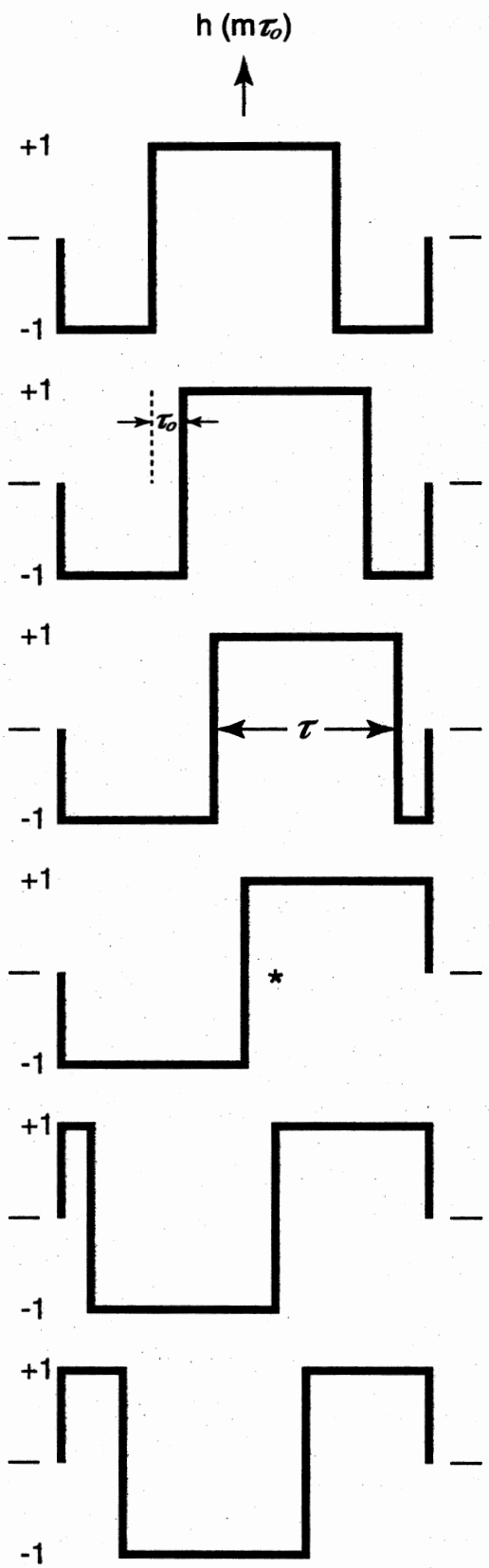


FIG. 1c

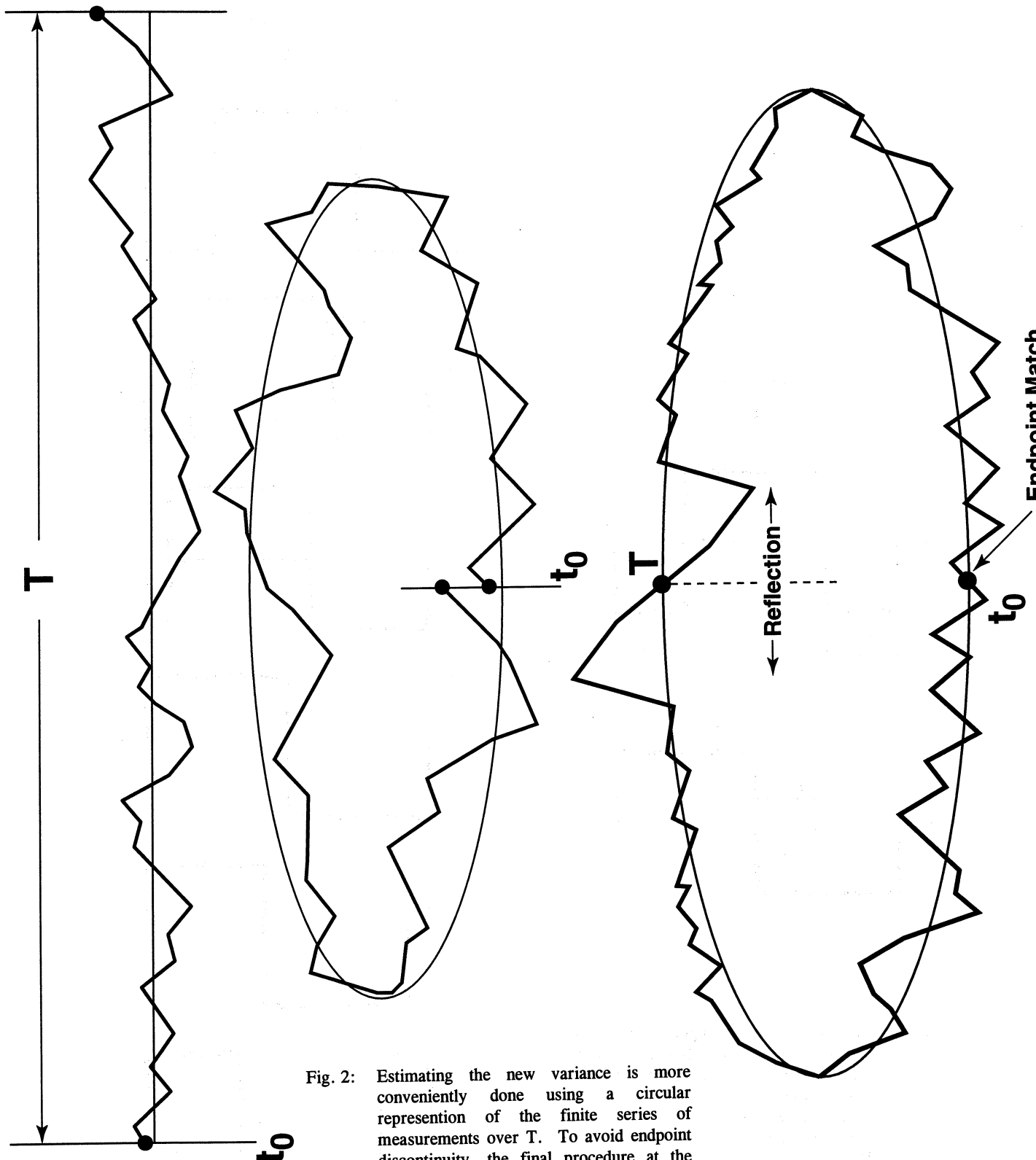


Fig. 2: Estimating the new variance is more conveniently done using a circular represention of the finite series of measurements over T. To avoid endpoint discontinuity, the final procedure at the bottom involves a circular wrap and reflection of the original data at the time origin t_o and end (T). This procedure makes the computation simple for statistics dubbed TOTALVAR eq (4) and TOTAL-TVAR eq (5) in the text.

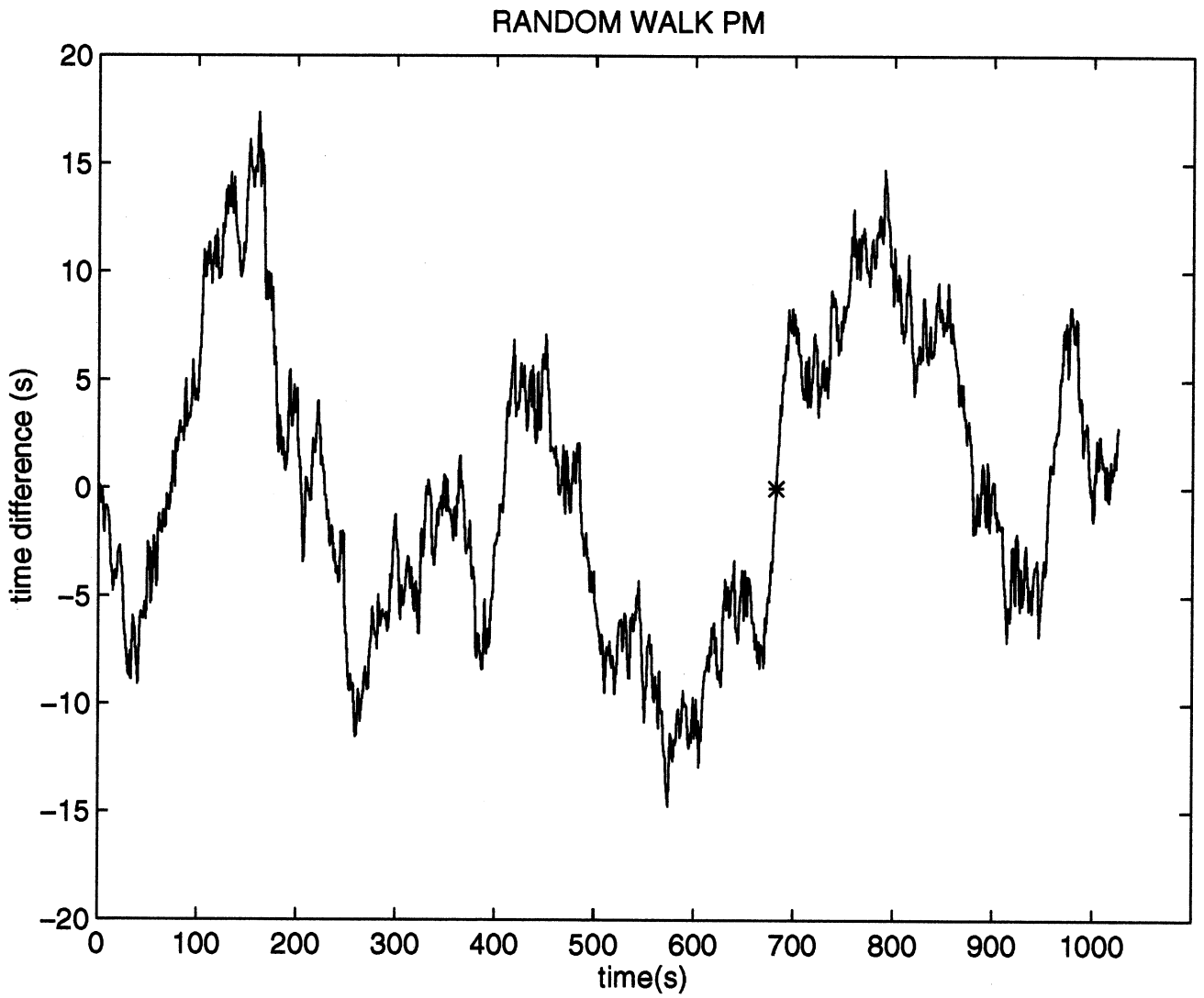


FIG. 3

Fig. 3: This figure shows typical data that is extended (adjoined at *) using endpoint matching and reflection about the last point.

OPTICAL FREQUENCY STANDARDS

Chairman: A. Clairon

Optical Frequency Standard Based on Trapped Ca Atoms

J. Helmcke, F. Riehle, H. Schnatz, B. Lipphardt, P. Kersten, G. Zinner, T. Trebst, T. Kurosu

Physikalisch-Technische Bundesanstalt
Bundesallee 100
D-38116 Braunschweig, Germany
Phone: +49-531 592 4300, Fax: +49-531 592 4305, e-mail: juergen.helmcke@ptb.de

Optical frequency standards allow at present the most accurate realization of the SI unit of length, the metre. Furthermore, they may lead to the development of optical clocks of unprecedented low uncertainties. We have developed an optical frequency standard stabilized to the $^3P_1$ - $^1S_0$ transition in ^{40}Ca at λ = 657.46 nm. In our standard, we have excited Ca atoms stored in a magnetooptical trap with three consecutive pulses of a standing wave leading to atomic interferences. The frequency of a prestabilized laser was then stabilized to the interference structure. Using separated field excitation, we have investigated the phase and frequency shifting influence of disturbing fields e.g. ac and dc-Stark effect, Aharonov-Casher effect, Sagnac effect, laser fields, and cold collisions to determine an uncertainty budget of the optical frequency standard. It can be expected that a relative uncertainty of a few parts in 10^{15} can be reached with this standard. Recent consecutive frequency stabilizations to two independent magnetooptical traps showed no difference between the two frequency values within the relative uncertainty of the experiment of 10^{-13}.

In order to relate the frequency of the Ca optical standard to the one of the primary standard of time and frequency, the Cs atomic clock, a phase-coherent frequency measurement chain was set up to bridge the frequency gap between the microwave and the optical standard. Four different measurements performed so far led to the combined result of (455 986 240 494.07 ± .35) kHz. As a result of this measurement, the Ca stabilized laser now represents the frequency standard in the visible spectrum with the lowest uncertainty for the realization of the metre. Its frequency will be suitable as a reference in this spectral range for precision wavelength measurements as well as for the determination of fundamental constants.

WIDEBAND OPTICAL PARAMETRIC OSCILLATORS FOR OPTICAL FREQUENCY MEASUREMENT

Takeshi Ikegami and Sergey Slyusarev

National Research Laboratory of Metrology,
1-1-4, Umezono, Tsukuba-Shi, Ibaraki-Ken 305, Japan
Tel: +81 298 54 4035, Fax: +81 298 50 1456
E-mail: ikegami@nrlm.go.jp

ABSTRACT

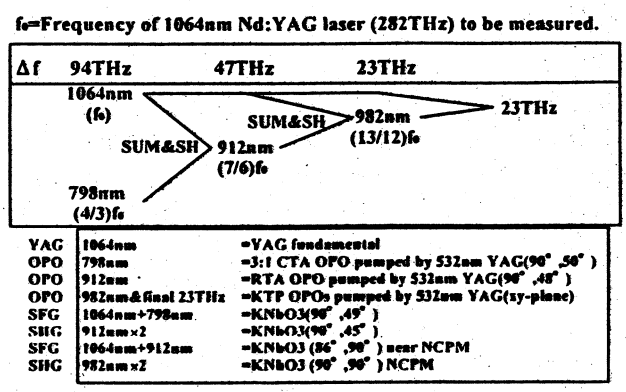
We developed an injection locked Nd:YAG laser using a monolithic nonplanar Nd:YAG ring oscillator as a master laser. The injection locked, unidirectional output of 9.5W was obtained. By the external doubling with an LiB_3O_5 crystal, we got 2.5W of green (532nm) light. Using this green light as a pump source of cw doubly resonant optical parametric oscillators (DROs), we got oscillation with three types of nonlinear crystals: (1)With a $KTiOPO_4$ (KTP) with cutting angle of (θ,ϕ)=(90°,30°), oscillation frequency was (1050nm(signal),1078nm(idler)) and the pump threshold was 30mW, (2)With KTP (90°,40°), they were (1001nm(s),1135nm(i)), 612mW, (3)With $RbTiOAsO_4$ (90°,48°), they were (904nm(s),1284nm(i)), 500mW. The efficiency was about 20%. To get higher stability, a monolithic KTP DRO was demonstrated. The threshold was 7mW and the efficiency was 16%. It was very stable and operated over 3 hours without any mode-hopping under free running condition. It was tunable by 2nm with temperature. The beat frequency between the signal and the idler could be successfully phase locked to a signal from a synthesizer by applying an E-field across the crystal.

Keyword: cw optical parametric oscillator, monolithic optical parametric oscillator, frequency chain

1. INTRODUCTION

Recently, the new types of frequency chains were proposed based on the development of the new solid state laser technology and the progress of nonlinear optical crystals (Refs.1-3). We also proposed a new frequency chain shown in Fig.1, considering the following situations (Ref.4). (1)Appearance of a commercial, reliable, very narrow linewidth (10kHz) Nd:YAG laser, (2)Availability of a high power (over 10W) Nd:YAG laser (Ref.5), (3) Appearance of transparant, non-hygroscopic, easy to treat, reliable, nonlinear optical crystals with large nonlinear constant, such as $KTiOPO_4$ (KTP) and its isomorphic crystals: ($RbTiOAsO_4$ (RTA), $CsTiOAsO_4$ (CTA), $RbTiOPO_4$ (RTP), $KTiOAsO_4$ (KTA)), $LiNbO_3$ (LN) and its periodically poled one (PPLN), and $KNbO_3$ (KN), (4)Possibility of practical use of cw doubly resonant optical parametric oscillators (DROs), which can produce any wavelength with the same principle (Ref.6), (5) Appearance of optical frequency comb generators (OFCGs) which produce wide span of the sidebands over 5THz with a modulation frequency of several GHz (Ref.7).

(a)



(b)

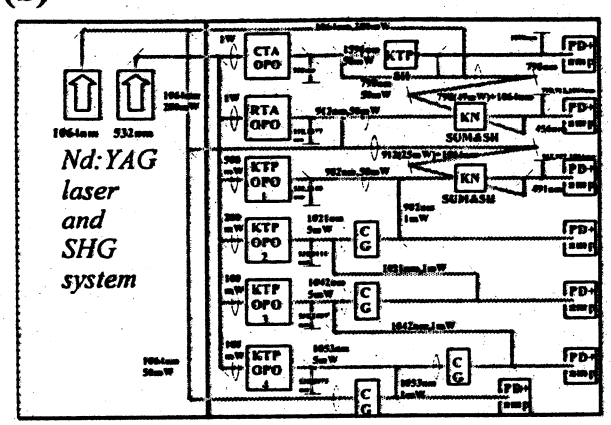


Fig.1 (a)Proposed NRLM frequency chain. SUM&SH: Frequency divider (Ref.1), (θ,ϕ): Cutting angle of nonlinear crystals. The wavelength 798nm is produced by 3:1 divider DRO and the final 23THz is measuremd by the parallel connection of KTP DROs and optical frequency comb generators (Ref.2).
(b)Actual image of the NRLM chain on an optical bench. CG: Optical frequency comb generator, PD: Photodiode. For the other parts, refer to the text.

We already succeeded in the oscillation of a nearly degenerate cw-DRO with the threshold of 10mW, the efficiency of 20%, and the tuning range of 20nm. The signle mode pair operation of signal and idler without any mode hopping over 3 hours was obtained (Ref.8). The phase locking of the beat signal between the signal and the idler was also obtained (Ref.9).

On the other hand, we developed an OFCG which had 5.5THz span at 1064nm with the modulation frequency of 6.2GHz (Ref.4).

For the realization of our frequeny chain, we need wide span cw-DROs which can operate between 800nm and 1600nm. To get this span, we developed a high power green (532nm) pump source described in section 2. Using this green light, we pumped wide-band cw-DROs (section 3). In addition, to get more stable and low threshold oscillation, we developed a monolithic DROs (section 4).

2. HIGH POWER GREEN LIGHT SOURCE

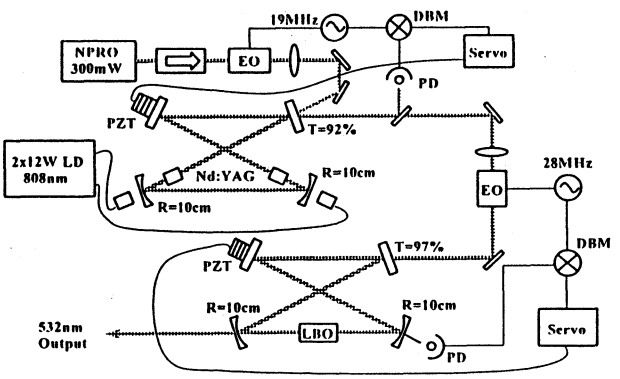


Fig.2 Setup of a high power green light source to pump cw-DROs.

To realize our frequency chain, we need a green (532nm) pump source satisfying the following requirements: (1) Over 4W of green (532nm) power, (2) Single logitudinal and transverse mode oscillation with the spectral linewidth smaller than 10kHz, (3) Stable single mode operation without any mode hopping over hours and days. Because we already have a commercial laser which satisfies the above spectral characteristics (a monolithic nonplanar Nd:YAG ring oscillator (NPRO), Lightwave Electronics Inc.), we decided to adopt an end pumped ring cavity high power Nd:YAG laser, which was injection locked to the NPRO (Ref.5). To get the green power from this IR, the second external ring enhancement cavity was adopted as shown in Fig.2.

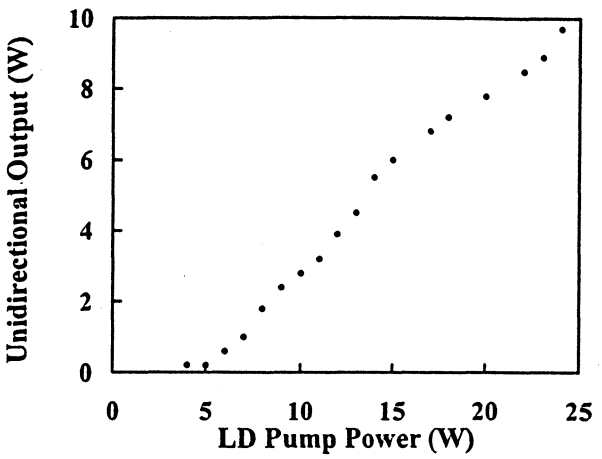


Fig.3 Injection locked, unidirectional output power vs LD pump power.

For the IR part, we adopted similar physical parameters to (Ref.5). The main cavity has a bow tie ring structure including a pair of Nd:YAG rods (3mm in diameter and 5mm long), which are end-pumped by a pair of fiber coupled laser diodes (808nm,12W each). The cavity was composed of 2 flat mirrors and 2 concave mirrors with a curvature of 10cm. Because the laser oscillated bi-directionally with multi longitudinal modes, it was injection locked to the NPRO with output power of 700mW to improve the spectral characteristics. The transmittion of the input=output mirror for 1064nm was 92%. After the injection locking, we obtained an uni-directional, single longitudinal mode output of 9.5W as shown in Fig.3. To keep the injection locking, the cavity length was locked to the master laser frequency with a Pound Drever method with a modulation frequency of 19MHz. For the cavity length control, two PZTs (slow and fast) which were attached to each mirror were used. The locking was kept over 1 hours as shown in Fig.4(a).

For the SHG part, we adopted an external enhancement cavity with the similar bow tie ring structure. The parameters for the cavity was similar to the main cavity except that the transmission of the input coupler was 97%. We used a LiB_3O_5 (LBO) crystal in the angle tuning configulation. We obtained a maximum SH power of 2.5W at 532nm. To keep the resonance, the Pound Drever method was adopted again. The modulation frequency of 28MHz was selected to avoid the interference between the main cavity modulation. The total system including the master laser, the main laser cavity and the SHG cavity operated stably over 2 hours without any mode hopping as shown in Fig.4(b).

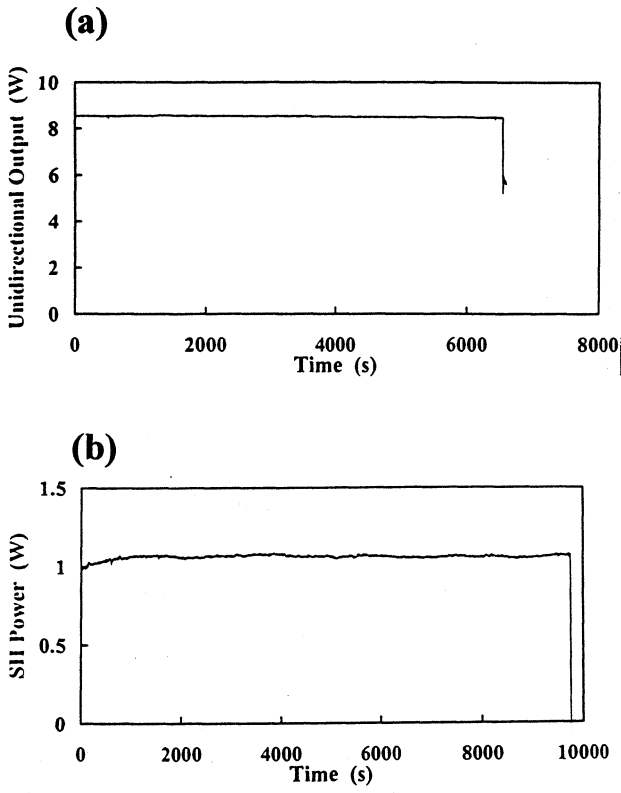


Fig.4 (a)Unidirectional IR (1064nm) output from the slave laser, (b)Green (532nm) SH output from the 2nd cavity.

We are trying to improve the stability of the system now. Also, we are comparing several nonlinear crystals including an angle tuning, critical phase matching LBO, a temperature tuning NCPM LBO, KTP, Na doped KTP to get higher efficiency.

3. WIDEBAND CW OPTICAL PARAMETRIC OSCILLATORS

Next, by using the developed high power green light source, we demonstrated wide band cw DROs. We can expect the low threshold oscillation in the full range between 750nm and 1900nm using three KTP isomorphic crystals, KTP, RTA, CTA because the walkoff angle is small in the xy plane phase matching ($< 0.5°$), as shown in Fig.5.

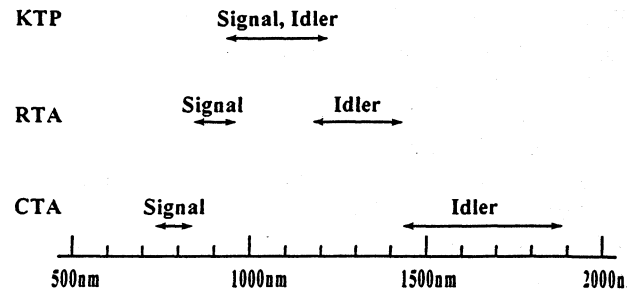


Fig.5 Wavelength range obtainable by 532nm pumped DROs using KTP, RTA, and CTA crystals in the xy plane phase matching.

The DRO setup is shown in Fig.6. Three types of nonlinear crystals with 8mm long, 3mmx3mm cross section were used with the type-II phase macthing configulation. Both sides of the crystal were AR coated for the pump wavelength (532nm), the signal wavelength, and the idler wavelength. The typical reflection for the pump, the signal and the idler were 5%, 0.15%, 0.15%, respectively. The input mirror had a 25mm radius of curvature and was coated for maximum reflection (typically 99.9%) for signal and idler. The transmission of the input mirror for 532nm was 96%. The output coupler had a 25mm radius of curvature and its transmission was 0.4% for the signal and idler. It was coated for maximum reflection for 532nm light. A PZT was attached to the output mirror for cavity tuning. The typical distance between the mirrors was 49mm.

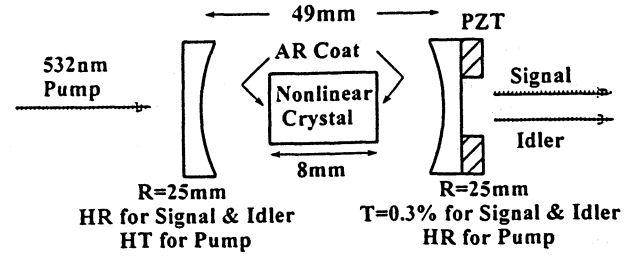


Fig.6 Experimental setup of wideband DRO.

We first tried KTP crystal with the cutting angle of $(\theta, \phi) = (90°, 30°)$. The oscillation wavelength was 1050nm and 1078nm as shown in Fig.7(a). The

pump threshold was 30mW and the efficiency was 30%.

Next we tried KTP with the cutting angle of (90°, 40°). The oscillation wavelength was 1001nm and 1135nm as shown in Fig.7(b). The threshold was 612mW.

Thirdy, we tried RTA crystal with the cutting angle of (90°, 48°). The oscillation wavelength was 904nm and 1284nm as shown in Fig.7(c) and the pump threshold was 500mW.

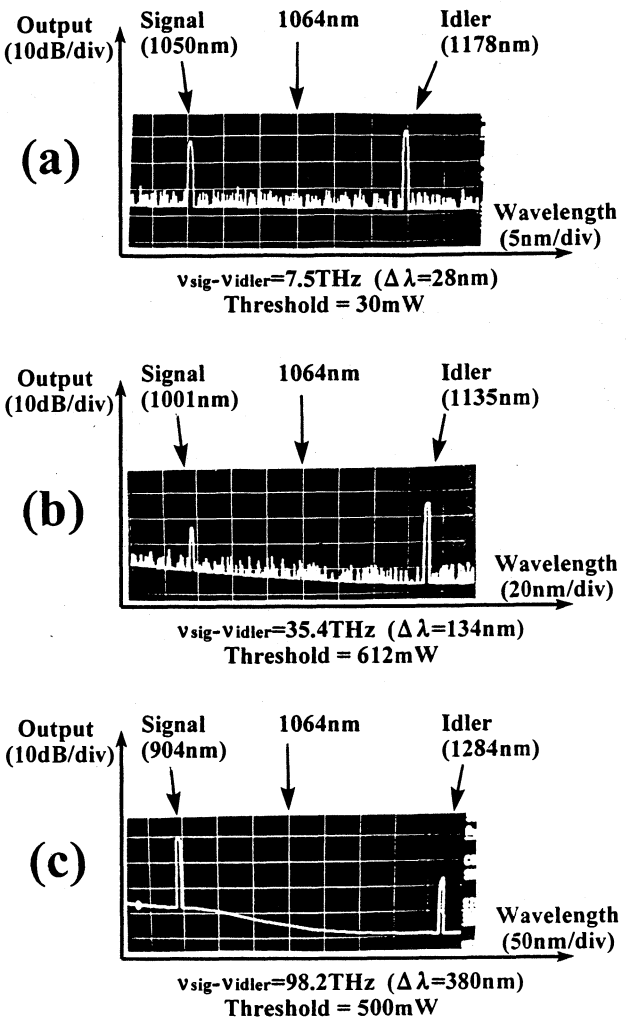


Fig.7 Output spectrom from the DROs (a) KTP (90°,30°), (b) KTP (90°,40°), (c) RTA (90°,48°).

4. MONOLITHIC CW OPTICAL PARAMETRIC OSCILLATOR

Owing to the adoption of the type-II phase matching, our DRO operated in signle longitudinal and transverse mode pair of signal and idler (Ref.10). However, due to the same reason, the high stability is requiered to the cavity length to maintain the simultaneous resonance for the signal and the idler. By the adoption of the semimonoithic cavity which does not use the adjustable element such as mirror mounts including springs, we could obtain the enough mechanical stability for long term operation (Ref.8). However, the allignment became much diffucult to get the oscillation in this case.

Second, for the low thredhold operation, very severe specification is required to the coating. Very high transmission, triple wavelength anti reflection (AR) coating to the crystal, and very high reflectivity, triple wavelength mirror coating to the mirrors are necessary. The present DRO composed of discrete elements includes 4 relevant faces (2 mirror surfaces and 2 crystal suraces). If we adopt a monoithic structure, in which the two faces of the crystal is polished spherically, the loss will be significantly reduced. In addition, we can expect much higher mechanical stability. However, if we adopt the monlithic structure, the tunability of the oscillation frequency may be limited. We will discuss this point later.

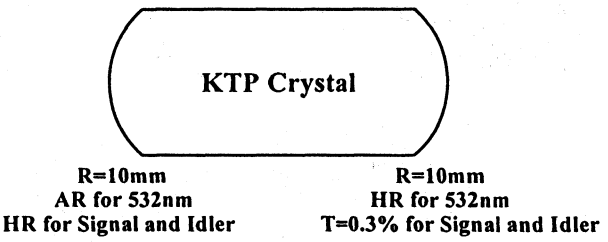


Fig.8 Structure of monolithic KTP DRO.

We adopted the following monolithic degenerate KTP DRO as shown in Fig.8. The crystal size was 3x3x8mm. The two faces of the crystal was polished spherically with the radius of curvature of 10mm. The input surface was AR coated for 532nm and high refrection (HR) coated for 1064nm. The output face has a transmission of T=0.3% for 1064nm and HR coated for green. We obtained 7mW of threshold and 16% of efficiency. It operated in single longitudinal mode pair of signal and idler over 3 hours without any mode hopping under free running condition as shown in Fig.9. For the tunability we anticipated, we could confirm that it is tunable by 2nm by changing temperature from 46°C to 20°C as shown in Fig.10. The tunability was 0.1nm/deg. By tunig temperature, we could operate it close to the degeneracy and we could observe the beat signal between the signal and the idler. By detecting the beat signal with a high speed photodetector, mixing it with a signal from a synthesizer, and by feedbacking to the E-field applied to the crystal, we could obtain the phase locking of the beat signal as shown in Fig.11. The beat frequency was 12GHz.

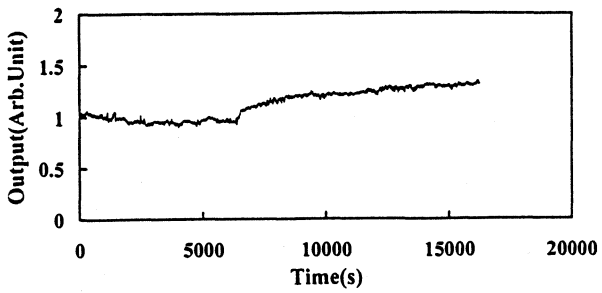


Fig.9 Output intensity from the monolithic DRO.

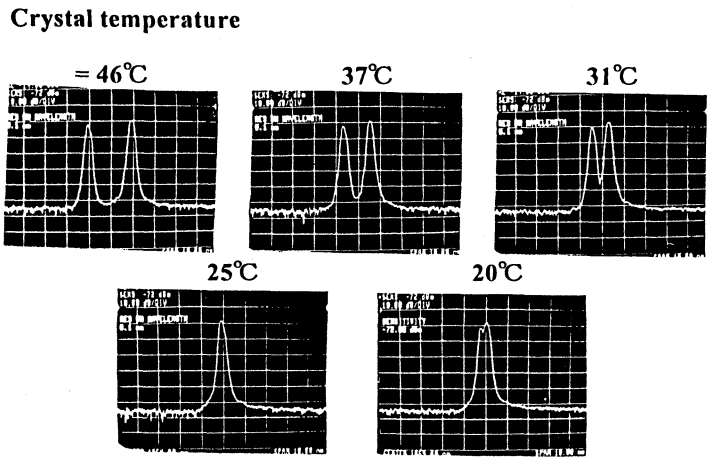


Fig.10 Dependence of the output spectrum from the DRO on temperature.

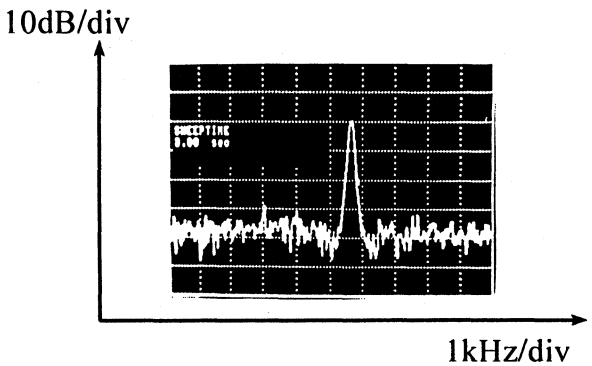


Fig.11 Phase locked beat spectrum of monolithic DRO between signal and idler.

6. SUMMARY

1. We developed a 2.5W, stable, narrow linewidth, green light source to pump several OPOs at once.

2. With this green light source as a pump source, we demonstrated wideband cw-DRO between 905nm and 1284nm using the isomorphics crystals of the KTP.

3. To improve the stability and to the threshold, the monolithic DRO was demonstareted. It was stable, and tunable with temparature (slow) and E-field (fast).

For the construction of our frequency chain, following things are remained to be done.

1. Increase of the green power using higher efficiency crystals such as noncritical, temperature phase matching LBO, KTP, Na doped noncritical phase matching KTP.

2. Demonstration of 3:1 divider (532nm pumped 798nm signal 1596nm idler cw-DRO). This is considered to be possible with the (90°, 50°) CTA crystal or the PPLN crystal.

3. Demonstration of the parallel connection of cw DROs using the OFCGs.

4. Construction of the system using SUM&SH dividers with KN crystals.

REFERENCES

(1) H.R.Telle et al, Opt.Lett. **15**, 532 (1990).
(2) N.C.Wong, Opt.Lett. **17**, 1156 (1992).
(3) K.Nakagawa et al, Appl.Phys.B **57**, 425 (1993).
(4) T.Ikegami et al, Proc.5th Symposium on Frequency Standards and Metrology, ed.J.C.Bergquist, World Scientific, 1996, p.333.
(5) S.T.Yang et al, Opt.Lett. **21**, 1676 (1996).
(6) J.Opt.Soc.Am.B **10**,No.9 (1993), the special issue on OPO and OPA.
(7) M.Kourogi et al, IEEE J. Quantum Electron. **29** (1993) 2693.
(8) T.Ikegami et al, Jpn.J.Appl.Phys. **35**, 2690 (1996).
(9) S.Slyusarev et al, Jpn.J.Appl.Phys. **35**, 3459 (1996).
(10) D.Lee et al, J.Opt.Soc.B **10**, 1659 (1993).

A FAST EXTERNAL LASER FREQUENCY STABILIZER

H. R. Telle

Physikalisch-Technische Bundesanstalt, D-38023 Braunschweig, Germany

1. ABSTRACT

A novel external frequency stabilization scheme is presented which is applicable to almost any single-mode cw laser. It consists of a self-oscillating opto-electronic loop whose oscillation frequency is determined by the frequency detuning between the optical input signal and the resonance frequency of a high-Q reference cavity. The method has been used for frequency stabilization of a laser diode (P=50 mW, $\lambda \cong 850$ nm) with respect to the resonance of a stable Fabry-Perot interferometer

Keywords: FM noise reduction, electro-optic modulation, laser diodes

2. PRINCIPLE

Light sources with narrow emission line width. i. e. low FM noise level, are required for most applications in spectroscopy, interferometry, optical frequency synthesis and so on. Although the quantum-limited emission line width is sufficiently small in most cases, fast servo loops are needed to reduce technical noise contributions from acoustic perturbations of the laser resonator, plasma oscillations, thermal 1/f-noise and so on. However, locking the emission frequency to a stable reference requires a fast frequency control input which is rarely available. To overcome this problem, external frequency stabilizers have been proposed and realized [1],[2] which employ acousto-optical modulators as variable frequency shifters. However, owing to complicated error signal processing loop delay times are quite long, yielding slow servo response: FM-spectroscopy is used to generate an error signal from a narrow resonance (optical resonator or atomic absorption line). This error signal in rf domain is converted to a baseband signal with the help of a mixer. Then, the baseband signal is amplified and employed as the input signal of an voltage controlled oscillator. Its rf output signal is used to drive the AOM. Finally, after an acoustical delay of the order of µs the control signal reaches the optical beam. Due to this large number of components, the overall loop delay time limits the control bandwidth to the order of MHz.

This paper describes a novel scheme which entirely operates in rf domain and thus minimizes the number of loop components. The experimental scheme is depicted in Fig. 1. The upper part of this set-up is used for spectral diagnostics: a small portion of the output beam of the laser under consideration (LD1) is coupled out and heterodyne-detected with the help of a local oscillator (LD2) and a fast photo diode (PD1). Its output spectrum is monitored with a rf spectrum analyzer (SA). The remaining part of the emission of LD1 is used for frequency stabilization. It is sent through an electro-optic modulator (EOM) and focused into a high-Q Fabry-Perot resonator (HQC).

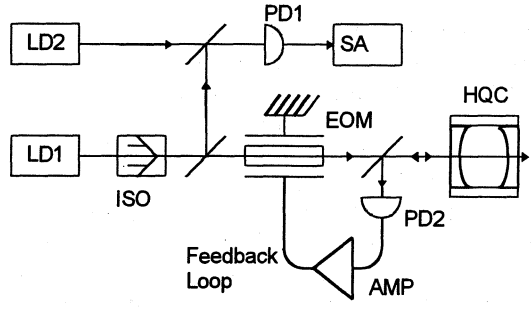


Fig. 1: Experimental scheme
LD: laser diodes, ISO: isolator, EOM: electro-optic modulator, SA: rf spectrum analyzer, PD: photo diodes, AMP: rf amplifier, HQC: high-Q-cavity

This reference cavity acts, in conjunction with a fast photo diode (PD2), which detects the reflected signal, as a FM/AM-converter for FM noise components which coincide with the offset between carrier and cavity resonance. The rf-component of the photo current is amplified and fed back to the EOM. For properly chosen feedback phase, the initial FM noise component is filtered and amplified yielding a larger rf photo current signal and vice versa. Hence, sustained oscillation sets in and the optical frequency of the resonant FM sideband (RS) becomes almost independent of fluctuations of the carrier (C) frequency whereas the FM noise amplitude of the non-resonant sideband becomes two times larger than the carrier noise. However, this unwanted , non -resonant sideband (NR) and the

carrier are not present in the transmitted signal of the cavity. The oscillation amplitude grows for sufficiently large rf gain until the sideband power reaches -5 dBc, i. e. the maximum of the first order Bessel function.

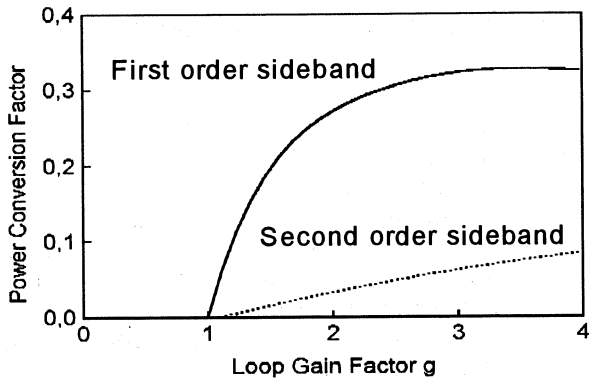


Fig. 2: Input-to-output conversion factor

Fig. 2 depicts the power of the first and second order sidebands, normalized to the input power, as a function of the loop gain g, normalized to its value at threshold,

$$g \approx \frac{\Delta\varphi}{2\,J_0(\Delta\varphi)\,J_1(\Delta\varphi)},$$

where $\Delta\varphi$ denotes the intra-loop oscillation amplitude and J_0 and J_1 the Bessel functions of zeroth and first order, respectively.
The parameter g combines factors like optical input power, quantum efficiency of the photo diode, gain of the rf amplifier, modulation sensitivity of the EOM and so on. Fig. 2 indicates that the optimum operation point of this oscillator is at gain values about 2.5 times (~8 dB) above threshold: the power of the desired first order sideband is already close to maximum whereas the second order sideband is still weak. Fig. 2 reflects a remarkable characteristic of this oscillator: gain saturation does not arise from nonlinear behavior of the gain medium but from mathematical properties of the Bessel functions.

3. EXPERIMENTAL RESULTS

Fig. 3 shows typical experimental output spectra of the frequency stabilizer, measured with the heterodyne-detection scheme shown in Fig.2. under the following experimental conditions:

- cavity with free spectral range of FSR= 2 GHz and resonance width of 30 kHz (FWHM), deduced from decay time of 10 µs
- EOM: 1*1*25 mm^3 $LiNbO_3$-crystal
- broad band rf power amplifier with 12 W output power
- LD1: free-running GaAlAs laser diode STC050, $\lambda \cong 850$ nm, $P_{out} = 50$ mW
- LD2: extended-cavity GaAlAs laser diode STC050, $\lambda \cong 850$ nm, $P_{out} = 10$ mW
- loop delay time $\tau_d \cong 40$ ns
- loop oscillation frequency $f \cong 150$ MHz.

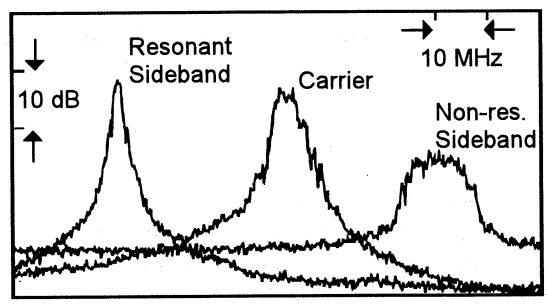


Fig.3: Heterodyne-detected emission spectrum of resonant sideband, carrier and non-res. sideband, detection bandwidth: 300 kHz

One clearly sees in Fig. 3 the expected behavior: the spectral width of the resonant sideband is substantially reduced with respect to the carrier whereas the non-resonant line is broadened. The width of the resonant line is already limited by the emission line width of the local oscillator. The frequency noise spectrum of the LO emission was measured to consist of an almost white frequency noise level of $S_\nu \cong 10^4$ Hz^2/Hz at high Fourier frequencies (f>100 kHz) and relatively strong 1/f-noise and acoustic noise contributions at low Fourier frequencies.
The considerable narrowing of the line wings of the resonant sideband, seen in Fig. 3, indicates that the effective control bandwidth is of the order of 10 MHz. For a more quantitative determination of the loop bandwidth, the carrier was frequency modulated at a few MHz adding a small ac current to the injection current of LD1. The unity gain frequency was deduced from the modulation width reduction of the resonant sideband. A typical spectrum for f_{mod} = 4.2 MHz is shown in Fig. 4.

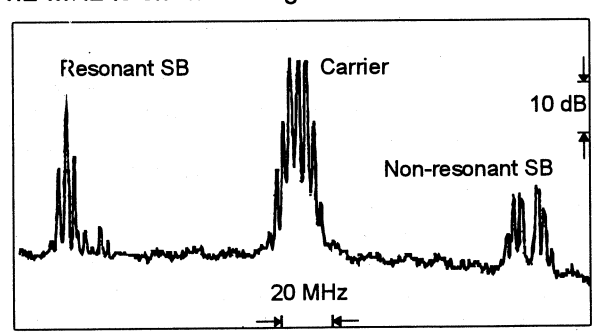


Fig. 4: Heterodyne-detected emission spectrum, carrier frequency modulated @ 4.2 MHz

The carrier displays a central line and first order sidebands of equal magnitude. The power of these lines is proportional to the square of the zeroth and first order Bessel function, yielding

$$J_0^2(\Delta\varphi_C) = J_1^2(\Delta\varphi_C)$$

and hence a phase deviation of the carrier of $\Delta\varphi_C \cong 1.4$ rad. The RS displays a central line which is 10 dB stronger than the first order sidebands,

$$J_0^2(\Delta\varphi_{RS}) / J_1^2(\Delta\varphi_{RS}) \cong 10$$

which leads to $\Delta\varphi_{RS} \cong 0.5$ rad whereas the central line of the NS is negligible small,

$$J_0^2(\Delta\varphi_{NS}) \cong 0,$$

yielding $\Delta\varphi_{NS} \cong 2.3$ rad. Thus, the initial modulation width of the carrier is reduced by 0.9 rad for the RS and increased by the same amount for the NS, as expected. From these values one calculates a unity gain frequency of

$$f_{ug} = f_{mod}\sqrt{(\Delta\varphi_C)^2 / (\Delta\varphi_{RS})^2 - 1} \cong 11\ \text{MHz}$$

Such a large control bandwidth leads to almost perfect FM noise cancellation at low Fourier frequencies. For example, one may expect a suppression ratio of

$$S_{\nu,RS} / S_{\nu,C} = (\Delta\nu_{cavity})^2 / (f_{ug})^2 \cong -57\ \text{dB}, \quad f \le \Delta\nu_{cavity}$$

where $S_{\nu,RS}$ and $S_{\nu,C}$ denote the frequency noise spectral density of the resonant sideband and the carrier, respectively and $\Delta\nu_{cavity}$ = 15 kHz the resonance width of the cavity.

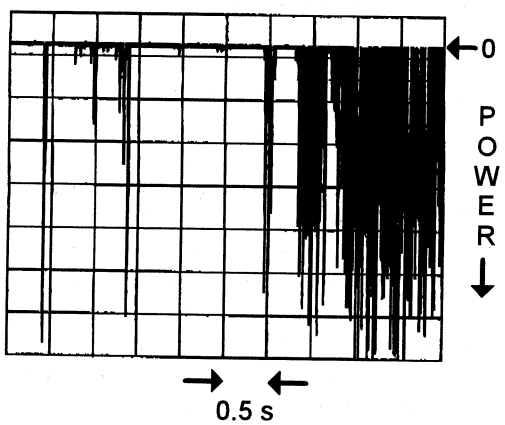


Fig.5: Typical oscilloscope trace of cavity output power, servo loop open, detection bandwidth ~1.2 MHz.

As a result, the emission line width of the carrier is reduced by some orders of magnitude, as seen in Fig. 5 and 6.

Fig. 5 shows a typical oscilloscope trace of the cavity transmission signal under open loop conditions. The frequency of the carrier was tuned to the vicinity of a cavity resonance. Due to its large frequency noise, the carrier accidentally meets the cavity resonance for a short period, yielding stochastic spikes as the output signal. The duration of these spikes is determined by the cavity decay time of 10 µs.

The situation changes dramatically under closed loop conditions, see Fig. 6. Now, the carrier frequency of the carrier was detuned by about 150 MHz from the cavity resonance. One sees, that the rms amplitude modulation of the transmitted signal which stems from the RS is much smaller than 1, even for detection bandwidths much wider than the cavity resonance width. This clearly indicates that the emission line width of the resonant sideband is substantially smaller than the cavity resonance width of 15 kHz.

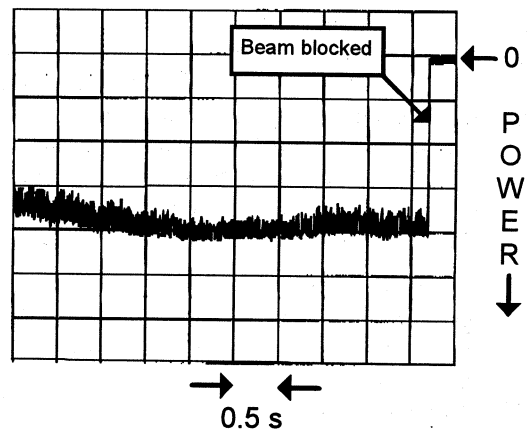


Fig.6: Typical oscilloscope trace of cavity output power, servo loop closed, detection bandwidth ~1.2 MHz.

4. SUMMARY

The paper presents a novel laser frequency stabilizer scheme which can be used for almost any cw laser. Furthermore, it can be extended to saturated absorption lines as frequency references. The device is inherently fast since it operates in difference frequency domain and requires no components for baseband conversion. Thus, the group delay of the feedback path is minimized yielding control bandwidths of 10 MHz and more. This makes the system particularly useful for noisy lasers like dye lasers and laser diodes and applications where the ~5 dB insertion loss can be tolerated. As a first demonstration, the emission line width of a GaAlAs laser diode has been reduced from the megahertz region to values below 10 kHz.

5. REFERENCES

[1] J. L. Hall and T. W. Hänsch, Optics Lett. 9, 502 (1984)

[2] M. Zhu and J. L. Hall, J. Opt. Soc. Am. B10, 802 (1993)

SUB-HERTZ STABILISATION OF CO_2 LASERS IN THE 30 THZ SPECTRAL REGION

V. Bernard, Ch. Daussy, G. Nogues, P.E. Durand, L. Constantin, A. Amy-Klein, and Ch. Chardonnet.

Laboratoire de Physique des Lasers, URA 282 du C.N.R.S., Université Paris 13, Avenue J.-B. Clément, 93430 Villetaneuse - France.

1. ABSTRACT

We developed a frequency stabilisation scheme of CO_2 lasers using only external modulations *via* an electrooptic modulator (EOM). One of the laser sidebands which are generated by the EOM and frequency-modulated is set in resonance with a Fabry-Perot cavity, filled with OsO_4. The saturation signal of an OsO_4 line detected in transmission of the Fabry-Perot cavity is used for the stabilisation. We obtained a stability of 0.1 Hz ($\Delta\nu / \nu = 3.5\,10^{-15}$) on a 100 s time scale, and a reproducibility up to 10 Hz with the strongest OsO_4 reference lines. As an application we measured 20 frequencies of CO_2 laser transitions relative to the OsO_4 grid with a precision of 100 Hz. The absolute frequencies differ by several kHz from results obtained with the saturated fluorescence technique and we attribute these differences to pressure shifts of several kHz in the conditions of this last method.

Keywords: CO_2 laser, stabilisation, electrooptic modulator, CO_2 frequency grid

2. INTRODUCTION

High-resolution spectroscopy and metrology have made important progress over the last twenty years and require better and better stabilized lasers (Refs. 1-3). During the last two decades, our group performed molecular spectroscopy with CO_2 lasers in the 10 µm spectral region. A typical resolution of 1 kHz limited by the transit width (Ref. 3) was routinely obtained and we were able to resolve most of the hyperfine structures of molecules. Concerning the metrological aspects, the most widely used secondary frequency grid in the 30 THz spectral region is based on the CO_2 laser transitions, for which the saturated fluorescence method allows for an accuracy of the order of several kHz (Refs. 4, 5). We demonstrated the possibility of using directly CO_2 saturated absorption signals as frequency standards (Ref. 6). There is also a more precise frequency grid based on OsO_4 saturation signals which reaches accuracies from 1 kHz (Ref. 7) to 50 Hz (Ref. 8).

In this paper, we present a frequency stabilization scheme of a CO_2 laser using only external modulations *via* an electrooptic modulator (EOM). One laser sideband generated by the EOM is stabilized onto a saturation peak of OsO_4 detected in transmission of a Fabry-Perot cavity (FPC). The modulation frequencies required for the detection of the cavity resonance and the molecular line are directly applied to the sideband through the synthesizer which drives the EOM. Thus, we removed the distortions induced by the modulations with piezo-electric transducers (PZT) and obtained a reproducibility of about 3.10^{-13}. In addition, this method gets rid of any modulation of the powerful stabilized laser carrier which can be more easily used as a local oscillator. Finally, the carrier can also be used as a tunable source by scanning the frequency of the synthesizer which drives the electrooptic modulator. As an application, we also present the comparison of 20 new absolute frequency measurements of CO_2 laser lines with previous values obtained with the saturated fluorescence technique.

3. NEW STABILISATION SCHEME

Let us first recall the characteristics and performances of our previous stabilization system (Ref. 9) which are essentially preserved in the new scheme. To ensure reproducibility and accuracy of the laser frequency, one requires an atomic or molecular transition as a frequency reference. Our experimental set-up was based on the stabilization of a CO_2 laser onto the saturated absorption peak of a strong OsO_4 line. We used, as an absorption cell, a Fabry-Perot cavity (FPC) filled with OsO_4 which improved the signal contrast by a factor of the order of the cavity finesse. Moreover, the cavity allows a better control of the beam geometry.

In order to detect the molecular line, the cavity resonance was locked to the laser frequency with a servo loop bandwidth of 100 Hz. This was accomplished using a 30 kHz frequency modulation of the FPC length *via* a PZT. The laser frequency also was modulated at 4.4 kHz *via* a PZT to detect the third harmonic of the molecular saturation signal. The laser frequency was locked onto this line with a servo loop bandwidth of 1 kHz. The performances of the stabilization method are summarized as following : the

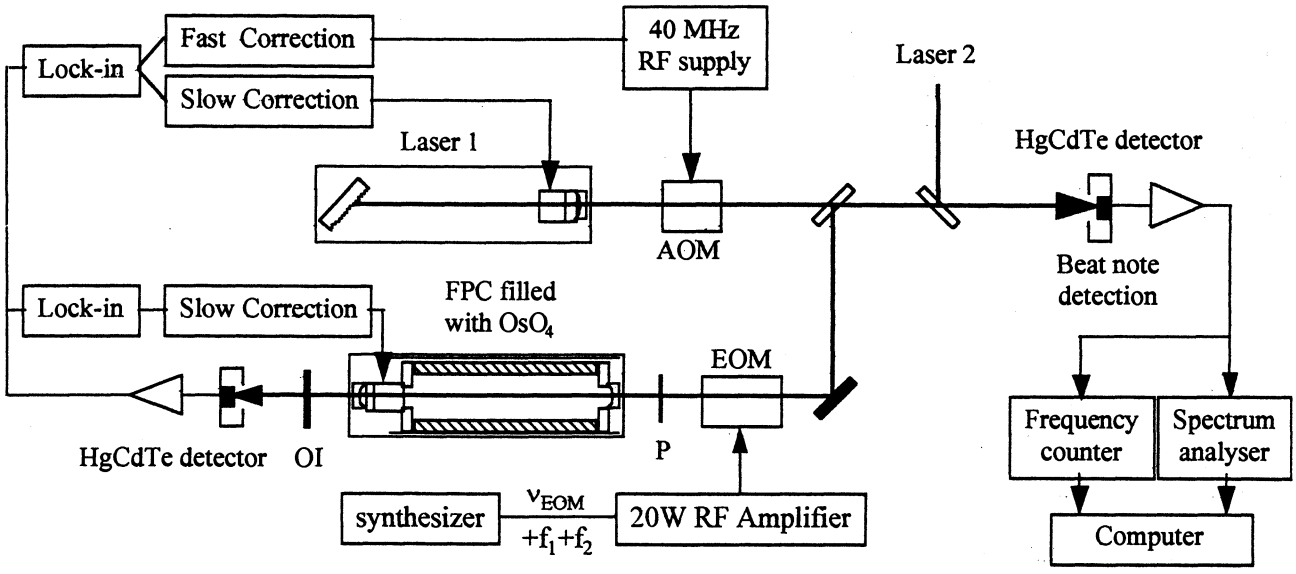


Figure 1.

square root of the Allan variance of the laser frequency reached a Flicker plateau at a level of 0.7 Hz between t=1 s and 50 s while we obtained a laser linewidth (FWHM) of 3.4 Hz.

Although the excellent long-term stability of the laser frequency, when reproducibility and accuracy were considered, we observed frequency offsets up to 500 Hz between two identical stabilized lasers that we attributed to modulation distortions induced by the piezo-electric transducers. The new scheme, presented now, overcomes most of the limitations associated with the usual modulation techniques in the 30 THz spectral region.

The basic idea of our new stabilization set-up (Ref. 10) is to use one of the two laser sidebands generated by an electrooptic modulator (EOM) as a new carrier. Thus, this sideband can be modulated by modulating the radio-frequency (rf) which drives the EOM. Then, the FPC is fed by one of the sidebands which can be locked onto the OsO_4 line. Figure 1 shows the experimental set-up, that differs from the previous one mainly by the introduction of an EOM and by the modulation method.

The principle of the electrooptic modulator is the following. It is a traveling-wave EOM, working as an amplitude modulator. The 2*2*40 mm^3 CdTe crystal is driven with a rf field which generates two sidebands. We obtained easily a bandwidth of 500 MHz which is still far from the theoretical limit. The frequency modulations that are required for detecting both, the cavity and molecular resonances are simply applied to the sidebands by modulating directly the rf synthesizer. This method allows the setting of a strong and easily controlled modulation index, and a broad choice of the modulation frequencies without facing the distortion problems induced by PZT's. With a 20 W rf power, we transfer up to 0.03 % of the laser power in each sideband.

The sidebands are generated with a polarization orthogonal to that of the carrier. Thus, we can pick up the sidebands as the reflection on a Brewster plate which transmits and, thus, rejects 99.8 % of the carrier power. However, the residual carrier reflected along the sidebands is yet about 10 times stronger than each sideband, and its frequency often coincides with higher-order modes of the FPC, so that it is not always efficiently filtered by the FPC. Note than an optical isolator set just after the FPC reduces the feedback of the residual reflection on the detector into the FPC and the laser cavity.

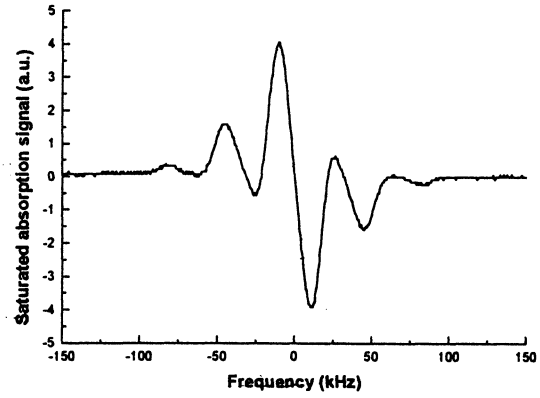


Figure 2.

With the new method of modulation, we completely avoid any amplitude modulation induced by the modulation of the laser gain profile when the cavity length is modulated by a PZT. The modulation frequencies were set to 90 kHz to lock the external

cavity resonance onto the laser sideband and to 4.9 kHz (with a 3f-detection) to lock the sideband on the molecular line. These frequencies are sufficiently different to filter efficiently each signal in each detection channel. Figure 2 presents a spectrum of the molecular frequency discriminator : it appears a comb of lines separated by 45 kHz due to the combined frequency modulations. The laser is locked onto the central component : we typically obtained, with a few 10^{-2} Pa molecular gas and 50 μW power inside the cavity, a peak-to-peak linewidth of 20 kHz and a signal-to-noise ratio (SNR) of 1000 with a 1 ms time constant. Note that, by constrast with the modulation applied by PZT, no asymmetry of the signals as well as no phase drifts were detected. Slow corrections (bandwidth of about 100 Hz) of the cavity length and laser frequency are applied via the PZT's, as previously. The fast correction of the laser frequency is applied to the rf source of an AOM with a correction bandwidth of 4 kHz, which could be improved in the future.

4. RESULTS

Figure 3 shows the beat note between two independent lasers stabilized on the P(46) A_1^2 (–) OsO_4 line located at + 115 209,15 kHz from the center of the P(14) CO_2 laser line. Note that, actually, the beat note of the laser carriers is recorded or measured while one of the sidebands of each laser is locked onto the molecular line. The FWHM is 12 Hz corresponding to 6 Hz per laser.

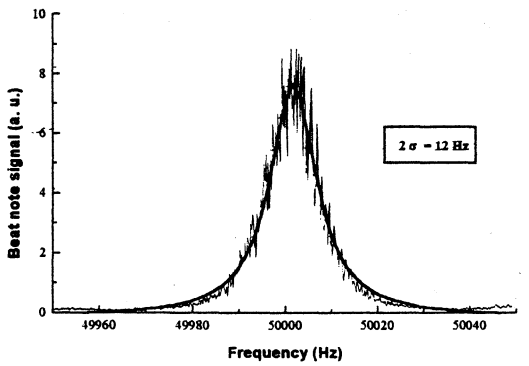


Figure 3.

Figure 4 gives the square root of the Allan variance of the beat frequency. It goes down with a slope of - ½ up to $\tau \approx 100$ s indicating a white frequency noise and thus a correct functioning of the servo loop. The minimum is reached at a level of 0.1 Hz, that is $\Delta\nu/\nu = 3.5\,10^{-15}$, at τ = 100 s. This improves by a factor 7 the results obtained with a PZT modulation. For $\tau \geq 100$s and up to 500 s, the square root of the Allan variance stays below $\Delta\nu/\nu = 10^{-14}$, and goes up with a slope near + 1 which is the signature of a linear frequency drift.

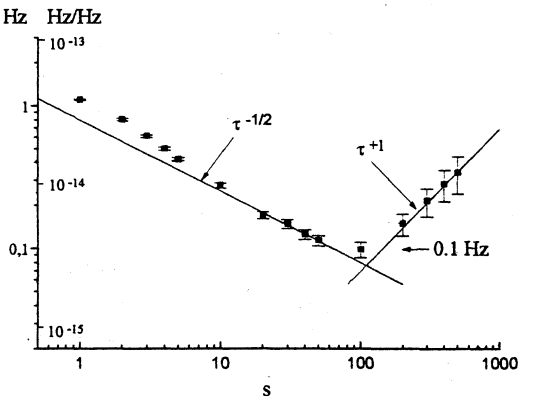


Figure 4.

To complete this analysis, we recorded the long-term frequency fluctuations (Fig. 5) : the laser frequency stays within an interval of 5 Hz for 5000 s. For longer time, the frequency exhibits oscillations whose peak-to-peak amplitude is less than 20 Hz, with a few hours period. We attribute them to residual interference fringes due to an imperfect optical isolation, which is the weak point of our set-up. These fringes are probably localized between the EOM and the Fabry-Perot cavity and between the cavity and the detector.

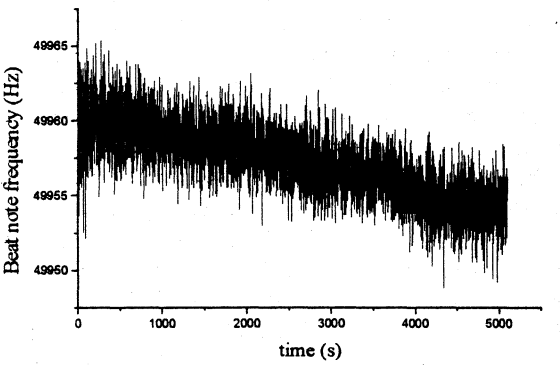


Figure 5.

Then, we analyzed the performance of our system in terms of reproducibility. Figure 6 shows the histogram of beat note frequencies when the sidebands of the two systems are locked, both, onto the P(46) A_1^2 (–) OsO_4 line. These measurements were performed within a period of 20 days. The radio-frequencies of the two EOM's differed by 50 kHz in order to measure beat notes around 50 kHz instead of zero.
The mean frequency is 50.002 kHz with a standard deviation of 10 Hz. This demonstrates an excellent reproducibility of our system (3.10^{-13}). We observed that, when slightly changing some experimental conditions like the optical isolation or the cavity alignment, the center frequency could be shifted up to 50 Hz. Note that these optical problems could be

seriously limited with a better rejection of the carrier and with a modulation of the optical path length to scramble the fringes.

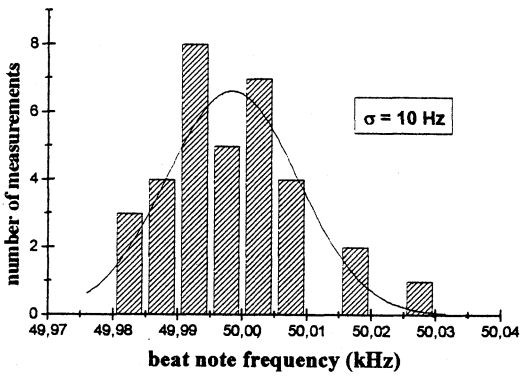


Figure 6.

However, this new modulation method largely improves the performances of our stabilization system. We did not observe any modulation distortion, proving the clean characteristics of this external modulation technique. Another crucial advantage of this set-up is the possibility to shift continuously the sideband frequency up to 500 MHz away from the CO_2 gain profile. The tunability of our system is thus largely enhanced and this allows us to reach some new and better molecular reference lines, not accessible with CO_2 lasers tunable over only 150 MHz for each laser line.

5. CO_2 FREQUENCY MEASUREMENTS

As an application to this new stabilisation scheme, we measured by beat note experiment the frequency difference between 20 CO_2 laser lines and the close OsO_4 lines belonging to the OsO_4 frequency grid.

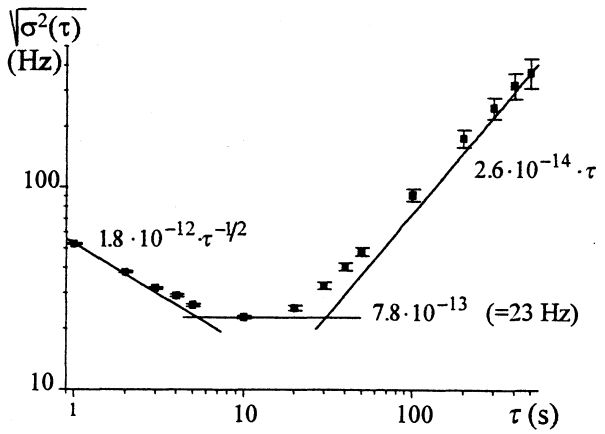


Figure 7.

The standard conditions for the detection of the CO_2 lines were a pressure of 0.1 Pa and a laser power of 1 mW. A first harmonic detection was preferred because of the weakness of the signal. The servo-loop bandwidth was reduced down from 1 kHz to about 100 Hz because of the lower signal-to-noise ratio than for a OsO_4 line; the laser linewidth was then of a few hundreds of Hertz. The square root of the Allan variance of a laser stabilized on the R(16) CO_2 laser line is displayed on Figure 7, the flicker plateau of 23 Hz was reached at $\tau = 10$ s.

We recorded twenty OsO_4/CO_2 frequency differences from P(24) to R(26) on the 10 μm branch of our CO_2 lasers, where OsO_4 references are available, by using the frequency measurement procedure which is detailed in a forthcoming paper (Ref. 11). The typical uncertainties of our measurements are from 50 to 200 Hz and comparable to the day-to-day repetability. We could also control the quality of our measurements by comparing the results for P(12), P(14) (Ref. 6) and R(10) (Ref. 12) obtained a few years ago with a slightly higher precision due to the narrower linewidth of the CO_2 lines obtained by using our 18 m long cell. The two sets of measurements, although they were obtained with two very different experimental methods and conditions, are fully compatible. Finally, the absolute frequencies are deduced using the OsO_4 frequency grid with accuracies between 50 Hz and 1,4 kHz.

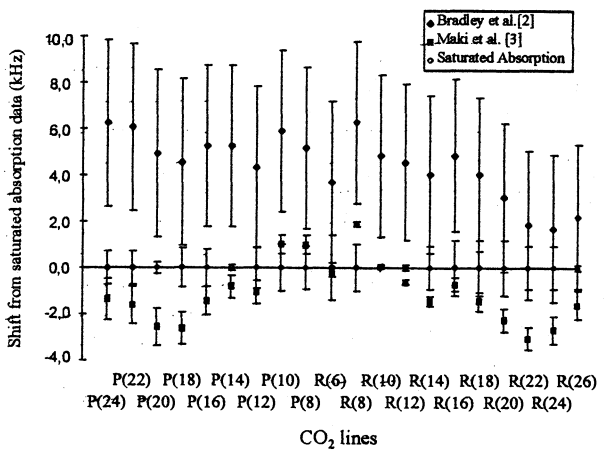


Figure 8

Figure 8 displays the frequency shift of the CO_2 grid of Bradley *et al.* (Ref. 4) and Maki *et al.* (Ref. 5) relative to our data. These grids are mainly obtained from CO_2 measurements with the saturated fluorescence technique. It is remarkable that, roughly speaking, our measurements are between the values of Maki and Bradley although these data come mainly from the same measurements. We can also observe that the error bars of these data are underestimated if we consider that our measurements are certainly more precise that those performed with the saturated fluorescence technique for which the line is wider by about two orders of magnitude. It must be noticed that the frequencies of these two grids are not the direct result of measurements. In contrary, the frequencies were re-calculated from the molecular constants obtained after a global fit of all the measurements. The uncertainties take into account only the statistical residuals coming

from the fit and no systematic error was included, as it should be to estimate properly the actual accuracy of the data. For example, let us remind that a red pressure shift of -2 to -4 kHz under the standard conditions of the saturated fluorescence technique is predicted following the saturation measurements of (Ref. 12) in agreement with theory (Ref. 13). However, an anomalous blue shift of the saturated fluorescence signal was reported in (Ref. 14) but not taken into account for the calculation of the frequencies of (Ref. 4) because important instrumental errors were simultaneously observed. For these reasons, we believe that the discrepancy of a few kHz between our measurements and those based on the saturated fluorescence technique is an indication of the present limitation of this last method in terms of accuracy.

6. CONCLUSION

We developed a frequency stabilization scheme using only external modulations *via* an EOM, that improves significantly the performances of our previous system. The square root of the Allan variance reaches the 0.1 Hz level at 100 s and stays below 0.3 Hz up to 500 s. This corresponds -to our knowledge- to the best long-term stability obtained in that frequency range (Ref. 15). The cleaner characteristics of the modulations prevent from any molecular line distortions and lead to a few tens of Hertz reproducibility of the system. To improve these performances the point is certainly to have a better optical isolation. Besides the stability performances, the external modulation method proved to be very convivial as there is no real technical limitations for the modulation index and frequency, so that the experimental parameters can be very easily adjusted.
These new developments with its various advantages allow us to consider a new generation of experiments. For instance, the use of an unmodulated local oscillator will be essential for the observation of 50 Hz-Doppler-free two-photon Ramsey fringes on a supersonic beam of SF_6, that we plan to perform. With this ultra-stable laser, it can now be considered to test parity violation in molecules : we plan to measure very small frequency differences in the spectra of two enantiomers of a chiral molecule. Recent theoretical estimations (Ref. 16) indicate that the present stability of our lasers is compatible with the observation of such an effect. For that purpose, we plan to use a second EOM as described above, working with the same principle as the first one but with a tunability range of 8-18 GHz on each side of the laser carrier. Finally, the stability and reproducibility of our system open the way to significant progress in the field of metrology. We were able to carry out new measurements of the frequency differences between OsO_4 lines and CO_2 laser resonances. This led to an enhancement of more than one order of magnitude of the accuracy of the CO_2 frequency grid and confirmed that the detection of the cavity-enhanced saturation signal is a good alternative to the usual saturated fluorescence method.

7. ACKNOWLEDGMENTS

This work has been supported by ETCA-DRET contract # 20.398/92, by Bureau National de Métrologie (contracts # 93-2-46-0032 and 96-3-001), by the European Union (CHM contract # ERBCHRXCT94-0633 and -0603) and by the Conseil régional d'Ile de France (LUMINA contract #E-946-95/98).

8. REFERENCES

[1] C. Freed, D. L. Spears, R.G. O'Donnel and A. H. M. Ross, in *Laser Spectroscopy*, R. G. Brewer and A. Mooradian, Eds. New York Plenum, pp. 171-191, 1974.
[2] A. Van Lerberghe, S. Avrillier and Ch. J. Bordé, IEEE Journal of Quantum Elec., vol. QE-14, pp. 481-486, 1978.
[3] Ch. Salomon, Ch. Bréant, A. Van Lerberghe, G. Camy and Ch. J. Bordé, Appl. Phys. B, vol. 29, pp. 153-155, 1982.
[4] L. C. Bradley, K. L. Soohoo and C. Freed, *IEEE J. of Quant. Electron.*, vol. 22, pp. 234-267, 1986.
[5] A. G. Maki, C.-C. Chou, K. M. Evenson, L. R. Zink and J.-T. Shy, *J. Mol. Spectrosc.*, vol. 167, pp.211-224, 1994.
[6] Ch. Chardonnet, A. Van Lerberghe, and Ch.J. Bordé, Opt. Comm., vol. 58, pp. 333-337, 1986.
[7] A. Clairon, A. van Lerberghe, Ch. Salomon, M. Ouhayoun and Ch. J. Bordé, Opt. Comm., vol. 35, pp. 368-372, 1980.
[8] A. Clairon, B. Dahmani, A. Filimon and J. Rutman, IEEE Trans. Instrum. Meas., vol. IM-34, pp. 265-268, 1985.
[9] V. Bernard, P. E. Durand, T. George, H. W. Nicolaisen, A. Amy-Klein and Ch. Chardonnet, IEEE Journal of Quantum Elec., vol. 31, pp. 1913-1918, 1995.
[10] V. Bernard, C. Daussy, G. Nogues, L. Constantin, A. Amy-Klein and Ch. Chardonnet, *to be published in IEEE J. of Quant. Electron.*
[11] V. Bernard, G. Nogues, C. Daussy and C. Chardonnet, submitted to Opt. Comm.
[12] A. Clairon, O. Acef, Ch. Chardonnet and Ch. J. Bordé, *Frequency Standards and Metrol.*, A. de Marchi Ed., pp. 212-221, 1989.
[13] D. Robert, J. Bonamy, *J. Phys.* (Paris), vol. 40, p. 923, 1979.
[14] K.L. SooHoo, C. Freed, J.E. Thomas, and H.A. Haus, *Phys. Rev. Lett.*, vol. 53, pp. 1437-1440, 1984.
[15] O. Acef, *Opt. Commun.*, vol. 134, pp. 479-486, 1997.
[16] L.N. Ivanov, and V.S. Letokhov, to be published in Proc. of Russian Acad. Sci (1997).

ABSOLUTE FREQUENCY STABILIZATION OF SOLID STATE ER-YB LASERS AT 1.5 μm

E. Bava and C. Svelto

Dipartimento di Elettronica e Informazione del Politecnico di Milano,
P.zza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: bava@elet.polimi.it Phone: +39(2)239936-09/01 Fax: +39(2)23993635

P. Laporta and S. Taccheo

Centro di Elettronica Quantistica e Strumentazione Elettronica del CNR,
Dipartimento di Fisica del Politecnico di Milano,
P.zza Leonardo da Vinci 32, 20133 Milano, Italy

ABSTRACT

1. INTRODUCTION

Owing to their intrinsic characteristics of short term frequency stability (improved by absolute stabilization to a molecular reference) and relatively high output power, diode-pumped solid-state Er-Yb:glass lasers are attractive sources for applications in optical communications, spectroscopy and metrology. The Er-Yb microlaser developed at Politecnico of Milan [1] can provide for output powers up to 20 mW, in a pure TEM_{00} mode under single frequency operation with a wide wavelength tunability range. A single laser device can also be operated in a multi-longitudinal-mode operation allowing for a comb of 4 to 10 optical carriers with a frequency spacing of 40 to 20 GHz between adiacent ones. For both high-density wavelength division multiplexing (HDWDM) [2] and frequency metrology the absolute stabilization of the emitted wavelength(s) is of great interest. In this work the obtained results of frequency stabilization by two different techniques with respect to molecular lines of the acetylene molecules will be reported.

2. EXPERIMENTS

The Er-Yb microlaser (Fig. 1) consists of a plano-spherical resonator longitudinally pumped by the 978 nm radiation emitted by a InGaAs laser diode (Spectra Diode Laboratories Mod. 6460). The active medium is a phosphate glass disk with diameter of 6 mm and 2.5 mm thick containing dopant concentrations of 10^{19} ions/cm^3 Er^{3+} and 10^{21} ions/cm^3 Yb^{3+}. The overall cavity length, including the anular PZT transducer for fine frequency tuning, is less than 5 cm. The output beam is linearly polarized, diffraction limited with circular shape.

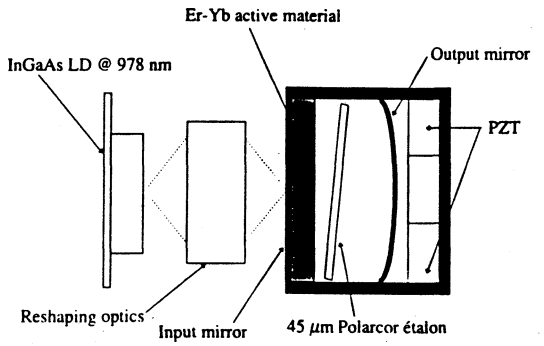


Fig. 1 The Er-Yb microlaser.

In a first set of experiments, see Fig. 2, two identical Er:Yb:glass lasers have been stabilized at two different points of the P(13) vibrational-rotational line of C_2H_2 (λ=1532.828 nm) by the fringe-side locking technique [3].

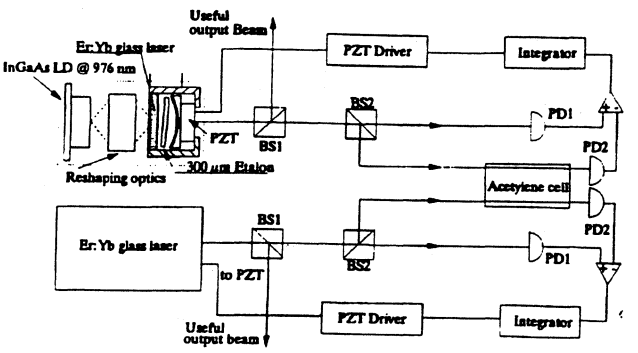


Fig. 2 The fringe-side-locking experimental setup.

Upon detecting the beat note between the two stabilized lasers, a long term (4-hours) fluctuation

of the frequency difference of about ±1 MHz has been measured. By analysis of the beat note spectrum, the short term (~1 ms) linewidth of each laser has been evaluated to be narrower than 50 kHz. The stability measurements are reported in Fig. 3.

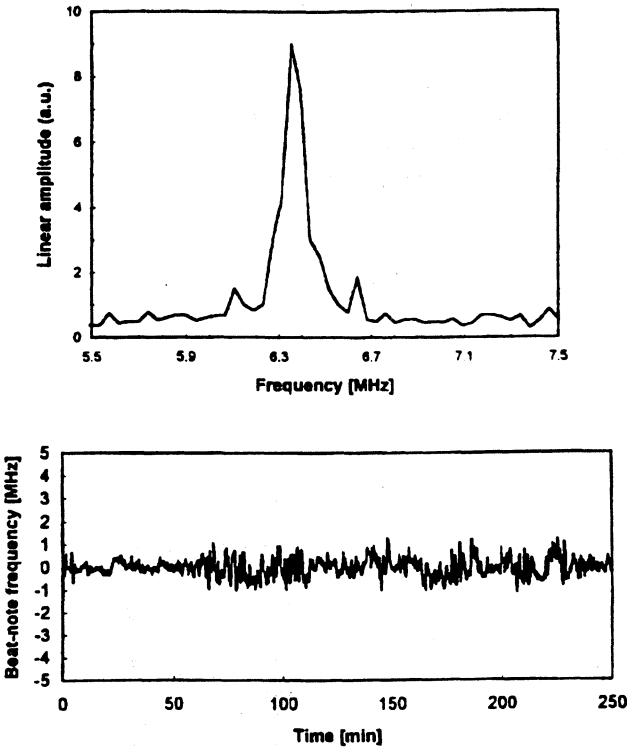


Fig. 3 Short-term linewidth and long-term frequency stability of the Er-Yb laser locked to the side of the P(13) line of C_2H_2.

Absolute frequency stabilization by external phase modulation and synchronous detection after the gas cell, FM-eterodyne technique, has also been demonstrated for one source at 1534.097 nm. The experimental setup is depicted in Fig. 4. The time analysis of the error signal showed a peak-to-peak frequency deviation lower than 80 kHz over several hours of robust locking (Fig. 5). A second Er-Yb laser will now be locked to the center of P(15) line of C_2H_2 at 1534.097 nm to allow for beat note measurements and for an evaluation of the Allan variance. The aim is to achieve a long term frequency stability and reproducibility of less than $\Delta\nu \cong 100$ kHz. The corresponding frequency stability and reproducibility ($\Delta\nu/\nu \approx 5\times10^{-10}$) should establish a stable reference in the 1.5 μm spectral region based on this novel solid-state laser source [4]. By means of the tunable Er-Yb laser some fine spectroscopic measurements of absorption lines at 1.5 μm can also be performed. As an example, in Fig. 6, we report the measured absorption profile of P(15) line of C_2H_2 at a pressure of 10 mbar.

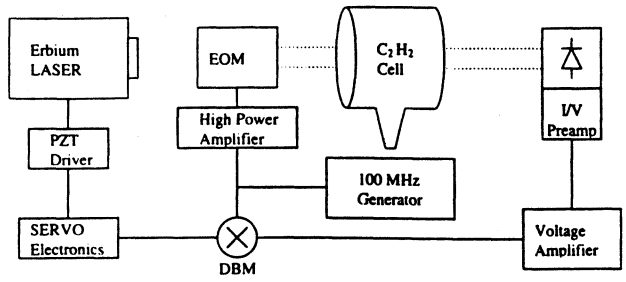


Fig. 4 Experimental setup of the FM-eterodyne locking.

In another set of experiments, the Er:Yb:glass laser has been successfully tuned around the 1550 nm wavelength [5]. The wide continuous tunability interval obtained, spanning over 36 nm, suggested to investigate the possibility of frequency locking these sources to the molecular lines of $^{13}C_2H_2$ [6]. With the same laser, several absorption lines have been matched from P(13) at 1540.057 nm to P(29) at 1550.870 nm. Preliminary fringe-side locking experiments, at these higher wavelengths, have also been conducted.

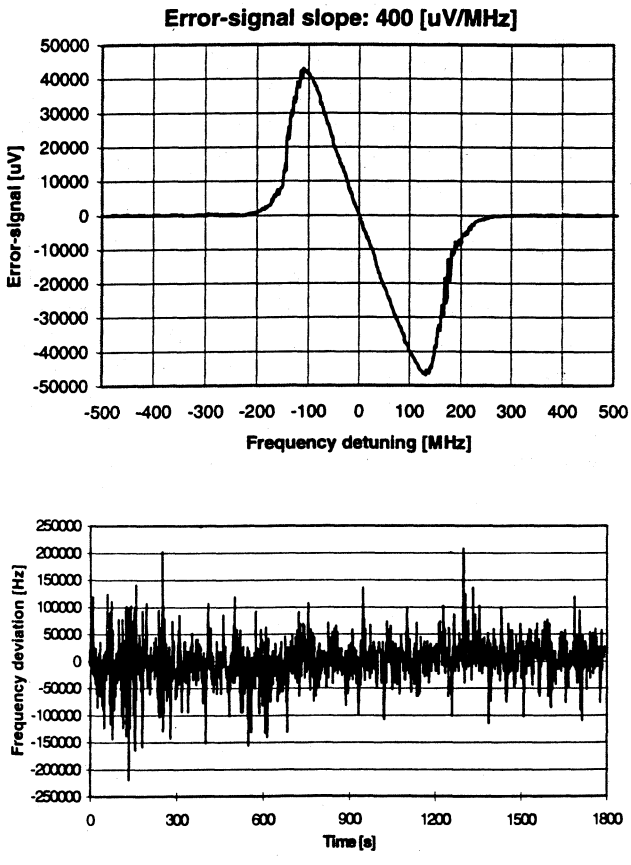


Fig. 5 Frequency to voltage discriminator curve and time behavior of the error signal in the FM-eterodyne experiment on the P(15) line of C_2H_2.

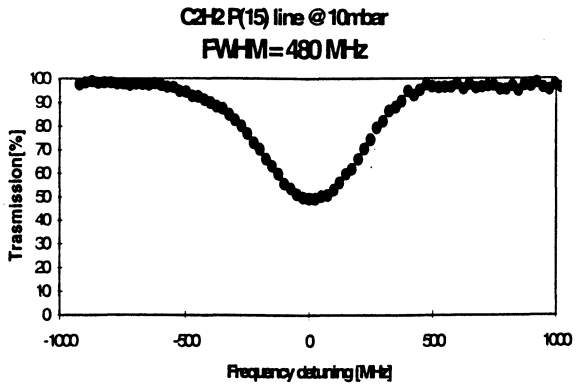


Fig. 6 Absorption profile of line P(15) of C_2H_2 at 1534.097 nm.

3. CONCLUSIONS

The Er-Yb microlaser has been frequency stabilized to different molecular lines of the acetylene molecule at 1.5 μm. The wide wavelength tunability (continuous tuning range from 1528 nm to 1564 nm) in single frequency operation and the frequency stability, when locked to the molecular reference ($\Delta\nu$<50 kHz @ 1 ms and $\Delta\nu\cong$100 kHz @ 10 h), open several possibilities of application in the fields of optical communications, spectroscopy and frequency metrology.

REFERENCES

[1] P. Laporta, S. Longhi, S. Taccheo, O. Svelto and G. Sacchi, Electron. Lett. **28**, 2067 (1992).
[2] A. H. Gnauck, A. R.Chraplyvy, R. W. Tkach and R. M. Derosier, Electron. Lett. **30**, 1241 (1994).
[3] P. Laporta, S. Taccheo, S. Longhi, L. Pallaro, E. Bava and C. Svelto, Opt. Lett. **20**, 2420 (1995).
[4] P. Hansell, and H. Leeson, NPL-Requirement Survey on "Absolute frequency standards for optical communications" Doc. PRO-AFS_FR-2 (1992).
[5] P. Laporta, S. Taccheo and C. Svelto, Appl. Phys. Lett. **68**, 2621 (1996).
[6] C. Latrasse, M. Breton, M. Têtu, N. Cyr, R. Roberge and B. Villeneuve, Opt. Lett. **19**, 1885 (1994).

SAW FILTERS

Chairman: P. Defranould

Main approaches to design of low loss SAW filters for mobile communications

(abstract)

C. Lambert, S. Kondratiev, V. Plessky, and T. Thorvaldsson
MICRONAS Semiconductor SA, CH-2022 Bevaix, Switzerland

As the wireless telecommunications industry continues to expand rapidly, there is an increasing demand for new surface acoustic wave (SAW) filter technology. To design approaches are currently dominating: coupled resonator filters (CRFs) and ladder type filters. The last idea was proposed about 10 years ago by Japanese scientists. The idea is actually very simple: long SAW transducer (or SAW resonator) has impedance dependence on frequency very similar to classic LC resonator, or to quartz resonators. It was known for many years how to build the filters using crystal quartz resonators. The authors proposed to used SAW resonators to build the same ladder filters, as using quartz bulk wave resonators. The idea was not appreciated by the SAW community for quite a long time, and only when other companies a few years ago also demonstrated very attractive results the real boom started.

This type of filters radically differ from all others. In these filters the SAW itself plays an auxiliary role: the signal is **not** transformed into SAW, as in classic SAW filter, and then re-converted back. Instead of that the change of impedance of SAW resonator is exploited, exactly like in quartz crystal filters. The separate resonators are not "talking" acoustically, through the SAW, so no signals is carried by SAW from the input to the output. Because of the primary role of the impedance changes we call these class of filters *"SAW impedance element filters (IEFs)"*, but the names "ladder" or "lattice filters" are also used.

In this paper we will compare CRFs and IEFs and we will try to show that the IEFs have essential advantages over all other types of low loss filters and so they will probably occupy the great part of the market, with some particular exceptions, and we will concentrate on Balanced bridge type of IEFs, recently invented in Micronas SA .

We will discuss the principals of operation and design of impedance element filters, estimate achievable parameters and limitations, discuss advantages of this design approach compared to others types of low loss SAW filters. A few examples of filters designed in Micronas (SAW division) will be presented.

We believe that in most cases the SAW filters which will be used in the next generations of mobile phones will be impedance element filters. For 2 GHz range it will be the only real possibility, but evident advantages of the IEFs will gradually replace all other types of RF filters in 900 MHz range too. As to IF filters the balanced bridge filters have significant advantages both in performances and in size/cost and they match perfectly to balanced circuitry already now widely used in IF stages of handhold phones.

A NEW LOW IMPEDANCE BALANCED DRIVE STRUCTURE FOR SAW TRANSVERSELY COUPLED RESONATOR FILTERS*

Marc SOLAL

THOMSON MICROSONICS
SOPHIA ANTIPOLIS, FRANCE

* This work was partially sponsored by France Telecom

ABSTRACT

Waveguide coupling of SAW resonators is widely used to make low loss compact filters. The classically used structure for this kind of device consists of two tracks separated by a common ground bussbar. To obtain a balanced drive using this structure, it is necessary to split the transducer in two parts connected electrically in series. This can result in load and source impedances which are too high for system requirements.

We present a new multitrack balanced drive waveguide coupled resonator filter structure exhibiting both low impedances and small crystal sizes. A general analysis procedure has been developed to analyse this new structure. This procedure is suitable for an arbitrary number of layers and for any number of electrical ports. Results of a GSM/DCS 1800 IF filter at 188 MHz using the new structure will be presented. Typical performances are a 3 dB bandwidth of 250 kHz, a 20 dB bandwidth of 500 kHz and a 6 dB insertion loss.

1. INTRODUCTION

One major application of SAW is mobile phone I.F filters. Increased IF frequencies make coupled resonator filters well suited for that application. Proximity coupled SAW resonators achieve both low insertion losses and very compact sizes, both needed in mobile phones. The evolution of integrate circuit technologies results in the need for balanced drive devices. The usual filter structure using a center ground bussbar is not well suited for balanced drive. In the past, two main approaches were used to handle balanced drive. The first one [3] consists of splitting one transducer into parts. This results in a large impedance increase, often incompatible with system requirements. A second approach, [5] consists of dividing the center bussbar in two parts. This approach solves the high impedance problem, however two filters have to be connected in parallel to avoid degraded rejection. This results in larger crystal sizes. We propose a third approach based on a three resonators configuration.

2. STRUCTURE DESCRIPTION

2.1 Unbalanced resonator filter

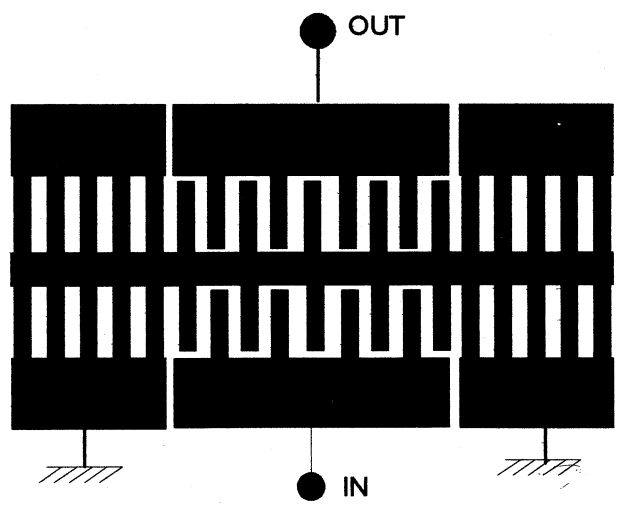


FIGURE 1 : CLASSICAL UNBALANCED DRIVE TWO POLE PROXIMITY COUPLED RESONATOR FILTER

Figure 1 shows the classical unbalanced two pole filter. By coupling, each resonator mode is split into two modes, one which is symmetric in the transverse direction (parallel to the metal strips) and one which is antisymmetric. By moving the two resonators closer, the coupling becomes stronger and the frequency shift between the two modes (and also the filter bandwidth) increases.

In this configuration, the center bussbar is common to the input and the output and has to be connected to ground. This does not directly allow balanced drive and different structures have to be found. Usually, 4 pole filters are made by cascade of 2 pole filters. So, only balanced drive input, unbalanced output structures are needed

2.2 Serial connection balanced drive structure

One approach to obtain balanced drive on one side is to split one of the transducer into two

parts. This gives the configuration given in the figure 2. Obviously, 4 or 6 poles can be obtained by cascading several 2 poles, with or without a parallel coupling inductor.

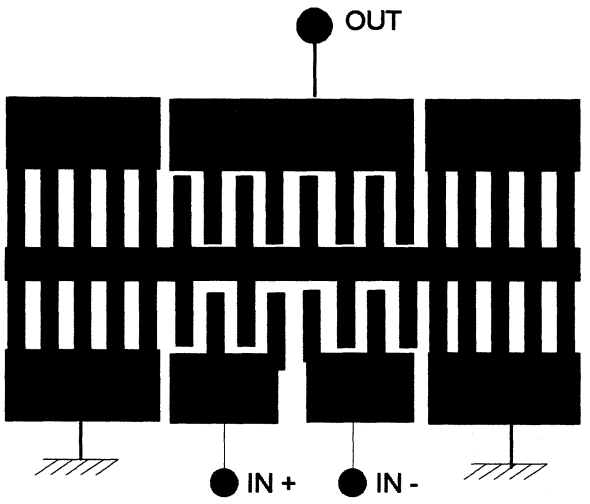


FIGURE 2 : SERIAL CONFIGURATION FOR BALANCED DRIVE

This configuration is equivalent acoustically to the unbalanced configuration, however the serial connection of two half transducers increases the electrical impedance by a factor of 4. For several applications, especially for the wide band applications, the resulting impedances are too large to be useful.

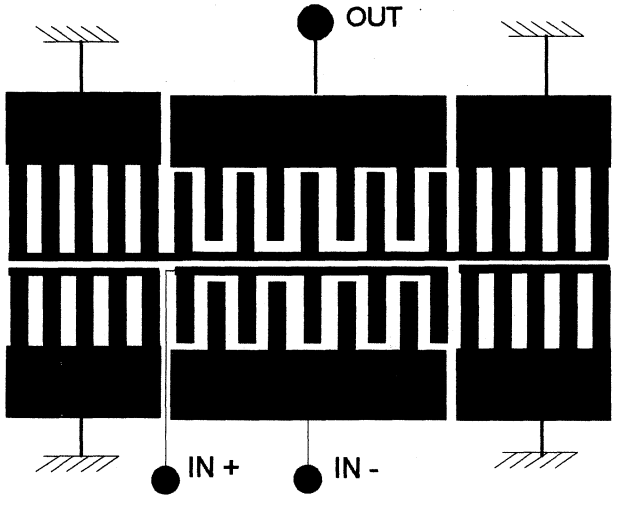


FIGURE 3 :SPLIT CENTER BUSSBAR CONFIGURATION

2.3 Parallel connection balanced drive configuration

A second approach is to keep the same acoustical structure and cut the center bussbar as shown in figure 3. This allows one transducer to be balanced driven with the same impedances as for the conventional unbalanced structure.

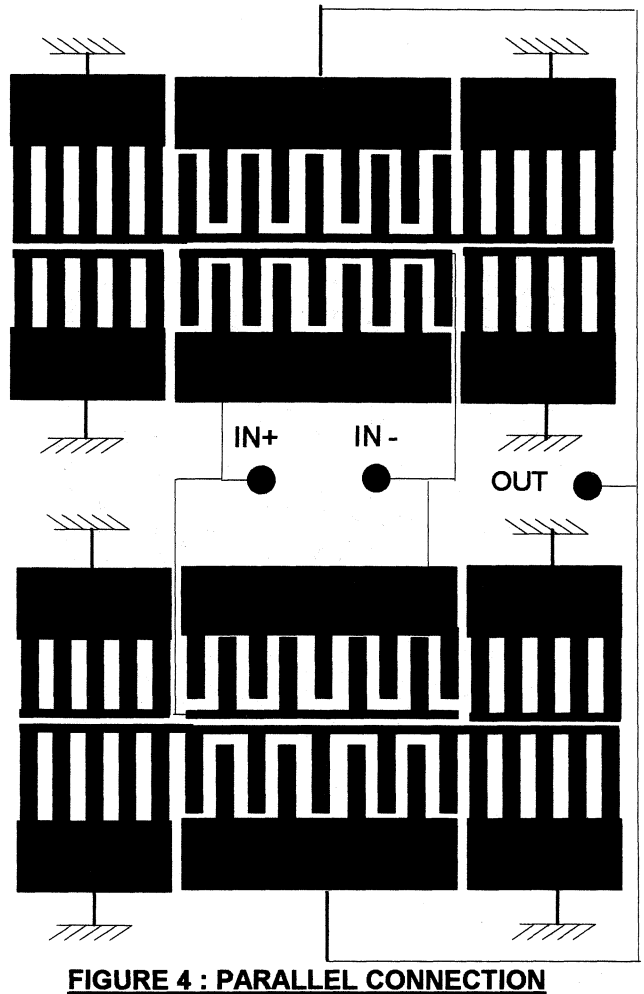


FIGURE 4 : PARALLEL CONNECTION BALANCED DRIVE STRUCTURE

As pointed out in [5], the main problem here is the electrical asymmetry in the capacitances between the two input ports and the output port, resulting in important direct feedthrough. This problem is solved in [5] by connecting two filters in parallel, thus making the structure symmetric again.

The feedthrough problem is now suppressed and the impedance is lowered by a factor of 2 due to the parallel connection. The only disadvantage of this structure is the need for 2 filters in parallel, requiring increased substrate width.

Next, we will present an alternate structure which has a low impedance with electrical symmetry and with a substrate size close to the classical unbalanced structure.

2.4 The three resonator balanced drive structure

Figure 5 shows the three resonator balanced drive structure. This structure keeps the split bussbar idea described in section 2.3 allowing low impedances but the symmetry is now obtained by using a three resonator structure.

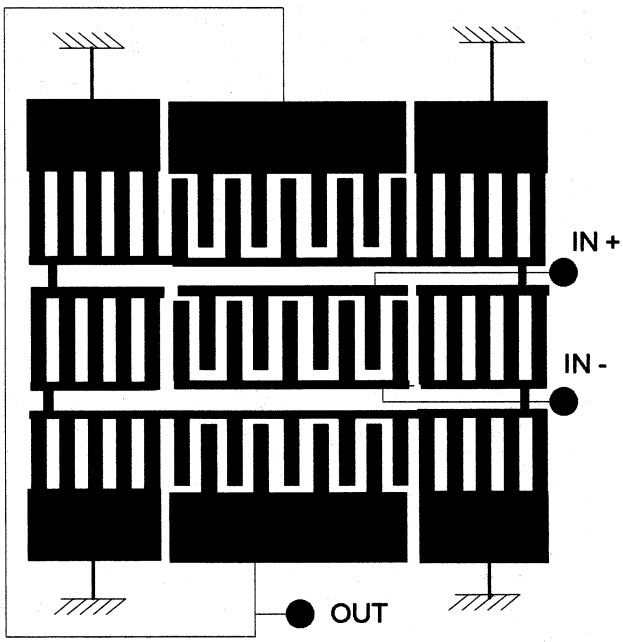


FIGURE 5 : THREE RESONATOR BALANCED DRIVE STRUCTURE

The center transducer of figure 5 is balanced driven while the two external transducers are connected in parallel. This structure is electrically symmetric and can be used as a two pole filter. A four pole filter is obtained a by cascading two.

For the small apertures generally used, the three resonator structure has two transversely symmetric modes and one antisymmetric mode. Due to the symmetric electrical excitation, only the two symmetric modes are coupled. Figure 6 shows typical transverse mode shapes for this structure.

As for the conventional two pole filter, coupling between the center transducer and the two external ones is of opposite sign for the two symmetric modes contribution. This means that with an adequate choice of geometry, the three resonator structure is compatible of a two pole design.

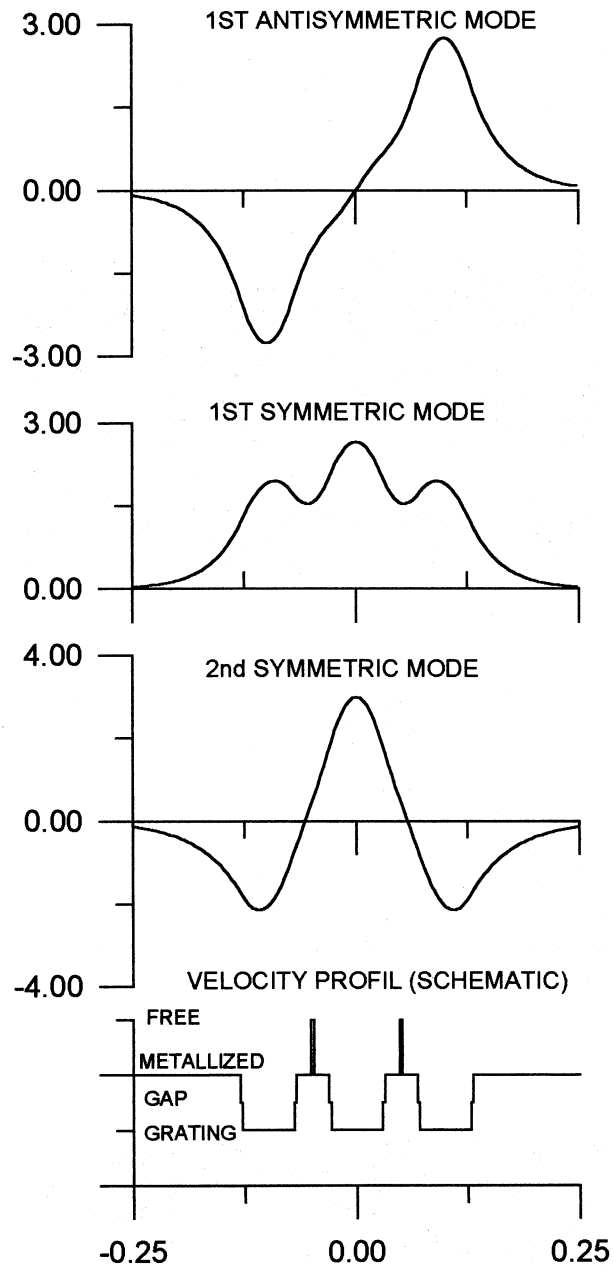


FIGURE 6 : EXAMPLE OF A THREE RESONATOR STRUCTURE AND CORRESPONDING MODE SHAPES

3. ANALYSIS PROGRAM

Due to the number of degrees of freedom of the three resonator structure, an analysis based on test device measurement data [10] was found insufficient for device optimization. Furthermore, this technique is structure dependent and does not allow a quick investigation of different structures. To alleviate these shortcomings, an analysis program was developed to work directly from the geometrical dimensions.

Following Haus et al. [6], the mode shapes and velocities vary inside the stop band and a correct model has to take into account the different reflectivities in the different regions. For synchronous resonators, the resonant frequency is at the beginning of the stop band and it is valid to consider the grating as an homogeneous material with simply a velocity change due to the reflectivity.

We choose to use a stack matrix theory [8] based model. This theory gives the mode velocities and shapes for an arbitrary geometry. As pointed out by Morgan [7], one key point is to use the correct velocity for the different regions. Inside the transducer-grating regions, our FEM BEM periodic model [4] was used to compute the correct velocity and reflectivity coefficient. This model, for example, can take into account the electrode shapes which are strongly dependent on the technology used to manufacture the filter.

Knowing the mode velocities and shapes, the filter is analyzed by using a classical P Matrix cascade model [11]. The basic principle of the model is to divide a transducer in several unit cells (usually the unit cell is a single period), to compute the P matrix of a single unit cell and then to cascade the P matrices to obtain the overall P matrix of the transducer. The transversely coupled device model follows the same approach as the basic model except that the unit cell P matrix now has a larger number of acoustical ports (due to the different propagation modes) and a larger electrical port number (due to the fact that electrodes in each track can be connected to different electrical ports). The unit cell P matrix has to take into account the mode shape of each mode. The program allows the use of an arbitrary number of electrical ports to analyze a general structure. For example, the analysis of the three transducer structure includes four electrical ports (two for the center transducer and one for each of the outer transducers).

The datas of Biryukov et al. [9] for the classical 2 pole resonator was used to test the analysis program. Figure 8 shows the comparison of the analysis to the measurement for a cascade of two 2 pole filters in a 50 Ω system. The fit between analysis and measurement is very good.

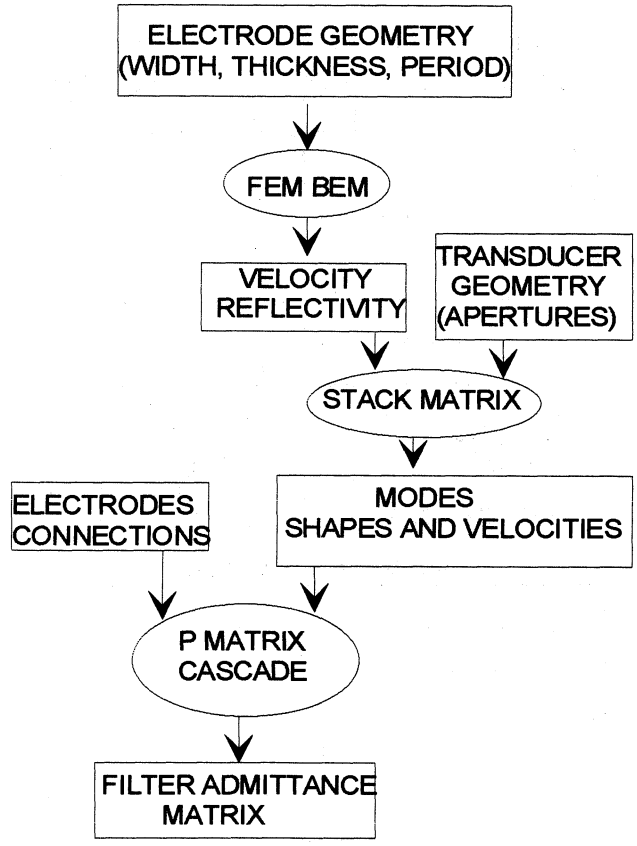


FIGURE 7 : PRINCIPLE OF THE FILTER ANALYSIS

4. RESULTS

4.1 . Geometry optimization

The first step when designing a filter is to determine the geometry of the structure. It was found that to obtain a symmetric frequency response, it is necessary to choose the acoustical widths of the three channels so that the coupling is equalized. This is obtained when, for each active area, the coupling to the two symmetric modes is nearly equal. This is true when the three resonators have the same aperture.

4.2. Filter results

The three resonator structure was used to design a filter at 188 Mhz for mobile phone applications. The filter works on 500 Ω with balanced drive and fits in a 9,1x7,1 mm^2 package. Figure 9 shows measurements and simulations of the filter. The measured insertion loss is 5.5 dB. It has a 3 dB bandwidth of 240 kHz a 20 dB bandwidth of 500 kHz and a 40 dB bandwidth of 900 kHz. Simulation and measurement are close.

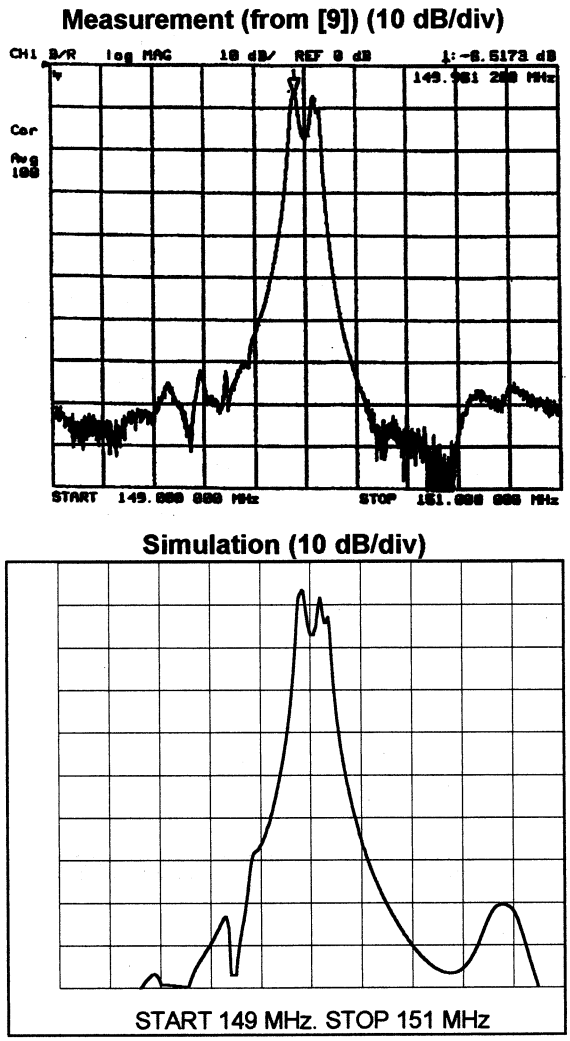
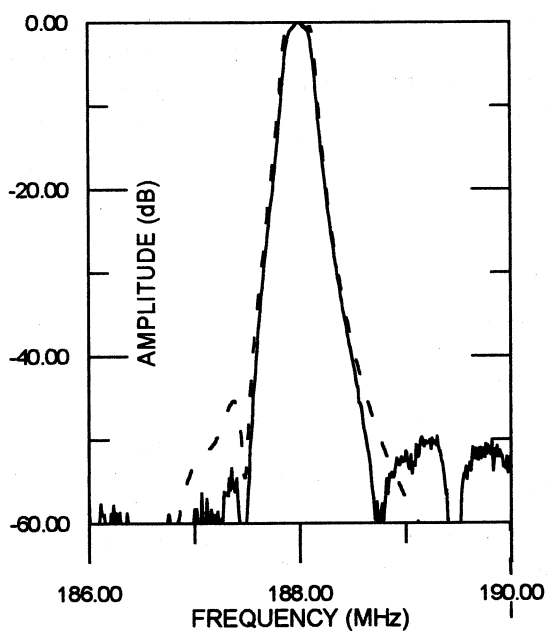


FIGURE 8 : COMPARISON BETWEEN OUR ANALYSIS AND [9] MEASUREMENT



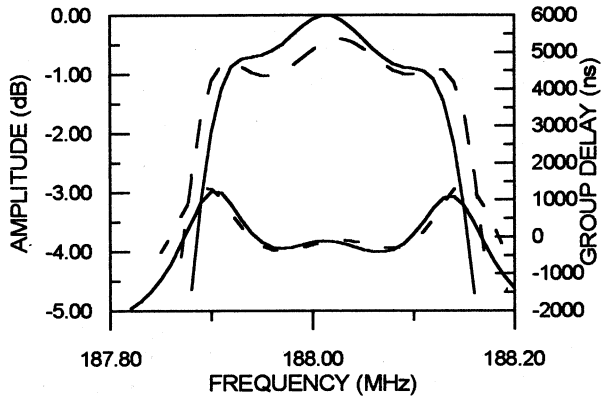


FIGURE 9 : MEASURE AND SIMULATION OF A 188 MHz 3 RESONATORS FILTER

Simulation (dashed line) : reference 5.7 dB
Measurement (full line) :reference 5.5 dB

5. CONCLUSION

A new balanced drive transversely coupled resonator filter has been presented. This structure allows for low impedances while keeping the electrical symmetry. A general analysis program for transverse coupling was developed and used to design a three resonator filter. Results are compatible with the system requirements as well as for the impedances and the frequency characteristics. Reasonable coherence between analysis and measurements were obtained demonstrating the ability of our analysis to be used as a design tool.

REFERENCES

[1] H. TIERSTEN and R. SMYTHE, Appl Phys Letters,28, 111-113 (1976)
[2] M. TANAKA et al., 38th Frequency Control Symposium, 1984, p 286-293
[3] C.B. SAW et al., US Patent 5,365,18, Nov. 94
[4] P. VENTURA,J. DESBOIS, L. BOYER, 1994 EFTF Proc.
[5] B. WALL, A.du HAMEL, 1996 IEEE Ultrasonics Symposium Proc., to be published.
[6] HAUS, J. Appl. Physics, Vol 48, No 12, Dec 77, pp 4955-4961
[7] D.P. MORGAN, S. RICHARDS, A STAPLES, 1996 IEEE US. proc, to be published
[8] W.D. HUNT et al., 1992 IEEE US proc., pp 45-50
[9] S. BIRYUKOV et al., 1995 IEEE UFFC Trans. vol 42, number 4, July 95
[10] D.P. CHEN et al., 1994 I994 IEEE US proc., pp 67-70
[11] C. RUPPEL et al., 1994 IEEE US proc., pp 313-324

SURFACE MOUNT HYBRID MODULES WITH SAW FILTERS FOR MOBILE TRANSCEIVERS

S.A.Dobershtein, A.V.Martynov, V.A.Malyukhov

ONIIP, Maslennikov str., 231, Omsk, 644063, Russia. Fax: (381-2) 57 90 38

At present for the packaging of the SAW filters the miniature SMD packages are widely used in order to combine the SAW filters with surface mount technology (SMT). At the same time for miniaturizing the mobile transceivers hybrid SAW modules (functional devices combining low loss SAW filters, amplifiers, mixers) are successfully used. It is promising to combine the two mentioned technologies [1]. This paper presents surface mount hybrid modules with SAW filters for 146-174, 440-470 MHz mobile transceivers; namely the receiver module (RXM) with RF amplifier (RFA), tramsmitter module (TXM) and receiver module (RXM) with voltage-controlled oscillator (VCO). For the modules we used 8.4x14.2x2.6 and 8.4x18.2x2.6 mm ceramic SMD packages. As the SAW filters for RXM with RFA we used previously developed self-matched SAW ring filters with the insertion loss of 1 dB on 128°YX and 64°YX $LiNbO_3$ [2]. The RXM with RFA contained two SAW filters: the first with low input/output impedances on 128°YX, the second with high input/output impedances on 64°YX, connected across RFA having only one bipolar transistor. Surface mount RXM's with RFA and the SAW filters have shown low amplitude ripple of 0.3 dB within the 2 dB bandwidth about 1.5%, 8 dB gain, suppression over 90 dB at ±10.7 and ±21 MHz offsets from the center frequency of 164.5 and 460 MHz respectively. Noise factor less than 2 dB. These RXM's were used for the local oscillator frequency and image frequency suppressions. We used VCO with SAW delay line and external phase shifter in the amplifier feedback loop [3]. For delay line we used previously developed low-loss leaky SAW ring filter with insertion loss of 1 dB having linear phase response with the phase shift of ±180° at 3-dB fractional bandwidth of 5% on 49°YX $LiNbO_3$ [2]. The hybrid module with VCO contained the SAW filter, amplifier, phase shifter. The amplifier of VCO contained a single bipolar transistor. The phase shifter with the miniature varactors and the leadless inductor [3] provided the phase shift about 100° when the control voltage varied from 0.5 to 12 volts. The optimization of the amplifier-phase shifter-SAW filter system was provided for achieving the maximal variable frequency range with specified maximal control voltage of 12 volts, current consumption of 5 mA and the carrier to noise ratio C/N no less than 70 dB (at 25 kHz carrier offset, 3 kHz bandwidth). Surface mount 164.5 and 460 MHz TXM's and 185, 480 MHz RXM's with VCO provided a variable frequency range of 1-2% with the control voltage varying from 0.5 to 12 volts. C/N was 80 dB . We developed new surface mount modules with SAW filters much smaller than the previous ones [1], and we shall use them in small mobile transceivers assembled using SMT.

REFERENCES

1. S.Doberstein et al. Proc. IEEE Ultrasonics Symp.,1995, pp.147-150.
2. S.Dobershtein, V.Malyukhov. Proc. IEEE Freq. Control Symp., 1994, pp.330-336.
3. T.E.Parker, G.Montress. IEEE Trans. on UFFC, v.35, 1988, pp.342-364.

A New Design of SAW Filters for Clock Recovery Application in Optical Communication Systems

Dong-Pei Chen, A. Vishwanathan, V. Narayanan, D. Mahoney, Bob Potter, and C.S. Lam
Vectron Technologies, Inc.
267 Lowell Road, Hudson, New Hampshire 03051, USA

Abstract

A new design of clock recovery SAW filter using a conventional λ/8 split-finger transducer in combination with a λ/4 finger resonant transducer has been developed. Such a new structure has the small chip size as a full resonant filter but retains many merits of a transversal filter- smooth passband, ease in phase adjustment, less sensitive to metallization thickness and mark to period ratio, etc. This paper presents the results of a 155.52 Mhz clock recovery SAW filter using this design.

1. Introduction

For many years, transversal SAW filters have been used widely in providing accurate clock extraction from the corrupted non-return to zero (NRZ) digital streams within the timing recovery circuitry of modern high speed optical communication systems from 50 Mhz to 2.5 Ghz[1~4].

Of the many available designs, SAW transversal filter is most popular because of its ease in design and manufacture. The absolute phase of conventional transversal filter can also be easily adjusted through varying the separation of the transducers. For clock recovery application, precise phase matching is needed prior to the decision circuit for retiming. This design, in general, has between 12 to 20 dB insertion loss. Quartz substrate is exclusively used because of its temperature stability for narrow band application.

Another popular design uses two long λ/4 finger resonant transducers which create strong resonant peak for filtering. Though this configuration features low-loss, small size, and wider finger feature size, it is considerably more difficult to design and manufacture because of its sensitivity to metallization thickness, mark to period ratio, and transducers separation which are inherent to resonant filter designs.

We have developed a new design of clock recovery SAW filter using a conventional λ/8 split-finger transducer in combination with a λ/4 finger resonant transducer. Such a new structure has the small chip size of the λ/4 finger resonant filter but retains some of the merits of a transversal filter- smooth passband, ease in phase adjustment, less sensitive to metallization thickness and mark to period ratio, etc. This paper presents the results of a 155.52 Mhz clock recovery SAW filter using this new structure.

2. Conventional Transversal SAW Filters

For clock extraction application in the timing recovery circuitry of modern high bit rate optical communication systems, SAW filters of 3-dB bandwidth Q of 400 to 1200 are most commonly used[1]. Since insertion loss of up to 20 dB is in general acceptable, conventional transversal SAW filters using quartz as the substrate material can easily offer the bandwidth with high temperature stability. If two identical transducers are used, the sidelobe rejection will be 26 dB which is more than enough for clock recovery applications. If needed, one of the transducers can be slightly weighted (e.g. raised-cosine apodization) to provide additional sidelobe rejection. Transversal SAW filters are easy to design and manufacture. However, they are in general quite lengthy because of the Q requirement.

As discussed before[1,5], the passband smoothness can impact the amount of jitter peaking of the clock recovered. The magnitude of the passband ripple depends on the amount of triple transit echo and reflectivity (primarily mechanical for aluminum transducer on weak coupling quartz). Since the insertion loss is high and matching is not needed, triple transit echo is in general not of concern. However, transducer finger reflection has to be suppressed. The most common way of suppressing finger reflection is to use λ/8 split-finger type transducers. This mechanical reflection can be canceled within every period of 4 fingers. This method limits the frequency of operation as λ/8 fingers are needed.

Later on, transversal filters using the group-type transducers with λ/4 fingers became popular[6,7]. The mechanical reflection can be canceled out within every four groups of fingers (Figure 1). The merit of such a structure is obvious since λ/8 fingers are no longer needed. This allows operation into higher frequency.

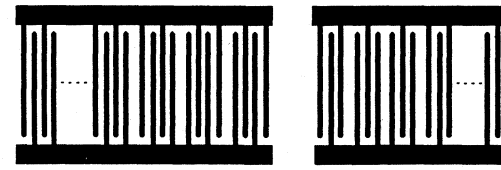


Figure 1. Transversal SAW Filter Using One 4-finger and One 3-finger Group-type Transducers.

To further ensure the suppression of mechanical reflection, fingers can also be buried since the mechanical impedance of aluminum is close to that of quartz[8]. One demerit of transversal filters using the group-type transducers is the existence of many spurious responses stemmed from sampling[9].

3. Resonant SAW Filters

Another class of clock recovery SAW filters was developed over a decade ago and has been used extensively in some of the products from Japan[10]. Clock recovery filters up to 2.5 Ghz using this structure are available[11]. The structure of this design is in fact a resonant type with two closely aligned long λ/4 finger transducers (Figure 2). Using the surface λ/4 fingers, the transducers generate a strong resonant peak which serves as the filter passband.

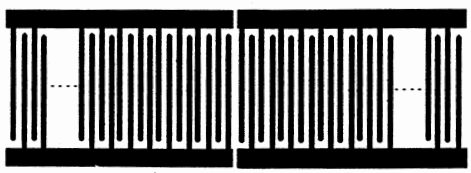


Figure 2. Resonant SAW Filter

Figure 3 compares the frequency responses of two SAW filters using the conventional transversal filter design and the resonant filter design respectively. The latter offers lower insertion loss (<10 dB) and smaller chip length. However, an only slightly attenuated spike (10~15 dB) appears in the high frequency side of the passband. This spike will appear on the low frequency side if the separation of the two transducers increased or decreased by λ/4.

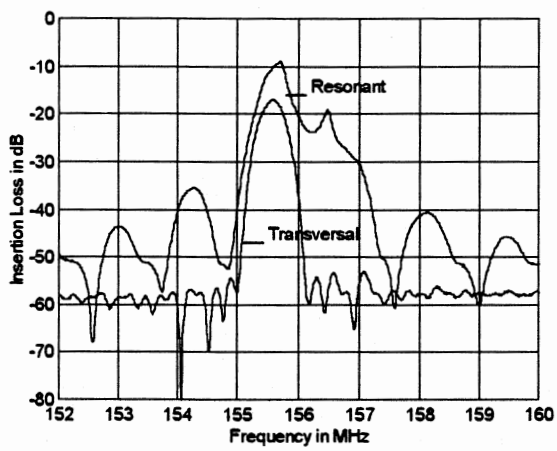


Figure 3. Frequency Responses of the Conventional Transversal and the Resonant SAW Filters

Since the structure is inherently resonant type, the frequency response is very sensitive to the metallization thickness and mark to period ratio. The passband is in general asymmetric (Figure 4) and is unfavorably sensitive to the separation of the transducers[12]. In addition, it is difficult to achieve low Q.

Phase matching is especially critical in timing recovery unit applications where additional phase shifter is not implemented due to the compactness of the devices[3,13]. Transversal filter allows continuous adjustment of phase through changing the separation between the transducers. For example, an increase or decrease of λ/8 in separation corresponds to about 45° in phase change. Though resonator filter has comparable phase slope, it's not easy to adjust the absolute phase since the frequency response is very sensitive to the separation of the transducers. Figure 5 compares the transmission phases of two SAW filters using the conventional transversal filter design and the resonant filter respectively.

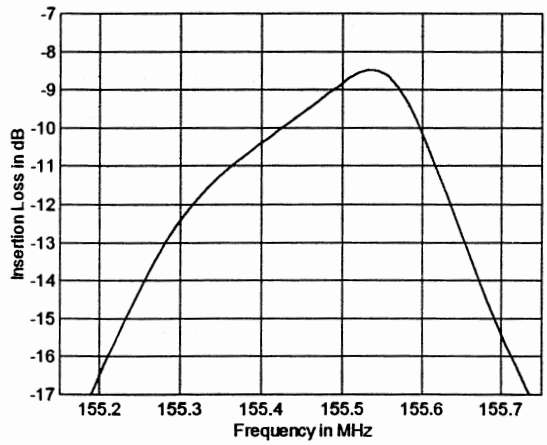


Figure 4. Passband of Resonant SAW Filter

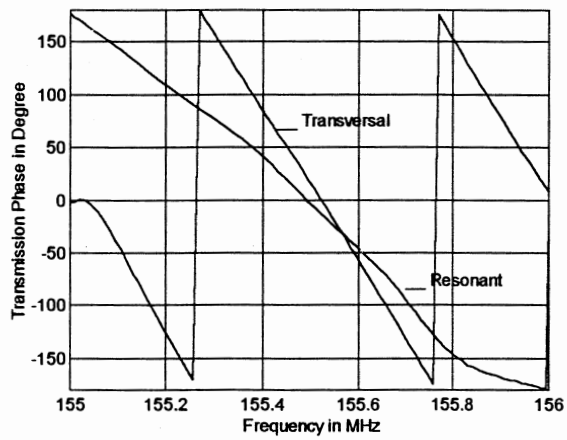


Figure 5. Transmission Phases of the Conventional Transversal and the Resonant SAW Filters

4. New Structure Using a Conventional λ/8 Split-finger Transducer with a λ/2 Resonant Transducer.

The lengthy chip size of transversal filter was not of concern in the past. Nowadays, new applications like SAW-based timing recovery units and optical receivers require small packaged SAW filters. To achieve the small chip size, a resonant filter structure is most likely needed. However, the problems as described in the previous section are of concern. We have developed a new 155.52 Mhz clock recovery SAW filter which can be encapsulated into a 9mmx7mm leadless chip carrier (LCC). The new structure uses a conventional λ/8 split-finger transducer in combination with a λ/4 finger resonant transducer (Figure 6). Such a structure retains the small chip size of a resonant filter. However, it has the merits of the transversal filter- smooth passband, ease in phase

adjustment, less sensitive to metallization thickness and ratio, etc. Figures 7 & 8 compare the frequency responses and transmission phases of two SAW filters using the conventional transversal filter design and the resonant filter design respectively.

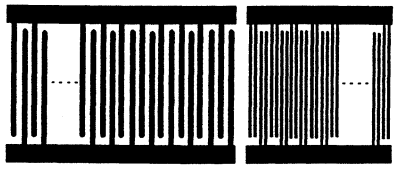


Figure 6. New Structure of SAW Filter for Clock Recovery Application

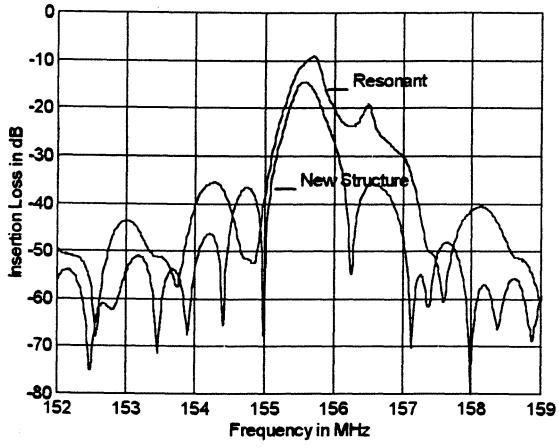


Figure 7. Frequency Responses of the New Structure and the Resonant SAW Filters

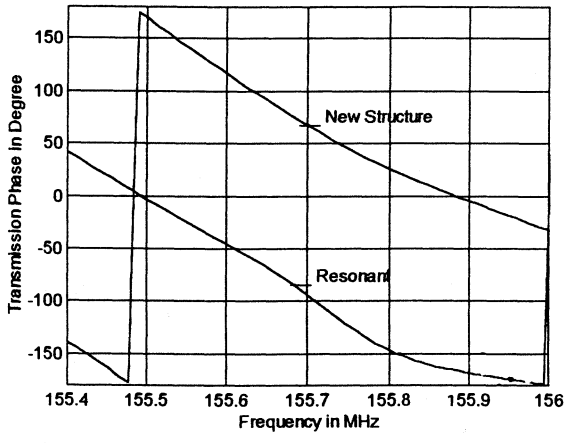


Figure 8. Transmissions Phases of the New Structure and the Resonant SAW Filters

In addition to the smaller chip size as compared with the transversal filter of the same Q, the new structure also has lower insertion loss because one of the transducers is a resonant one. Note that the sidelobe rejection of the new structure in general is not as much as that of the transversal filter (Figure 9). However, the rejection can be improved if the split-finger transducer is weighted. The passbands compare well (Figure 10).

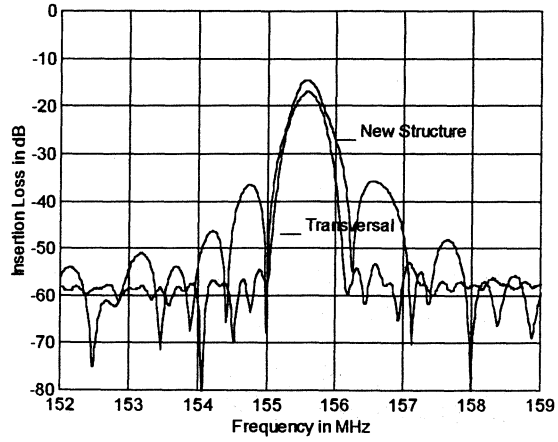


Figure 9. Frequency Responses of the New Structure and the Transversal SAW Filters

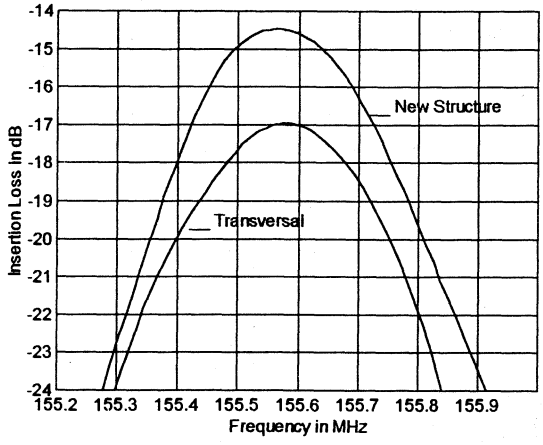


Figure 10. Passbands of the New Structure and the Transversal SAW Filters

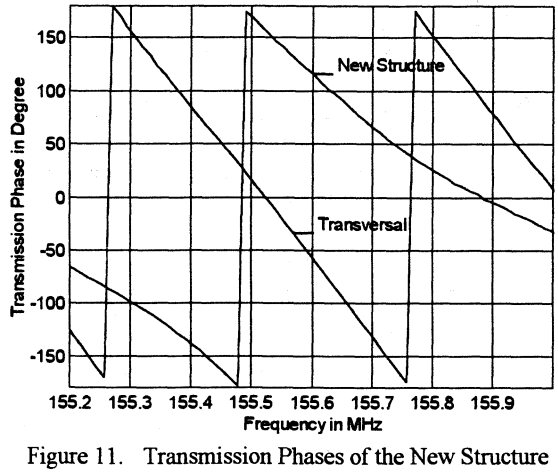


Figure 11. Transmission Phases of the New Structure and the Transversal SAW Filters

The transmission phase of the new structure compares well with that of the transversal filter (Figure 11). Figure 12

compares the ring time of the two designs. Though the chip size of the new structure in this example is about half of that of the transversal filter, it has the same ring time to ensure clock recovery in case of long transitionless signal string[1].

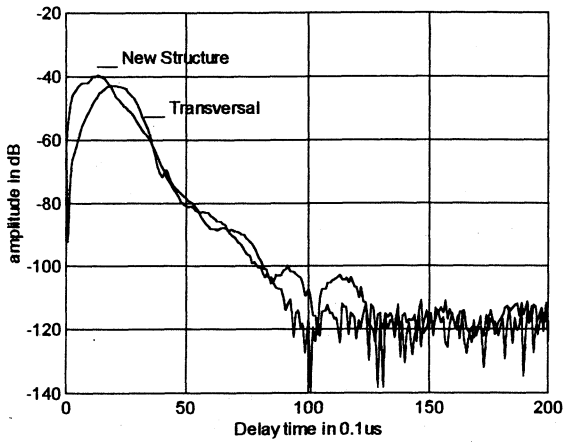


Figure 12. Ring Time of the New Structure and the Transversal SAW Filters

It is of interested to note that, in addition to the fundamental mode, the split-finger transducer supports the third harmonic also. By using the third harmonic, the finger width of the split-finger transducer would be about 33% wider than that of the λ/4 finger resonant transducer (Figure 13). We have demonstrated this with a new 622.08 Mhz clock recovery filter. Results will be reported later.

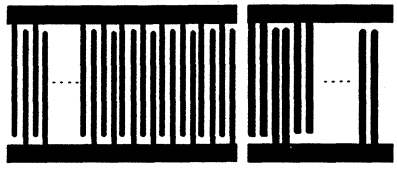


Figure 13. New Structure of SAW Filter Using the 3rd Harmonic of the Split-finger Transducer

5. Conclusion

We have developed a new 155.52 Mhz SAW filter for clock recovery application in optical communication systems by using a conventional λ/8 split-finger transducer in combination with a λ/4 finger resonant transducer. Such a structure has the small chip size as a full resonant filter but retains many of the merits of a transversal filter- smooth passband, ease in phase adjustment, less sensitive to metallization thickness and mark to period ratio, etc.

In addition, the λ/8 split-finger transducer can also be operated at the 3rd harmonic. For a 622.08 Mhz clock recovery SAW filter, submicron finger width will not be needed for both the split-finger transducer and the λ/4 finger resonant transducer.

References

1. "Optical Fiber Repeatered Transmission Systems Utilizing SAW Filters," R.L. Rosenberg et al., Proc. 1982 Ultrasonics Symp., pp. 238-246.

2. Data Book, Vectron Technologies, Inc., Hudson, New Hampshire, USA, 1996.

3. "SAW-Based Frequency Control Products for Modern Telecommunication Systems," C.S. Lam et al., RF design, June 1996, pp. 76-78.

4. "SAW Timing-Extraction Filter and Investigation of Submicron Process Technology for Gb/s Optical Communications System," K. Asai et al., Proc. 1995 Ultrasonics Symp., pp. 131-135.

5. "Analysis of Jitter Peaking Effects in Digital Long-Haul Transmission Systems using SAW-Filter Retiming," D.A. Fishman et al., IEEE Tran. Communications, Vol. Com-33, No. 7, July 1985, pp. 654-664.

6. "Interdigital Electrode Configurations for Cancellation of Internal Reflection," (in Japanese), H. Sato et al., Reports of 1975 Spring Meeting, Acoustical Society of Japan, May 1975, pp. 305-506.

7. "800 Mhz SAW Timing Filter for Optical Fibre Transmission System," J. Minowa et al., Elecronics Letters, 3rd Jan. 1980, Vol. 16, No. 1, pp. 35-36.

8. "Packaging and Reliability of SAW Filters," T.R. Meeker and W.R. Grise, 1983 Proc. Ultrasonics Symp., pp. 117-124.

9. "Simple Spectral Analysis of Single-Finger Transducer without Internal Reflection," (in Chinese), Chen Xiaobing & Cao Liang, International Workshop on Ultrasonics Application, Nanjing, China, Sept. 1996.

10. "SAW Filters and Resonators for Pubic Communication Systems," Y. Yamamoto et al., 1993 Proc. Ultrasonics Symp., pp. 95-103.

11. "SAW Filter for Timing Extraction," M. Kasagi et al., OKI Technical Review 141 Vol., 58, October 1991, pp. 31-36.

12. "A Study of Narrow Band Low-Loss SAW Filters," (in Japanese) T. Watanabe et al., 1982, private communication with Y. Yamamoto.

13. "Understanding Timing Recovery & Jitter in Digital Transmission Systems, Parts I & II," K.J. Bures, RF design, Octo. & Nov. 1992.

COMPENSATION OF WAVEGUIDING EFFECTS IN SURFACE ACOUSTIC WAVE FILTERS

Piotr Nagłowski, Elżbieta Dąbrowska, Bogumiła Niewczas

Institute of Electronic Materials Technology
Wólczyńska 133, 01- 919 Warsaw, Poland

ABSTRACT

A method for broadband compensation of waveguiding effects in SAW filters of "in-line" structure is presented. The method is suitable for filters supporting any number of propagating modes. According to the compensation method constrained optimization techniques are used to appropriately modify the apodization pattern. Theoretical and experimental compensation results are presented for TV-IF filters designed on 128° YX $LiNbO_3$ substrate.

Keywords: SAW filters, waveguiding effects, SAW diffraction.

1. INTRODUCTION

Numerous modern applications of SAW transversal filters require precise control of filter parameters like stopband rejection, passband amplitude ripple and transition bandwidth.
Various second-order effects degrade the parameters of SAW filters. These parasitic effects should be accurately predicted by the filter mathematical model and efficiently compensated for during the design process.
Interdigital transducers of narrow acoustic aperture (less than 20λ) are widely used in SAW filters to reduce the cost of the devices. SAW diffraction severely degrades the frequency responses of such filters, especially in the stopbands. Unfortunately, it is difficult to develop a mathematical model correctly predicting diffraction effects in a filter containing narrow aperture IDTs. Most published methods on diffraction compensation assume that the metallization pattern (IDT fingers and bus bars) has no influence on the SAW wavefronts once they have been excited. Thus the above methods ignore the slowing effect of the metal which is the cause of guiding of SAW energy. This assumption of a "free diffraction" model is particularly unjustified for strongly piezoelectric materials like e.g. 128° YX $LiNbO_3$.
The methods of waveguiding effects compensation were published previously (Refs.1,2). According to the methods it is assumed that the fundamental mode of the IDT guiding structure is the dominating one and the compensation concerns only this single mode. Since it is performed at the center frequency the described methods are suitable only for narrow-band SAW transversal filters.
To meet stringent specifications wideband compensation of parasitic effects is often necessary. Since it seems difficult to ensure absolute domination of the fundamental mode over a broad band of frequencies other propagating modes should be included in the mathematical model used for the compensation. This paper proposes a wide band compensation method which can be used for filters supporting any number of propagating modes.

2. MATHEMATICAL MODEL OF THE SAW FILTER

A typical application circuit containing a SAW filter is presented in Fig.1.

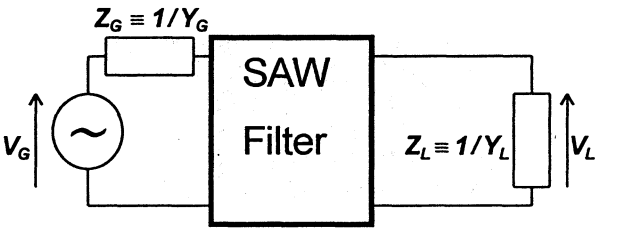


Fig.1. Circuit containing a SAW filter

The filter transmittance D can be expressed using the elements of filter's admittance matrix [Y], and terminations Y_G and Y_L (Ref.5). Neglecting the triple transit signal we obtain:

$$D \equiv V_L / V_G = F_{in} \cdot Y_{21} \cdot F_{out} \qquad (1)$$

where $F_{in} = Y_G/(Y_{11} + Y_G)$, $F_{out} = -1/(Y_{22} + Y_L)$.

The analyzed structure of a SAW filter is shown in Fig.2. It contains a section of earthed dummy fingers between the transducers (Ref.4), which forms the electromagnetic shield and makes the guide almost continuous. All the electrodes of the structure are of the same width (w) and the electrode pitch (p) is constant and identical in both IDTs and the shield. Double fingers are used to make the internal reflections off the electrodes negligible. The structure has slow (metallized, $a < |x| < b$), intermediate (semi-metallized, $|x| < a$) and fast (free surface, $|x| > b$) velocity regions.

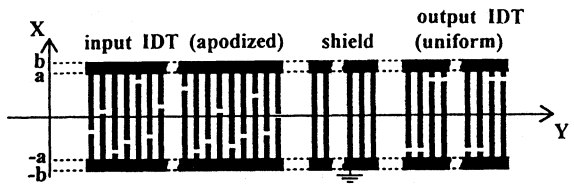


Fig.2. Analyzed interdigital structure and the coordinate system

We determined the modes propagating in the waveguide as described in Refs.3,4. The only difference in our approach is the use of the electric potential on the surface of the piezoelectric substrate to describe the modes instead of the components of particle displacement vector used in Refs. 3,4. The electric potential on the surface of the piezoelectric substrate caused by the m-th mode ($m = 1, ..., M$) propagating in the $+y$ direction is assumed to have the form:

$$\varphi_m (x, y) = f_m (x) \exp(-j k_m y) \qquad (2)$$

where k_m is the propagation constant and $f_m(x)$ describes the x-dependence of the m-th mode. The $\exp(j\omega t)$ dependence of each mode is assumed but omitted in all the formulas for clarity. In the finger region $f_m(x)$ takes the form of cos, sin, cosh or sinh function. In the bus bar region it has a sinusoidal form and in the free surface region exponentially decays with $|x| \rightarrow \infty$.
The dispersion relation of the waveguide is obtained by demanding the continuity of $f_m(x)$ and its first derivative.

The charge densities on the electrodes at and near the IDT ends are distorted. These distortions are known as end effects. They affect the frequency response of the IDT especially in the case of relatively short transducers. For the in-line filter structure shown in Fig.2 the apodized transducer is usually much longer than the uniform one (this is normally true e.g. for TV-IF filters). Therefore end effects in the uniform IDT are expected to cause more severe distortion of the filter frequency response than analogous effects in the apodized IDT. Exact modelling of end effects requires the computation of charge distribution on the IDT fingers. For long apodized transducers this computation is both time and computer memory consuming. For the above reasons end effects are taken into account only for the uniform IDT in this paper. Inter-electrode gaps are assumed to be elementary SAW sources in the apodized transducer. In the uniform IDT each electrode is considered to be an elementary SAW source. Charge densities on all the fingers of this IDT are determined numerically.
Let a voltage V be applied to the j-th IDT (j =1 corresponds to the apodized IDT, j = 2 to the uniform one) with the other IDT short-circuited. Figs.3,4 show the location of n-th elementary SAW source within the apodized IDT ($n = 1, ..., NS_1$; where NS_1 is the number of inter-electrode gaps in the IDT) and within the uniform IDT ($n =1, ..., NS_2$; where NS_2 denotes the number of fingers in the IDT) respectively.

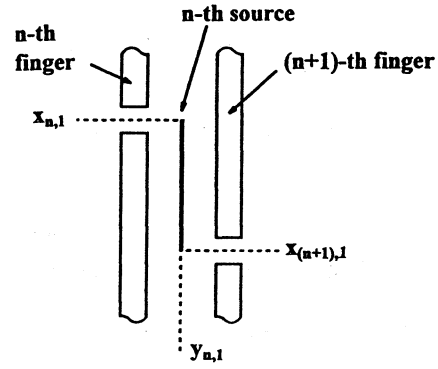


Fig.3. Location of SAW source within the apodized IDT

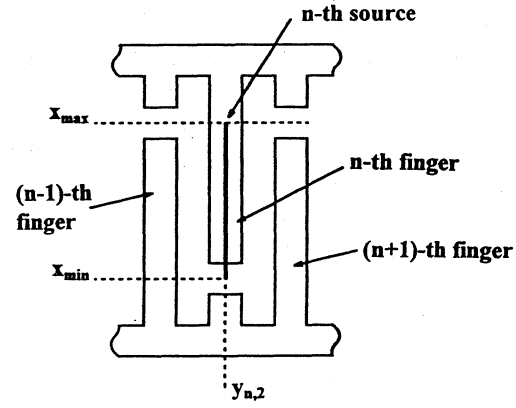


Fig.4. Location of SAW source within the uniform IDT

The SAW potential $\varphi_{nj}(x,y)$ generated by the n-th source is assumed to be a linear combination of all the propagating modes of the structure. To evaluate this potential at the exact source location the assumption of the source homogeneity throughout its length was made and quasi-static approximation was used. Infinite number of fingers in the apodized IDT was assumed (Ref.7) whereas the moment method (Ref.8) was applied to evaluate the charge distribution on the fingers of the uniform transducer. Application of the Galerkin method (as in Refs.3,4) yielded:

$$\varphi_{n1} (x,y) = j \Gamma_s \overline{\sigma_g}(k) \operatorname{sgn} (y-y_{n1}) V \sum_{m=1}^{M} B_{mn1} \cdot f_m (x) \cdot \exp (-j k_m |y-y_{n1}|) \qquad (3)$$

where:

$\Gamma_s = \frac{1}{\varepsilon_s(\infty)} \frac{\Delta v}{v_0} \left(1 + \frac{3}{2} \frac{\Delta v}{v_0}\right)$ (see Ref.9),

v_0 - free surface SAW velocity, v_m - metallized surface SAW velocity, $\Delta v = v_0 - v_m$, $\varepsilon_s(s)$ - effective permittivity function, $k = \omega / v_0$, $\bar{\sigma}_g(k)$ - gap element factor (see Ref.7),

$$B_{mn1} = \int_{x_{n,1}}^{x_{n+1,1}} f_m(x)\,dx \;/\; \int_{-\infty}^{\infty} f_m^2(x)\,dx ;$$

$$\varphi_{n2}(x,y) = C_n(y) \cdot V \cdot \sum_{m=1}^{M} B_{m2} \cdot f_m(x) \cdot \exp(-j k_m |y - y_{n2}|)$$

where: (4)

$$C_n(y) = \begin{cases} j\Gamma_s \bar{\sigma}_n^*(k) & \text{for } y > y_{n2} \\ j\Gamma_s \bar{\sigma}_n(k) & \text{for } y < y_{n2} \end{cases},$$

$\sigma_n(y)$ - charge density on the n-th finger when $1V$ potential difference is applied to the IDT,

$$\bar{\sigma}_n(k) = \int_{y_i - w/2}^{y_i + w/2} \sigma_n(y - y_{n2}) \exp(-j k y)\,dy,$$

$\bar{\sigma}_n^*(k)$ - conjugate of $\bar{\sigma}_n(k)$,

$$B_{m2} = \int_{x_{min}}^{x_{max}} f_m(x)\,dx \;/\; \int_{-\infty}^{\infty} f_m^2(x)\,dx .$$

The admittance matrix [Y] can be expressed as a sum of contributions from the IDT static capacitances and a motional part [Y^{mot}] due to the current induced piezoelectrically by the SAW (Refs.3,4). Applying the reciprocity relation between SAW launching and reception (Ref.7) and using Eqns.3,4 we obtained:

$$Y_{11}^{mot} = -\omega\Gamma_s \bar{\sigma}_g^2(k) \sum_{m=1}^{M} \int_{-\infty}^{\infty} f_m^2(x)dx \cdot \sum_{n=1}^{NS_1} \sum_{q=1}^{NS_1} B_{mn1} B_{mq1} \cdot \exp(-j k_m |y_{n1} - y_{q1}|) \quad (5)$$

$$Y_{22}^{mot} = \omega\Gamma_s \cdot \sum_{m=1}^{M} B_{m2}^2 \int_{-\infty}^{\infty} f_m^2(x)\,dx \cdot \sum_{n=1}^{NS_2} \sum_{q=1}^{NS_2} \cdot \cdot \bar{\sigma}_{max(n,q)}(k)\, \bar{\sigma}^*_{min(n,q)}(k) \cdot \exp(-j k_m |y_{n2} - y_{q2}|) \quad (6)$$

$$Y_{21}^{mot} = -\omega\Gamma_s \bar{\sigma}_g(k) \sum_{m=1}^{M} \int_{-\infty}^{\infty} f_m^2(x)\,dx\, B_{m2} \cdot \cdot \sum_{n=1}^{NS_1} \sum_{q=1}^{NS_2} B_{mn1}\, \bar{\sigma}_q(k) \cdot \exp[-j k_m (y_{q2} - y_{n1})] \quad (7)$$

Static capacitance of the apodized IDT can be computed from quasi-static theory assuming infinite number of the IDT electrodes (Refs.10,11). We determined the static capacitance of the uniform IDT by using the moment method to evaluate the charge distribution on its electrodes when $1V$ potential difference is applied to the IDT.

3. WAVEGUIDING EFFECTS COMPENSATION

To compensate for waveguiding effects we used the algorithm, outline of which was presented in Ref.5. Let's introduce the following notation:

$f_1, f_2, \ldots, f_L$ - dense grid of frequencies,

N_1 - number of apodized IDT electrodes,

$\underline{x} \equiv [x_1, x_2, \ldots, x_{N_1}]$ - vector of gap positions in apodized IDT,

$D(j\omega, \underline{x})$ - filter frequency response computed according to Eqn.1,

$D_k(\underline{x}) \equiv D(2\pi j f_k, \underline{x})$ - complex amplitude of frequency response for $\omega = \omega_k = 2\pi f_k$,

$C_{ki}(\underline{x}) \equiv \left. \frac{\partial D_k(\underline{X})}{\partial X_i} \right|_{\underline{X} = \underline{x}}$ - response sensitivity to i-th gap position change for $\omega = \omega_k = 2\pi f_k$.

For most practical cases $|Y_G| >> |Y_{11}|$ and the following approximation is justified:

$$C_{ki}(\underline{x}) = F_{in}(f_k, \underline{x}) \cdot F_{out}(f_k) \cdot \left. \frac{\partial Y_{21}(f_k, \underline{X})}{\partial X_i} \right|_{\underline{X} = \underline{x}} \quad (8)$$

It is assumed that before compensation the initial gap positions in the apodized IDT ($\underline{x}^\circ$), the geometry of the uniform IDT and the values of desired frequency response (I_k) on the frequency grid are obtained.

The compensation algorithm has the following form:

1. Let $\underline{x} := \underline{x}^\circ$,
2. Compute $D_k(\underline{x})$ and $C_{ki}(\underline{x})$ for $1 \le k \le L$; $1 \le i \le N_1$,
3. If the response is good enough then stop, or if not, find a complex number $\alpha = \alpha_o$ minimizing the expression: $\sum_{k=1}^{L} w_k \left| D_k(\underline{x}) - \alpha I_k \right|^2$ (w_k are positive weights),
4. Find the gap position displacement vector: $\Delta\underline{x} = \Delta\underline{x}^\circ$ minimizing the expression:
   $\sum_{k=1}^{L} w_k \left| D_k(\underline{x}) + \sum_{i=1}^{N_1} C_{ki}(\underline{x})\, \Delta x_i - \alpha_o I_k \right|^2$
   under constraints:
   $x_{i\,min} \le x_i + \Delta x_i \le x_{i\,max}$ for $1 \le i \le N_1$
   ($x_{i\,min}$ and $x_{i\,max}$ are chosen by the designer, see the design example in Ref.5),
5. $\underline{x} := \underline{x} + \Delta\underline{x}^\circ$, go to 2.

To minimize the quadratic function in step no.4 the gradient projection technique combined with conjugate gradient method is used. The $C_{ki}(\underline{x})$ coefficients are determined analytically from Eqn.8. In step no.2 $D_k(\underline{x})$ and $C_{ki}(\underline{x})$ are computed on the dense grid of frequencies which requires the amount of computer time roughly proportional to

$NS_1 + NS_2$. This time is much longer (proportional to $NS_1 \cdot NS_2$) for the "free diffraction" SAW filter model using parabolic approximation.

4. EXPERIMENTAL VERIFICATION

The developed compensation method was used to design many TV-IF SAW filters on 128° YX $LiNbO_3$ substrate. In the examples presented below the source impedance of 50Ω and the load impedance of 2kΩ in parallel with 3pF were assumed. Amplitude distortions resulting from the output IDT response, load impedance and gap element factor were compensated in the initial stage of the design. The first example is the video filter for quasi/parallel sound applications. It consists of the apodized IDT, the uniform one and the shield with 269, 30 and 14 electrodes respectively. The acoustic aperture is approx. 15λ. Fig.5 illustrates considerable differences between the desired ideal response and the response taking into account waveguiding effects computed before the compensation.

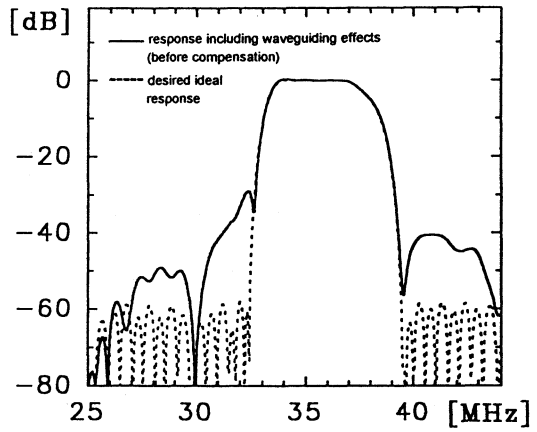


Fig.5. Theoretical amplitude responses of the TV-IF video filter

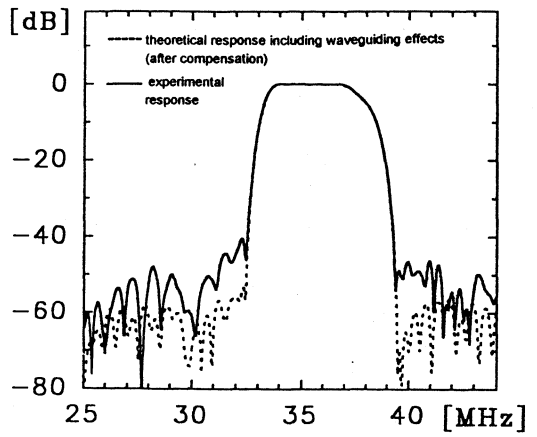


Fig.6. Theoretical and experimental amplitude responses of the TV-IF video filter

These differences are especially large in the low frequency stopband region adjacent to the passband. Since the zeroes (traps) of the uniform IDT appear approx. at 30MHz and 39.5MHz this transducer does not significantly contribute to the filter rejection in the above frequency region. The rejection in the region is practically determined by the input IDT apodization pattern. After 5 iterations of the compensation algorithm the theoretical amplitude response very close to the desired one was obtained. This response and the experimental one are shown in Fig.6. The stopband rejection of the experimental response is much better than the theoretical one corresponding to the uncompensated design. The discrepancies between the curves shown in Fig.6 are caused by many simplifications in the mathematical model of the filter. Some of these simplifications are: modelling waveguide modes by a scalar potential, neglecting the evanescent modes, parabolic approximation of SAW velocity anisotropy, use of quasi-static approximation, ignoring bulk waves and direct breakthrough signal.

Another example concerns two sound filters for parallel sound applications. In one of the filters (filter A) waveguiding effects were compensated according to the described method. In filter B only diffraction (and end effects in the output IDT) were compensated for. Parabolic "free diffraction" model was assumed for this filter and the compensation algorithm of Ref.5 was used. In both filters the acoustic apertures, uniform IDTs and the distances between the transducers are identical. The initial apodization patterns ($\underline{x}°$) and the desired responses (I_k) of the filters were also the same. In filter B the electromagnetic shield in the form of a single relatively wide metallic strip was used (instead of a 13-electrode array applied in filter A) to impede the propagation of waveguiding modes. The layout of filter B is shown in Fig.7.

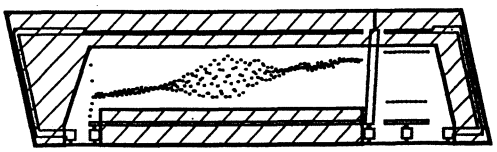


Fig.7. Layout of the TV-IF audio filter with compensated diffraction (filter B)

Waveguiding effects in this filter were suppressed by the application of SAW damping medium (epoxy resin) to IDT bus bars and adjacent areas (Ref.6). Screen printing techniques were used to apply the epoxy resin onto the substrate surface. The apodized IDTs in the filters contain 263 electrodes

and the uniform ones 38 electrodes. Acoustic aperture is approx. 11.5λ but the apodization pattern extent in the input IDTs is only approx. 5λ since considerable areas had to be reserved for acoustic absorber in filter B. The filters' theoretical amplitude response including waveguiding effects (before compensation) is shown in Fig.8.

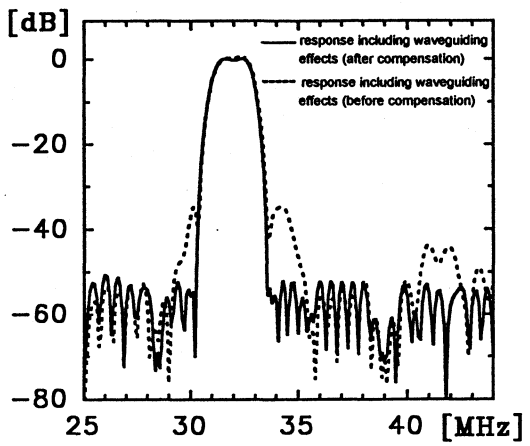


Fig.8. Theoretical amplitude response of the TV-IF audio filter with compensated waveguiding effects (filter A)

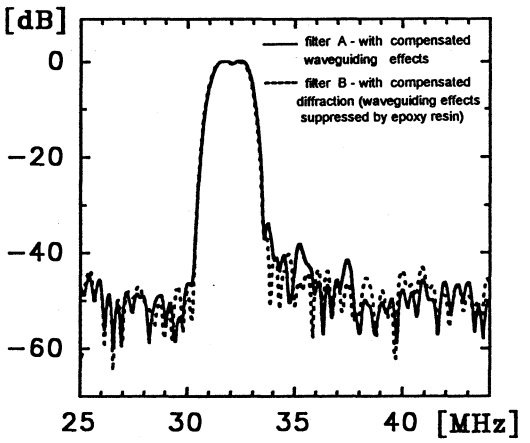


Fig.9. Experimental amplitude responses of the TV-IF audio filters

The same figure also illustrates the theoretical response obtained for filter A after 7 iterations of the compensation algorithm. This response is almost identical to both the desired ideal response and the theoretical response of filter B after compensation (computed using "free diffraction" model). The uncompensated response shown in Fig.8 exhibits two prominent sidelobes on both sides of the passband. Fig.9. illustrates experimental responses of the filters. Successful compensation of the left sidelobe is clearly observed for both designs whereas the other sidelobe was suppressed more efficiently in filter B. Probable cause are the simplifications adopted in the modelling of waveguiding effects. On the other hand the presented waveguiding effects compensation method gives quite satisfactory results. Contrary to the procedure used to develop filter B it does not require the precise application of SAW dampening medium onto the substrate surface. The method also allows to obtain smaller filter chips since there is no need to reserve IDT areas to be covered by SAW absorber.

5. CONCLUSION

A wideband waveguiding effects compensation method suitable for in-line SAW filters containing one apodized IDT and one uniform IDT has been developed. It was shown that the method makes it possible to compensate efficiently for waveguiding effects in narrow aperture SAW filters designed on 128° YX $LiNbO_3$ substrate. In the investigated case of the TV-IF sound filter the described method produced results almost as good as those obtained by the precise application of SAW absorbing medium over IDT bus bars and adjacent areas.

ACKNOWLEDGEMENTS

This work was partly supported by The State Research Project No. PBZ 12-05. The authors would like to express their gratitude to Mr. Marian Teodorczyk for the fabrication of the SAW devices and to Mrs. Jolanta Zawistowska for filter measurements. Thanks are also due to Mrs. Marianna Baranowska for typing of the manuscript.

REFERENCES

1. S.A. Wilkus et al., Proc. IEEE Ultrasonics Symposium 1985, pp. 43-47.
2. Y. Yamamoto and M. Morimoto, Proc. IEEE Ultrasonics Symposium 1985, pp. 48-52.
3. G. Clark and R.F. Milsom, Philips Journal of Research, vol.41, No.3, 1986, pp. 247-267.
4. G. Clark et al., Proc. IEEE Ultrasonics Symposium 1985, pp. 26-31.
5. P.J. Nagłowski, Proc. IEEE Ultrasonics Symposium 1989, pp.125 -128.
6. W. Faber, U.S. Patent No. 4,672,338.
7. D.P. Morgan, "Surface-Wave Devices for Signal Processing", Elsevier, 1985.
8. E.B. Savage, Ph.D. Dissertation, University of California, Santa Barbara, 1980.
9. C.S. Hartmann et al., Proc. IEEE Ultrasonics Symposium 1987, pp.161-167.
10. M. Politi et al., IEEE Trans. on UFFC, vol.35, No.6, Nov.1988, pp. 708-710.
11. P. Nagłowski, Ph.D. Dissertation, Warsaw Technical University, 1991, (in Polish).

COLD CESIUM ATOMS

Chairman: *A. Bauch*

A CONTINUOUS BEAM OF COLD CESIUM ATOMS FOR ATOMIC STANDARDS

Constance Valentin*, Eve Aucouturier*, Emmanuel Guillot*, Pierre Petit*, Stefan Weyers† and Noël Dimarcq*

* Laboratoire de l'Horloge Atomique - Unité propre du C.N.R.S., Bâtiment 220 - Université Paris-Sud, F-91 405 ORSAY Cedex - FRANCE

† present address : Physikalisch-Technische Bundesanstalt, Lab. 4.41, Bundesallee 100, D-38116 Braunschweig, GERMANY

1.ABSTRACT

Since last year, we have generated a continuous beam of cold Cesium atoms. This continuous beam is extracted from a two-dimensional Magneto-Optical Trap (2D-MOT) in a standard vapour cell. It has been generated by launching the atoms from the 2D-MOT either with a 1D moving optical molasses or a static magnetic field (magneto-optical train) in the vertical direction. The results will be presented and analysed. In order to understand these results, we have proposed some theoretical models : a good agreement has been shown for the launching with a moving molasses. In this case mean drift velocities between 1.9 to 5 m/s and up to 10^6 atoms/s in this beam can be obtained. Launching by a static magnetic field was similary efficient but exhibits a double-peaked structure in the velocity distribution of the launched atoms.
These experiments are a burning issue for the development of a new design for future atomic clocks.

Keywords : cesium atoms, atomic frequency standards, laser cooling and trapping.

2.INTRODUCTION

The advantages of using cold atoms for primary frequency standards is now well admitted (Ref.1). At the present time, most of the research on frequency stantards are based on pulsed operation. A continuous beam of cold atoms exhibits some advantages. First, the average atomic density during the ballistic flight would be reduced (typically by a factor of 100). As a result, the collisional shift (Ref.1) would be reduced by the same proportion. Secondly, a continuous beam could improve the frequency short-term stability, which is in the case of a pulsed operation limited due to the intermittent generation of an error signal in the servo-loop that controls the local oscillator (Ref.2). One may also use the properties of cold atoms to shorten the interrogation cavity size, while keeping the atomic clock stability to a level similar to the best cesium thermal beam frequency standards. It will lead to a compact and transportable frequency standard.

We expound here the experimental set-up for trapping the atoms in an anisotropic 2D-MOT. We consider launching the atoms using two methods : the moving molasses technique and the magneto-optical train. We present the first experimental results obtained with a continuous beam of cold atoms. We finally present the theoretical calcu-

lations of mean Doppler and sub-Doppler forces applied to the atoms in order to predict the mean velocity at which the atoms are launched.

3.EXPERIMENTAL SET-UP

In an usual MOT trapping and cooling of atoms from a vapour cell result from the application of three pairs of counterpropagating light waves emitted by a laser locked a few MHz below a "cycling" atomic transition (for Cs : $6^2S_{1/2}$, $F = 4 \rightarrow 6^2P_{3/2}$, $F = 5$ transition) and of magnetic field gradients in the region of intersection of the laser beams, which have to be appropriatly polarized ($\sigma^+ - \sigma^-$ configuration). In general the three-dimensional magnetic field gradient along the three axes of the cooling laser beams is generated by one pair of symmetric magnetic field coils supplied with opposite currents (anti-Helmholtz configuration).

Deviating from this set-up in our scheme the trapping process in one direction (Oz) of the MOT is suppressed in order to enable the extraction of atoms in this direction. For this purpose, we designed a 2D-MOT device, in which the magnetic field gradients in the intersection region are produced by two pairs of anti-Helmholtz coils as shown in Fig.1. The field gradients created by each pair of anti-Helmholtz coils add in the two horizontal directions (Ox and Oy) and compensate in the Oz-direction, so that the magnetic field is quasi zero on the Oz axis. For practical reason, we used octogonal coils as show in Fig.1. Finally no trapping occurs along the vertical axis of the 2D-MOT and the cold atoms are allowed to leak from the intersection region of the laser beam in this direction, while axial cooling in a 1D optical molasses is still provided.

Because the radiation pressure would disturb the trajectory of atoms on the Oz axis, it is not possible to use a pair of vertical laser beams for our 2D-MOT. Instead of this we use two pairs of counterpropagating laser beams of 2.5 mm diameter which makes an angle of $\phi = 14^0$ with the vertical axis ("crossed cooling and launching beams") (Fig.1). The two beams intersect in the intersection region of the horizontal beams, so that the atoms can be cooled and launched along the Oz axis, but no light appears along the atomic path. A repumping laser tuned to the $6^2S_{1/2}$, $F = 3 \rightarrow 6^2P_{3/2}$, $F = 4$ cesium transition is superposed to the horizontal beams to avoid the leakage of atoms to the state $6^2S_{1/2}$, $F = 3$.

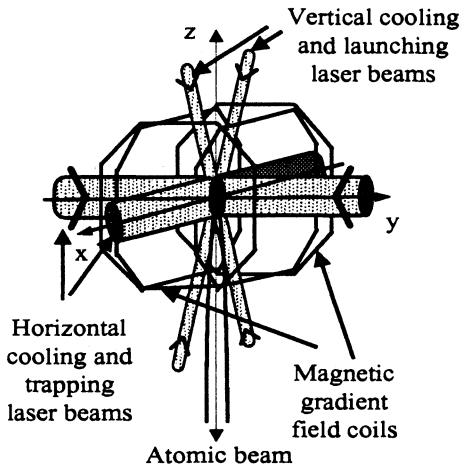


Figure 1: *Experimental set-up.*

We use two different techniques for launching the atoms. The first one consists in submitting the atoms to a moving laser wave generated by two counterpropagating progressive waves, whose frequencies are symmetrically shifted around the frequency of the horizontal laser beams by $\pm\Delta f$. The atoms are thus cooled in a frame which moves at the velocity

$$\bar{v}_d = \frac{\lambda \Delta f}{\cos \phi} \qquad (1)$$

where λ is the laser wavelength (Ref.3).

The frequency of the horizontal laser beams is ajusted to be -3Γ (Γ^{-1} is the excited state lifetime) below the F=4 $\rightarrow$ F=5 cycling transition, with the aid of acousto optical modulators. The crossed launching beams frequencies can be independently adjusted by means of acousto optical modulators around the frequency of the horizontal beams.

The second method consists in using a stationnary laser waves, whose counterpropagative components are respectivly circulary polarized σ^+ and σ^-, and applying a static homogeneous magnetic field along the Oz axis. In this case the magnetic field along the σ^+ (σ^-) beam has the same effect on the evolution of an atom as lowering (raising) the optical frequency by the Larmor frequency. In our experiment the stationnary laser wave is provided by the crossed cooling beams tuned to the same frequency as the horizontal beams. The magnetic

field is produced by a pair of symmetrical magnetic field coils (not shown on Fig.1) supplied with equal currents (Helmholtz configuration), so that a magnetic field amplitude B_z of up to 7 G could be obtained. The extraction mechanism is also studied at Observatoire de Neuchâtel (Switzerland) by Pierre Thomann's group (Ref.4).

4.EXPERIMENTAL RESULTS

After the trapping and cooling lasers have been switched on, we trap up to 4×10^6 atoms, whose fluorescence is detected by a calibrated photodiode. If one of the two described launching methods is applied, the atoms are extracted from the 2D-MOT with a drift velocity v_d and fall under the influence of gravity. They may be detected in a detection region situated 30 cm below the trap center (Fig.2), where the atoms cross a transverse standing laser wave tuned to the $F = 4 \rightarrow F = 5$ transition.
The experimental curves were obtained using the procedure described below. For measuring the continuous beam properties, the atoms cross 7 cm below the trap center a transverse "pushing" laser beam, to which a part of the repumping laser is superimposed. To begin with, the pushing laser is injection locked by the laser diode which provides the horizontal cooling and trapping beams (-3Γ below the $F = 4 \rightarrow F = 5$ transition) and it efficently pushes the atoms away. At $t_f = 0$ we quickly ($\ll 1\ ms$) detune the pushing laser by a current jump, so that the atoms arrive in the detection zone after a falling time t_f (Fig.2).
Fig.3 shows a typical signal for launching with a 1D moving optical molasses with $\Delta f = 3MHz$. The fluorescence is supposed to be proportional to the number of atoms N arriving in the detection zone at a certain instant t_f. The arrival time of the atoms in the detection region depends on their inital velocity distribution in the 2D-MOT along the Oz axis. The analysis of the launching process gives for the velocity distribution f(v) at the level of the pushing beam:

$$f(v) = \left(\frac{g}{2} + \frac{H}{t_f^2}\right)^{-1} \times \frac{dN(t_f)}{dt_f} \qquad (2)$$

where g is the gravity acceleration, H the distance between the pushing beam and the detection beam and $N(t_f)$ the number of detected atoms at a certain instant t_f.

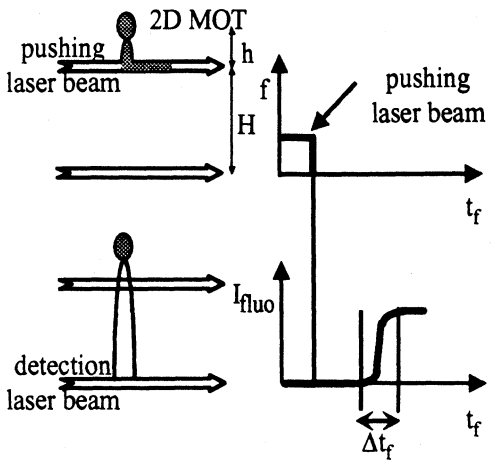


Figure 2: *Measurement of the atomic axial velocity distribution by the time of flight technique.*

In order to get the initial velocity distribution $f(v_d)$ at the trap level, v has to be replaced by $v_d = -(v^2 - 2hg)^{1/2}$ taking into account the falling distance h between the trap center and the pushing beam. The such derived velocity distributions are shown for the two launching methods in Figs.4a and 4b.
For launching with a 1D moving optical molasses the initial velocity distributions are well fitted by a gaussian curve for detunings Δf ranging between 1 and 4 MHz. In this range the temperatures deduced from the velocity distribution similar within the error and are on the order of 1 ± 0.5 mK. For largest detunings Δf the curve shapes become non-gaussian. We attribute this to an increasing misalignment of the crossed launching beams for increasing detunings due to the deflexion by the acousto optical modulators. In this case the 1D moving molasses does not cool the atoms properly. As we measure without any launching a temperature of 45 μK for the atomic cloud by

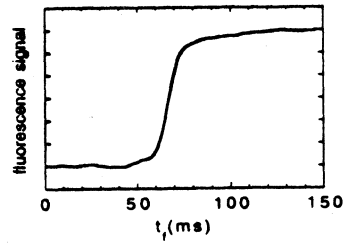


Figure 3: *Fluorescence signal measured in the detection region.*

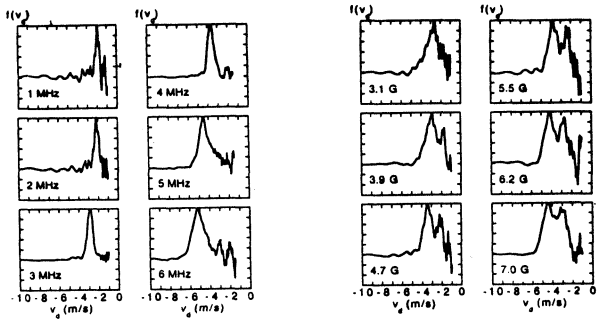


Figure 4: *Velocity distributions $f(v_d)$ of the atoms extracted from the 2D-MOT (a) by a 1D moving optical molasses for different detunings Δf and (b) by a static magnetic field for different magnetic field amplitudes B_z.*

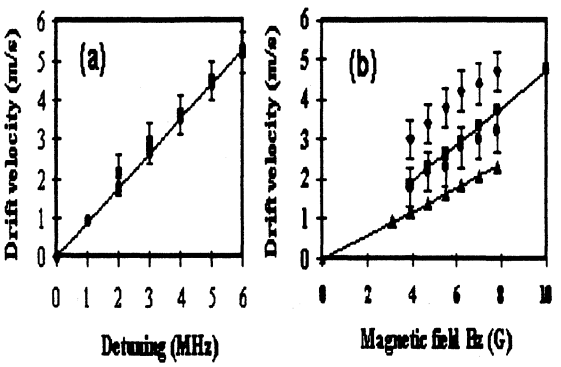


Figure 5: *Experimental and calculated slopes for launching by (a) a 1D moving molasses and (b) by a magneto-optical train.*

using a conventional time-of-flight technique (Ref.5), a heating process during the launching is obvious. A possible reason for this process is that the atoms when leaving the trap may cross a region, where they only interact with the horizontal beams. This would lead to a heating in the vertical direction. Further experimental investigations to clear up this problem are in progress. The measured mean drift velocities $\bar{v}_d$ are shown in Fig.5a as a function of the laser detuning Δf. They are in good agreement with the values calculated by eq.1, which demonstrate the efficiency of the launching process.

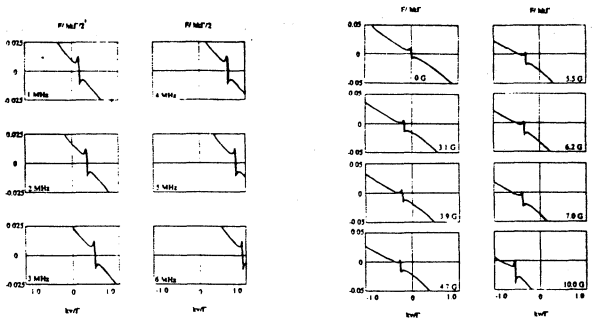
By launching with a static magnetic field the velocity distributions (Fig.4b) reveal a double-peaked structure, which suggests the existence of two different mean velocities for the atoms. As the two peaks are not completly resolved, it is difficult to determine the corresponding mean velocities. However, their position seems to depend linearly on the static magnetic field amplitude B_z with a similar slope of 0.41 ± 0.08 m/s/G (Fig.5b). The existence of a double-peaked strucure has already been experimentally proved in a 1D optical molasses (Ref.6).

For small values of B_z (< 3.1 G) our signal was so weak that the derivation of a velocity distribution was not possible because of noise. Our results prove the existence of a double-peaked structure in the velocity distribution by launching with magnetic field amplitudes between 3.9 and 7.0 G. The reasons for the unexpected large velocity distributions may be the same as for the high temperatures obtained by launching with a 1D moving molasses (see above), and its explanation needs further experimental studies. A better resolution of the two peaks would make it possible to determine the slopes more safely.

5.THEORETICAL CALCULATIONS

In order to explain these experimental results, we have developped a theoretical model using a $F = 1 \rightarrow F = 2$ transition system, but with the Cs Landé factors. This is the simplest atomic system if we want to observe the effects of Doppler and sub-Doppler cooling with circulary polarized light. We remind that sub-Doppler laser cooling appears when the light polarization varies and when the atoms exhibit a multilevel scheme for the ground level. This configuration allows us to achieve lower temperatures, well below the Doppler limit ($\simeq 125\,\mu K$ for Cs).

For our calculations, we first assume that the laser beams used for launching are vertically directed (not crossed as for the experiments). We also neglect the effect of the horizontal laser beams as photon recombination. So we present a simple 1D model along the Oz axis, defined as the quantization axis. We calculate the mean force versus the atomic velocity for the two launching techniques (Figs.6a and 6b). We cannot foresee the temperature from these curves but only the velocities at which the atoms are launched. For the experimental parameters, detuning $\delta = -3\Gamma$ and intensity per beam $I_0 = 5\ mW/cm^2$, the saturation parameter is well below the unity. So low



(a) (b)

Figure 6: *Calculated mean forces for launching (a) by a 1D moving molasses for different detunings Δf and (b) by a magneto-optical train for different static magnetic field amplitudes B_z.*

intensity approximation can be used to calculate the forces. In our model, we assume that $\delta = -3\Gamma$ and the saturation parameter $s_0 = 1$.

In the case of 1D moving optical molasses, the force becomes once zero for velocities different of zero. These velocity values depend linearly (Fig.5a) on the frequency detunings Δf. We clearly obtain only one velocity distribution because the Doppler and the sub-Doppler forces cancel out for the same velocity. There is a good agreement between the eq.1, the experimental results and these force calculations.

In the case of the magneto-optical train, the calculation results are very different (Fig.6b). For small magnetic field amplitudes ($B_z < 3.9\ G$), there is a single peak structure due to launching by a sub-Doppler mechanism (only the sub-Doppler force cancels out). The drift velocity is proportional to the Larmor frequency of the ground state and the calculated slope is about 0.30 m/s/G (Fig.6b). In an intermediate range ($3.9 < B_z < 10\ G$), a double-peak structure appears, for which the second peak is due to launching by a Doppler mechanism. For the second peak, the drift velocity is proportional to the Larmor frequency of the excited state and the calculated slope is about 0.48 m/s/G. Finally, for large B_z ($> 10\ G$) only the latter peak remains. The slope values are in agreement with the ones presented in Refs.7 and 8.

6.CONCLUSION

To conclude, we succeded in the generation of a continuous beam of cold atoms using a 2D-MOT and two different extraction mechanisms. At launching velocities of a few m/s we obtained up to 10^6 atoms/s in our beam. Our results suggest that launching with a 1D moving optical molasses is more promising than lauching with a static magnetic field, where the involved processes are more complex and lead to larger velocity distributions. Our source of a continuous beam of cold atoms may be an extremly valuable tool for atomic frequency standard.

ACKNOWLEDGMENTS

We acknowledge the assistance of G. Granger during a part of the measurements. This work was financially supported by the Bureau National de Métrologie (Paris) and the Direction des Recherches Etudes et Techniques (Paris). One of us (S.W.) acknowledges financial support by the european HCM-program.

REFERENCES

(1) A. Clairon, P. Laurent, G. Santarelli, S. Ghezali, S.N. Lea, and M. Bahoura, IEEE Trans. on IM **44**, 128 (1995)

(2) G.J. Dick, J.D. Prestage, C.A. Greenhall, and L. Maleki, in : Proceedings of the 22^{nd} Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Vienna, 487 (1990)

(3) P. Lett, W. Phillips, S. Rolston, C. Tanner, R. Watts, and C. Westbrook, J. Opt. Soc. Am. **B6**, 2084 (1989)

(4) G. Dudle, N. Sagna, P. Thomann, E. Aucouturier, P. Petit, N. Dimarcq, in : Proceedings of the 5th Symposium on Frequency Standards and Metrology 1995, Woods Hall, edited by J.C. Berquist, World Scientific, 121 (1996) and G. Dudle, N. Sagna, P. Berthoud, and P. Thomann J. Phys. **B 29**, 4659 (1996)

(5) P. Lett, R. Watts, C. Westbrook, W. Phillips, P. Gould, and H. Metcalf, Phys. Rev. Lett. **61**, 169 (1988)

(6) C. Valentin, M.-C. Gagne, J. Yu, and P. Pillet, Europhys. Lett. **17**, 133 (1992)

(7) M. Walhout, J. Dalibard, S.L. Roston, and W.D. Phillips, J. Opt. Soc. Am. **B9**, 1997 (1992)

(8) J. Werner, H. Wallis, and W. Ertmer, Optics Comm. **94**, 525 (1992)

AN ATOMIC RESONATOR OPERATING WITH A CONTINUOUS BEAM OF COLD CESIUM ATOMS

P Berthoud, G Dudle[1], N Sagna[2] and P Thomann

Observatoire Cantonal, Rue de l'Observatoire 58, CH-2000 Neuchâtel
[1]present address: NPL, Teddington, Middlesex TW11 0LW, UK
[2]present address: Laboratoire de Probabilité, Université Pierre et Marie Curie, Paris, France

ABSTRACT

Starting from a two-dimensional magneto-optical trap where Cesium atoms are permanently subjected to 3D sub-Doppler cooling and 2D magneto-optical trapping, we have produced a beam of cold atoms continuously extracted along the trap axis free of magnetic gradient. The simplest extraction mechanism, presently used, is the drift velocity induced by a constant magnetic field along the neutral axis of the trap.

We have used this continuous beam of atoms to produce Ramsey fringes in a traditional microwave cavity. The fringes are detected optically by the fluorescence from the atoms falling through a probe beam. This is the first demonstration of an atomic resonator operating continuously with laser cooled atoms. The flux density is 10^4 atoms s^{-1} mm^{-2}. Preliminary analysis of the Ramsey fringes is consistent with a sub-Doppler temperature in the atomic beam.

Keywords: atomic clock, laser cooling of atoms, continuous beam, Ramsey fringes

1 INTRODUCTION

Cold atoms are expected to bring considerable improvements in the accuracy and short-term stability of atomic frequency standards. In primary standards, cold atoms must be carried into - and out of - the microwave resonator, whose geometry does not allow on-site cooling or on-site detection of the resonance. In presently realised Cesium standards, the cooling and trapping laser beams are periodically shut off to allow launching of the cold cloud in a vertical ballistic trajectory (atomic fountain). Although this method already has shown its potential for accuracy [1], other approaches are worth investigating: one alternative consists in extracting atoms continuously from the trap [2]. Continuous extraction differs from pulsed extraction in several respects. While the resonator line Q is potentially lower in a continuous beam setup than in a fountain with the same size of the microwave interaction region, the average atomic density in a continuous beam is lower than in a fountain with the same average atomic flux by a factor equal to the inverse of the duty cycle in pulsed operation, typically 100. This is of interest since the collisional shift [1, 3] is reduced in the same proportion. Moreover, the continuous operation suppresses the short-term stability limitation [4, 5] which arises in a pulsed resonator from the intermittent presence of atoms in the microwave interaction region.

In this paper, we describe an atomic resonator using a continuous beam of Cesium atoms at sub-Doppler temperature to produce Ramsey fringes. The experimental setup is briefly outlined in section 2. Section 3 contains experimental results on the three following aspects:

1. Drift velocities induced in a 2D magneto-optical trap (MOT) by a homogeneous magnetic field. Three classes of atoms, with drift velocities centered at v_d = 0, 2.2 and 4.4 mm s^{-1} μT^{-1} are identified by their time-of-flight (TOF) spectra, the last one appearing only at high field ($B_z > 100\mu T$). All three velocity groups are characterized by sub-Doppler temperatures.

2. Production of a continuous beam of atoms, whose intensity is strongly correlated with that of the highest velocity group in TOF measurements.

3. Generation of Ramsey fringes in a short cavity. On the basis of these preliminary measurements, the temperature of the atomic beam is found to be in the sub-Doppler range. The shot-noise-limited stability of the resonator is estimated from the measured beam flux and resonance linewidth.

Section 4 concludes by pointing to possible improvements that will allow a short-term stability in the low 10^{-13} $\tau^{-1/2}$ range.

2 EXPERIMENTAL SETUP

We distinguish two parts in our vacuum system (figure 1): the top part, filled with a Cesium vapor at

room temperature (some 10^{-8} mbar) where atoms are cooled and trapped in a magneto-optical trap, and the bottom part where we interrogate cold Cesium atoms. The two parts are connected by a hollow cylinder of graphite, which acts as a getter for thermal Cesium atoms.

Our MOT uses three retroreflected cooling beams locked 2.5Γ (Γ is the natural width) below the $|F = 4\rangle \rightarrow |F' = 5\rangle$ transition and circularly polarized according to MOT requirements. Two of those propagate in the Oyz-plane at 45° with respect to the vertical direction. The third cooling beam is parallel to the Ox-axis. Each cooling beam has a power of about 20 mW and a diameter at $1/e$ of 18 mm. An additional repumping beam locked to the $|F = 3\rangle \rightarrow |F' = 4\rangle$ transition (not retroreflected) is superimposed to the Ox-beams.

To avoid trapping in the vertical direction we use an anisotropic magnetic gradient produced by four straight and vertical copper wires (up to 0.125 T/m). With this MOT configuration we get a trapping in the horizontal plane (Oxy) and a molasses vertically (Oz). The required compensation fields (Ox, Oy) and extraction field (Oz) are produced by three sets of Helmholtz coils.

Atoms falling from the trap pass through a Ramsey cavity placed vertically 200 mm below. This cavity, only 130 mm long, is surrounded by a C-magnetic field winding and two magnetic shields. To detect atoms after the cavity, we use a probe beam locked either to the $|F = 4\rangle \rightarrow |F' = 5\rangle$ or to the $|F = 3\rangle \rightarrow |F' = 2\rangle$ transition and measure the perpendicular fluorescence.

3 EXPERIMENTAL RESULTS

Cold atoms in anisotropic trap With respect to our previous MOT [6], we have improved several features: firstly we have increased the beam diameter and improved the homogeneity of the cooling beams. That leads first of all to a smoother shape of the cold atoms cloud (figure 2). Typical dimensions of the cloud are 30 mm long by 2 mm in diameter. Measured by the TOF technique, the temperature is about 25 μK for laser parameters optimizing the number of trapped atoms. For comparison, the temperature in a 3D trap with the same operating parameters is typically 50 μK.

Atomic drift in a magnetic field When a static magnetic field is applied during the cooling and trapping process, atoms thermalize around a drift velocity $v_d \neq 0$. The drift velocity is measured by the TOF technique 390 mm below the trap location. An evolution of TOF versus magnetic field B_z along Oz is shown in figure 3.

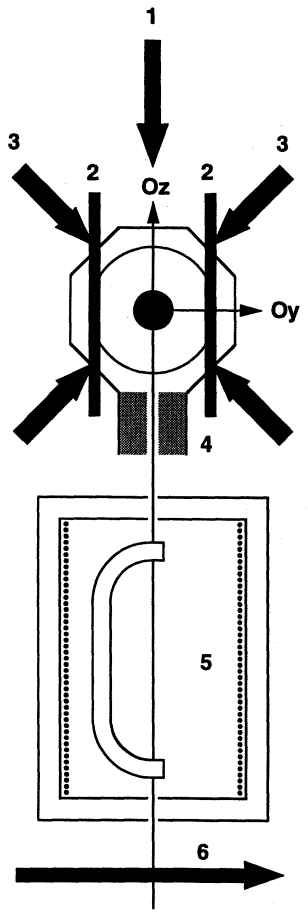


Figure 1: Continuous beam Cesium resonator. 1: static magnetic field; 2: 2D magnetic gradient; 3: cooling beams (Ox and repumping beams not shown); 4: thermal Cesium getter; 5: Ramsey cavity; 6: probe beam.

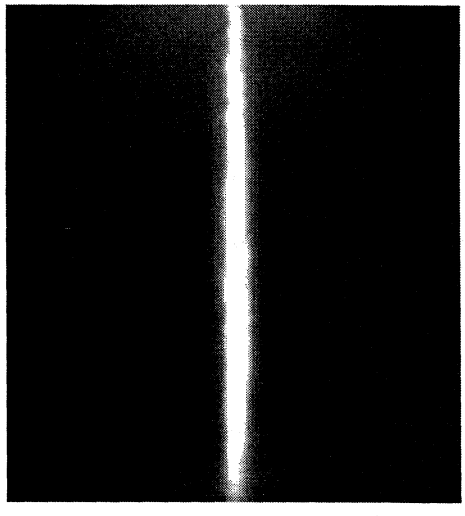


Figure 2: Cold atoms cloud in the anisotropic trap. (length: 30 mm)

At zero field the TOF looks very symmetric centered at 281 ms, corresponding to $v_d = 0$. By increasing the magnetic field we notice two important features : on the one hand a three-peaked velocity distribution appears, point on which we will come back later. On the other hand the signal decreases drastically between 0 and 100 μT (notice vertical scales in figure 3). A possible explanation of this loss of fluorescence signal could be a residual perpendicular drift velocity imparted to atoms leaving the trap when a magnetic field is applied.

An important improvement in the present experimental setup is a significant reduction of the thermal Cesium background in the detection zone. This allows to study weak TOF signals, especially in high magnetic fields. Indeed, over 100 μT, we clearly distinguish three peaks with different drift velocities: a weak peak always appearing after the same time as without magnetic field and two other peaks with drift velocities depending linearly on the magnetic field. It is interesting to notice that the integral of the total TOF profile, which measures the number of atoms detected, remains approximately constant above 100 μT. On the other hand, the proportion of the faster atoms increases with magnetic field. Over 200 μT only the fastest peak remains and we begin to see different DC levels before and after this peak. We will show further that this fact proves the existence of a continuous beam of cold atoms. The onset of the continuous beam coincides with the transition to the single-peak TOF profile where only the fastest atoms contribute. The widths of the TOF peaks provide an estimate of the longitudinal temperature according to the following relation, which depends on the initial velocity v_d:

$$T(v_d) = \frac{\Delta t_{1/2}^2 m g^2}{8 k_B \ln 2} C(v_d)$$

$$C(v_d) = \left(1 - \frac{v_d}{\sqrt{v_d^2 + 2gH}}\right)^{-2}$$

where $\Delta t_{1/2}$ is the TOF FWHM, m is the atomic mass, g is the gravitational acceleration, k_B is the Boltzmann constant and H is the falling height. All temperatures measured in these TOF spectra are sub-Doppler.

On figure 4 we have plotted the mean velocities of different peaks versus static magnetic field. As already pointed out, a small fraction of atoms does not undergo drift velocity. We explain this zero velocity peak by local potential wells in the vertical direction due to our 45° cooling lasers configuration. The second peak of TOF spectra, which dominates at low magnetic field (up to 100 μT) shows a drift velocity of 2.2 mm s^{-1} μT^{-1}. The 2D quantum model of Castin [7], which takes into account 45° configuration and is valid at low magnetic field, agrees well

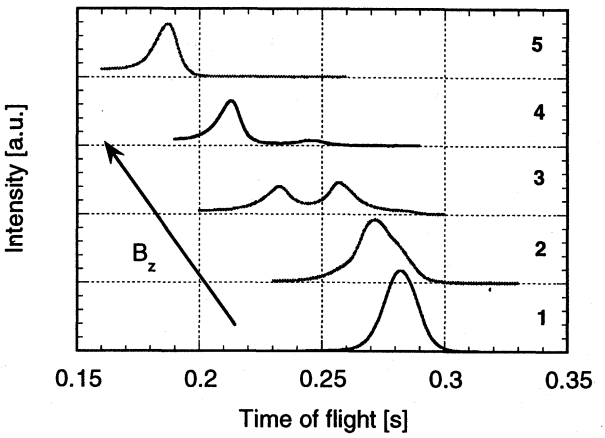


Figure 3: TOF spectra evolution *vs.* magnetic field. 1: $B_z = 0$ μT, scale 1:25; 2: 60 μT, 1:5; 3: 120 μT, 1:1; 4: 180 μT, 1:1; 5: 260 μT, 1:1.

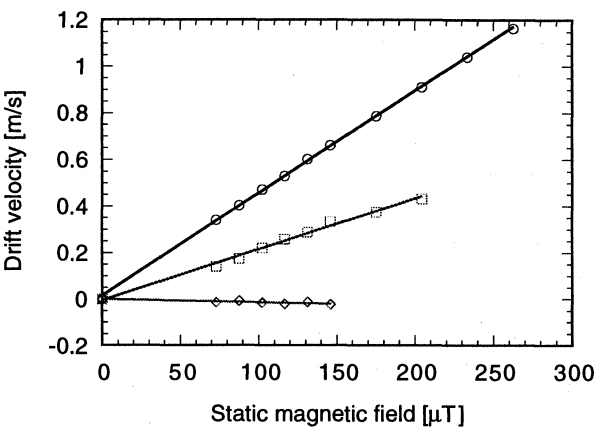


Figure 4: Drift velocities of figure 3 *vs.* magnetic field. Right peak is static, middle peak is drift at 2.2 mm s^{-1} μT^{-1} and left peak is launched at 4.4 mm s^{-1} μT^{-1}.

with our experiment (2.1 mm s^{-1} μT^{-1}). Concerning the fastest peak, which is mostly responsible for the presence of the continuous beam, we measure a linear dependance with magnetic field of 4.4 mm s^{-1} μT^{-1}, which is as yet not fully explained.

Continuous beam As already mentioned above, the static magnetic field provides a tool for the production of a cold atoms continuous beam. Figure 5 shows the fluorescent intensity transients from atoms crossing the probe beam (tuned to the $|F = 4\rangle \rightarrow |F' = 5\rangle$ transition) when the cooling lasers are shut off during 200 ms.

Before the TOF signal occurs (part 1 and 2), we detect different fluorescence levels depending on cooling lasers state. Indeed, by switching off the cooling light at $t = 0$, we reduce the depumping effect from the trap both on the continuous beam and on the thermal Cesium background, which results in an in-

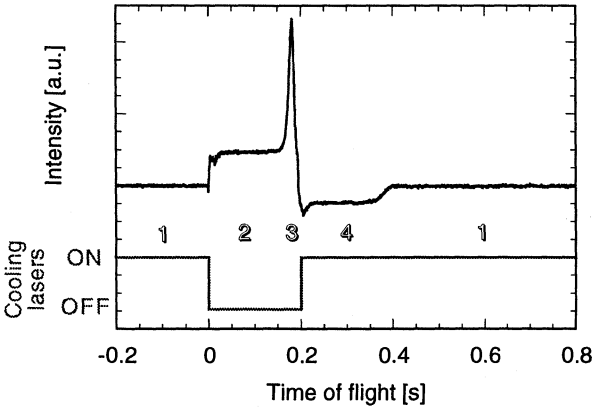


Figure 5: Continuous beam with a static magnetic field. The residual TOF signal occurs in 3 and the continuous beam intensity is the difference between 1 and 4.

crease of the fluorescence signal. In part 3 we observe the residual TOF signal 180 ms after light cut. This signal is due to atoms still in the trap when the cooling lasers are shut off. When we switch on the cooling lasers again in part 4, it takes the same 180 ms to see a step increase of the fluorescence signal. This increase is evidence of Cesium atoms being extracted continuously and reaching the probe beam from that time. The continuous beam intensity is measured by the difference of DC levels between parts 1 and 4. We estimate the atomic flux density to be about 10^4 atoms s^{-1} mm^{-2}. Notice also on this graph that the rise time of the continuous beam does not give any information about the longitudinal velocity distribution of the beam, but rather reflects the loading time of the trap.

Ramsey fringes We use an E-bend Ramsey cavity with a free precession length of 130 mm. Our goal is not for the moment to get the most precise and stable frequency standard, but only to show the ability of our cold atoms continuous beam to provide an atomic resonance signal. As the continuous fluorescence signal is very weak, we modulate the RF-power sinusoidally at 3 Hz and detect synchronously the fluorescence.

Figure 6 shows Ramsey fringes produced with a continuous beam of cold Cesium atoms. These fringes are obtained with the first optimal RF-power, which means an almost total π-pulse in the cavity at zero detuning. The total pulse is not exactly π, because the RF-power is modulated. The FWHM is 10 Hz, in agreement with the calculated width. This power only allows to distinguish nine fringes. Over 200 Hz detuning, the fringe contrast becomes too weak, because of the Rabi envelope behaviour.

In order to enhance the fringe contrast at high detuning, we repeat this measure with an approx-

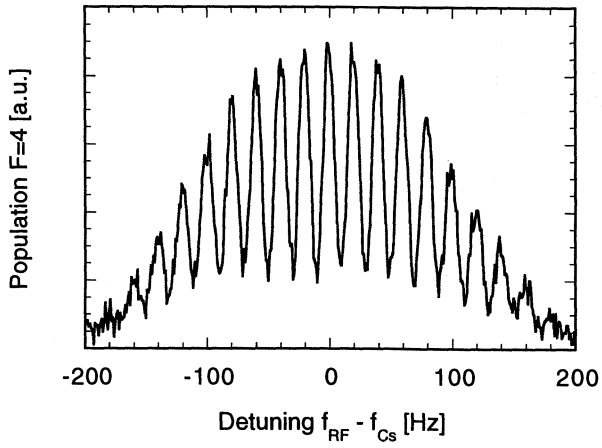


Figure 6: Ramsey fringes for the fisrt optimal RF-power. Measured and calculated FWHM: 10 Hz.

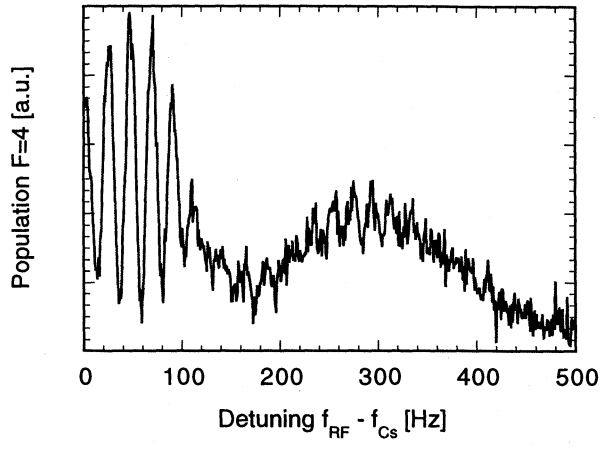


Figure 7: Ramsey fringes for the second optimal RF-power.

imately 3π-pulse (figure 7). We observe two types of contrast reduction: at low and at high detuning. The idea is to use these losses of contrast to get information about the velocity distribution in the beam along the vertical direction.

To explicit these contrast reductions we calculate steady-state Ramsey fringes. The RF-power modulation, necessary for synchronous detection, only affects the relative amplitude of the fringes below 200 Hz detuning. To account for the observed contrast reduction of fringes around 300 Hz detuning, we consider a Gaussian velocity distribution in the beam. Moreover, we take into account different interaction times in the cavity due to the gravitational acceleration (difference: 16%). Figure 8 shows theoretical Ramsey fringes calculated under the above assumptions and for a 3π-pulse. We observe a reduction of contrast at high detuning well understood by destructive interferences of different atomic velocity classes. At low detuning, the velocity distribution of atoms has, as expected, nearly no influence. The

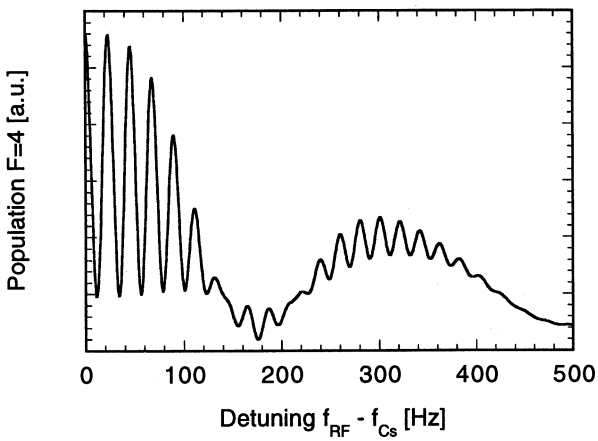


Figure 8: Ramsey fringes calculated for a 3π-pulse and using a longitudinal velocity distribution. ($T \leq 100\mu$K)

criterion that fixes the velocity distribution width is based on an equivalent fringe contrast around 300 Hz detuning with respect to experiment (figure 7). We obtain a satisfactory fit with a $v_{rms} \leq 80$ mm/s, which means a longitudinal temperature below 100 μK (sub-Doppler regime).

Resonator characteristics The magnetic field along Oz that optimizes the continuous beam intensity is about 300 μT. Atoms leave the trap with an initial velocity of 1.25 m/s and reach 2.6 m/s in the middle of the Ramsey cavity. Thus the effective atomic transit time is 50 ms. The estimated relative spread of longitudinal velocities in the continuous beam is 3%.

As already mentioned above, the total flux density is about 10^4 atoms s^{-1} mm^{-2} including all Zeeman sublevels m_F $(-3, \ldots, 3)$. The useful flux of $m_F = 0$ atoms (Ramsey fringes) is in good agreement with 1/7 of the total intensity of the continuous beam. Because of our short Ramsey cavity, the present fringes width of 10 Hz leads to an atomic quality factor $Q \approx 10^9$. So the theoretical stability $\sigma_y(\tau)$ of our resonator, if only limited by the atomic shot-noise, is $7 \cdot 10^{-12}\ \tau^{-1/2}$.

4 CONCLUSION AND PERSPECTIVES

We have shown in this paper the first results dealing with a sub-Doppler continuous beam. As extraction mechanism we have used a static magnetic field. As experimentally simple as it may be, this mechanism allows to extract continuously cold atoms from a magneto-optical trap. We have then used this continuous beam to get an atomic resonance signal (Ramsey fringes) as an alternative to the atomic fountain. These first results open the way to the development of a continuous beam frequency standard.

For the future, we plan to test potentially more efficient extraction mechanisms like the moving molasses and the dark channel. In addition, transverse cooling may prove useful to reduce the beam divergence and thus to further increase the beam flux. In parallel, use of a longer cavity to improve the atomic quality factor by 3 will allow a stability in the low $10^{-13}\ \tau^{-1/2}$ range.

ACKNOWLEDGEMENT

The authors wish to thank Y. Castin for the numerical computations he made for them. This work was supported by the Federal Office of Metrology and by the Swiss National Science Foundation.

REFERENCES

[1] A. Clairon, S. Ghezali, G. Santarelli, Ph. Laurent, E. Simon, S. Lea, M. Bouhara, S. Weyers, and K. Szymaniec. The LPTF preliminary accuracy evaluation of cesium fountain frequency standard. In *European Frequency Time Forum, 5-7 March 1996, IEE Conference Publication No 418*, pages 218–223, 1996.

[2] G. Dudle, N. Sagna, P. Thomann, E. Aucouturier, P. Petit, and N. Dimarcq. Generation of a continuous beam of cold cesium atoms. In *Proceedings of the Fifth Symposium on Frequency Standards and Metrology, 15-19 October 1995, Woods Hole, MA, USA*, pages 121–126. (James C. Bergquist, World Scientific, Singapore, 1996)

[3] K. Gibble and S. Chu. Laser-cooled cesium frequency standard and a measurement of the frequency shift due to ultracold collisions. *Phys. Rev. Lett*, 70(12):1771 - 1774, 1993.

[4] G. J. Dick, J. D. Prestage, C. A. Greenhall, and L. Maleki. Local oscillator induced degradation of medium-term stability in passive atomic frequency standards. In *Proceedings of the 22nd annual PTTI Meeting*, pages 487–508, 1990.

[5] C. Audoin, G. Santarelli, A. Makdissi, and A. Clairon. Dynamical behaviour of an oscillator slaved to a periodically interrogated atomic resonator. In *these proceedings*, 1997.

[6] G. Dudle, N. Sagna, P. Berthoud, and P. Thomann. Anisotropic magneto-optical trapping of atoms: capture efficiency and induced drift velocities. *J. Phys. B*, 29:4659–4673, 1996.

[7] Y. Castin. Private communication.

A NEW LASER SYSTEM FOR THE CS FOUNTAIN FREQUENCY STANDARD AT NRLM

G.A. Costanzo[#], S. Ohshima, K. Hagimoto, Y. Nakadan, Y. Koga

National Research Laboratory of Metrology
1-1-4, Umezono, Tsukuba, Ibaraki 305, Japan
e-mail: costanzo@nrlm.go.jp

1. ABSTRACT

The National Research Laboratory of Metrology is developing an atomic Cs fountain frequency standard. An experimental apparatus, useful to perform basic experiments on cooling and launching of Cs atoms, was realized.
In the past years we proposed a multipulse scheme and showed the feasibility of the reduction of some frequency inaccuracies (Ref.1) due to collisional frequency shift and the down conversion of the local oscillator noise.
To improve the laser beam quality, the MOT characteristics and the long term stability we realized a new optical set-up which will be an important part of the standard under development. The technical solutions adopted in the Cs atomic fountain standard and the related experiments on the cold atomic source are reported in this paper.

Keywords: laser system, cold atoms, Cs fountain.

2. INTRODUCTION

In the recent years the laser cooling and trapping of neutral atoms allowed a set of new apparatus for basic experiments in physics and spectroscopic fields (Ref. 2). Particularly the frequency and time community took advantage of these methods which seem necessary to overcome some inaccuracies in the *traditional* atomic Cs beam standard.
With the aim of a reduction of some causes of inaccuracies new technical solutions were adopted in many primary frequency standards (Ref. 3, 4) but at this moment the Zacharias idea of the atomic fountain seems the best solution to achieve an accuracy budget in the low 10^{-15} because of the characteristics which distinguishes this type of standard. Among these the long interaction time, the monokinetic velocity of the atoms and the intrinsic reduction of the cavity related shift are relevant benchmarks.
On the contrary the complexity of the Cs standards using the fountains scheme is increased of almost an order of magnitude because of the large number of lasers necessary to trap, pump and probe the Cs atoms. Consequently the performances of the laser system is crucial because the cold atomic source is strongly correlated to many parameters of the laser set-up.
Actually many laboratories are involved in the project of Cs fountain apparatus and a set of new solutions were proposed to overcome some sources of inaccuracies. Finally the LPTF laboratory already carried out many remarkable results on the fountain frequency standard and the preliminary evaluation of the operating standard reported an accuracy of the order of few parts in 10^{-15} (Ref. 5).

[#]G.A.Costanzo is a member of La.r.c.a. - Dip. Elettronica , Politecnico di Torino (Italy)

3. THE CS FOUNTAIN AT NRLM

NRLM planned to build a fountain frequency standard operating beside the optically pumped Cs clock which was realized in the late 80's (Ref. 6, 7, 8). Many papers have been already published (Ref. 1, 9, 10) and some basic experiments have been carried out to evaluate the feasibility and the performances of this new standard.
In fig.1 a rough set-up for the experimental apparatus is sketched out. Six independent laser beams, tuned to the red of the (F=4-F'=5) cycling transition, are introduced in the loading part of the vacuum chamber for trapping and cooling of the Cs atoms.

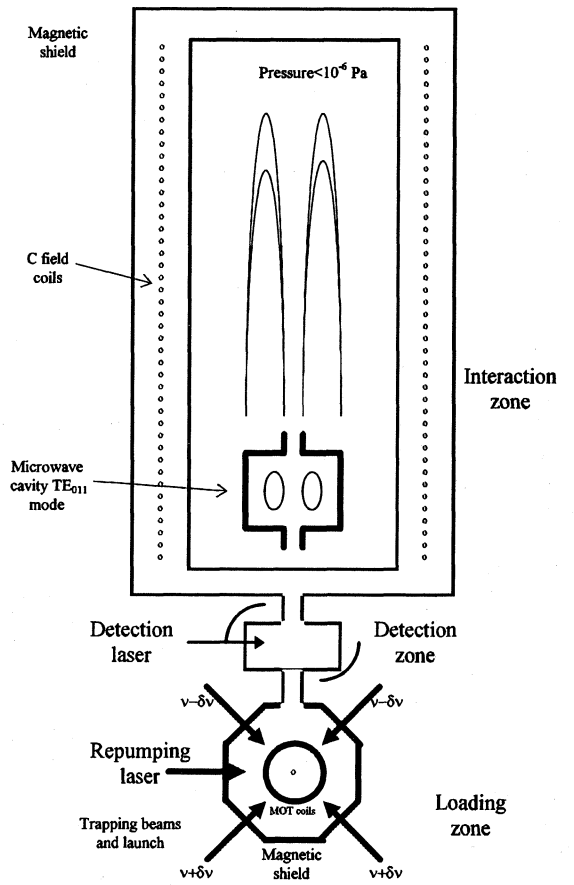


Fig.1 : Scheme of the Cs fountain at NRLM (horizontal beams orthogonal to the MOT coils are not shown).

The quadrupole magnetic field, useful for magneto-optical trapping, is generated by two anti-Helmholtz coils wound around the vacuum conflat flanges in the loading zone which is surrounded by a magnetic shield.

The horizontal beams are orthogonal to the Helmholtz coils plane whereas the vertical ones have an inclination of 45°. In this manner it is possible to introduce the vertical beams without the spatial reduction due to the passing holes in the end caps of the microwave cavity.
An additional laser beam, which is frequency locked on the (F=3-F'=4) transition, is then used to repump the atoms from the F=3 level.
The 3D σ^+ - σ^- MOT method is used to trap the atoms in the bottom part of the vacuum chamber connected through a valve to a Cs reservoir at the laboratory temperature.
After the launching sequence, which is performed with the moving molasses technique with a few millisecond post cooling period, the Cs atoms interact with the microwave frequency in the cylindrical copper cavity which resonates in the TE_{011} mode.
Thanks to their parabolic flight the atoms pass twice in the microwave cavity performing the two separate oscillatory field Ramsey resonance method. After that the atoms are detected by fluorescence using a probe laser, tuned on the (F=4-F'=5) cycling transition, and a light collection mirrors with an efficiency of several tens percent.
During the past years a couple of experiments have been realized using this apparatus with the purpose to increase the background knowledge towards this new standards.
During the first experiments the amplitude fluctuations of the lasers intensity and the beams positions were the main source of instabilities in the cold atomic source for the apparatus.
To overcome these problems we projected and realized a new optical set-up also with the aim to obtain a handling and modular structure which will permit an easier and faster maintenance of the final standard as delineated in the following paragraph.

4. THE OPTICAL SETUP

In the Cs fountains the laser system is an important part which has a remarkable complexity compared to the scheme used in the optically pumped Cs beam frequency standards.
To develop the optical set-up it is necessary to mind many parameters to achieve the goal of a highly stable trapped cold atoms. Frequency stability, beam intensity, beam quality, good flexibility in the tuning of the AOM devices are essential.
Nowadays semiconductor lasers in the near infrared region is a useful tool in the Cs spectroscopy. Extended cavity laser diodes are commonly used and commercial products are also available. Nevertheless the provided power is not exceeding several milliwatt.
An efficient solution (Ref. 11, 12) at this problem is achieved with the master-laser configuration which transfers the spectral purity of a low power stabilized master laser (ML) to the higher power of the slave injected laser (SL). For this purpose only several microwatt of the ML are necessary and, moreover, no servo electronics.
In Fig. 2 the new optical set-up for the Cs fountain is roughly depicted. In this scheme a commercially available diode laser system mounted in Littman-Metcalf configuration is locked on the crossover resonance between the (F=4-F'=5) and (F=4-F'=4) lines using the saturated absorption spectrum of a Cs cell which is surrounded by a magnetic shield to prevent the frequency instabilities due to the variations of the magnetic field in the laboratory environment. Few percent of the total power of the laser beam is inserted in an electro-optical modulator driven by a 30 MHz synthesizer.
In this manner the common FM sideband method is applied without the modulation of the diode laser current which introduce detrimental sidebands in the main beam. The saturated absorption signal is then detected by a PIN photodiode and subsequently demodulated in order to get the error signal which is used to control the laser current and the PZT voltage by the loop filter.
From the analysis of the Fourier Transform of the error signal in closed loop conditions it is possible to assume a value of several kHz for the laser linewidth which is dominated by the low frequency jitter due to the sensitivity of the extended cavity to the mechanical vibrations.

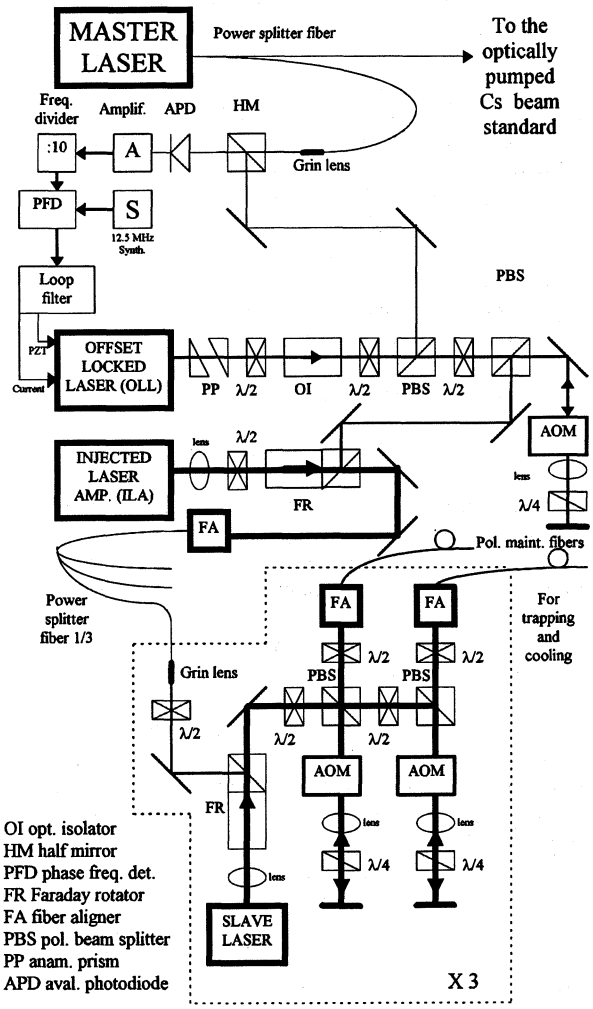


Fig.2 : The optical set-up for the Cs fountain (the dashed area shows one of the three injected slave lasers).

Afterwards the main power of the master laser is coupled into a single mode fiber and then splitted into two parts: one is used for the NRLM optically pumped Cs beam frequency standard and the other one for the fountain apparatus.
Each fiber provides about 1.5 mW useful to lock another laser (OLL).The offset frequency is about 125 MHz and it is accurately tuned by referring to the maximum fluorescence signal from the detection of the atomic Cs beam in the optically pumped Cs standard. The beat note between the ML

and the OLL is detected by an avalanche photodiode and then is fed in an AC coupled amplifier.
After a prescaling of the IF by a digital divider by 10 the signal is inserted in a digital phase and frequency discriminator (PFD) (Ref. 13, 14) which ensures a reliable lock for long time but showing a high sensitivity to the acoustic noise picked-up from the optical table. The output of the PFD is splitted in a fast component which control the laser current and a slow component for long term stabilization by the PZT voltage input of the extended cavity laser.
Fig. 3 shows the beat note between the ML and the offset locked laser: the signal to noise ratio is of the order of 30 dB in a bandwidth of the spectrum analyzer of 10 kHz. Thus the main power of the OLL is firstly up-shifted of (2x75)MHz using an AOM in double pass configuration and cat's eye combination and finally is injected in a single mode laser amplifier (ILA) with a CW power of 150 mW.

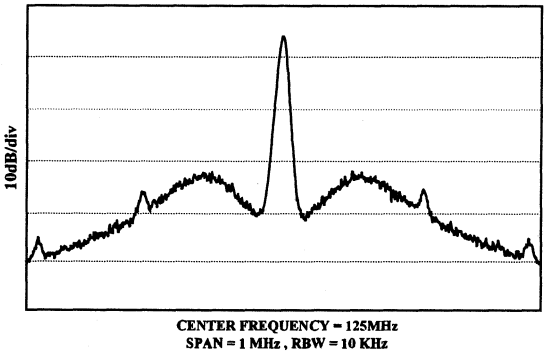


Fig. 3 : Beat note between the master laser and the offset locked laser in closed loop condition.

To monitor the laser spectrum we used a scanning Fabry-Perot interferometer with a full spectral range of 2 GHz. The injection beam (P=2mW) is aligned in the laser amplifier and, with the purpose to optimize the locking bandwidth, we measured the induced photocurrent when the ILA was switched off.
In this simple manner the maximum induced photocurrent is about 400 μA and it is possible coarsely evaluate the amount of the total injected power which is of the order of 400μW in the hypothesis of unitary quantum efficiency. A nearly optimum mode matching is performed by the temperature setting of the Peltier cooling element in the laser package.
Because during the launching sequence the OLL frequency undergoes a large frequency shift, the efficiency of the AOM and, consequently, the power useful to inject the ILA is decreased. Following that we drove the ILA at lower power with the aim to obtain the highest locking bandwidth for this laser.
When the ILA power is about 20 mW and the injected power is several hundreds of microwatt, we obtained about 10 GHz of locking bandwidth which ensures a good and robust stabilization for long term operation and still when we produce a frequency shift of the order of 10-15 MHz from the operating frequency (75 MHz) of the AOM.
Finally we coupled the ILA in a singlemode fiber which distributes the power to three independent slave lasers. Each beam coming up from the fiber is collimated by a grin lens and then injected in the SL. The three SLs emit about 150mW with a locking bandwidth of more than 2 GHz. The main laser power is splitted into two parts and each one undergoes a negative frequency shift using AOM in double pass configuration. Finally the laser beam is fed in an optical fiber. An half wave plate is used before the fiber aligner to accurately match the laser beam polarization with the polarization axis of the fiber.
After the beam expander the beam intensity has a full gaussian shape with a diameter of 30 mm ($1/e^2$ points). A quarter wave plate is placed close to the fiber output to achieve the desired σ^+ or σ^- polarization.
All the AOMs are driven by independent direct digital synthesizers (DDS) which frequencies are set by remote control with a computer. The output of each DDS is fed in a RF power amplifier to achieve the 1W power level necessary to properly drive the AOM.
For the fluorescence detection we used a DBR laser which is stabilized by injection locking using few percent of the OLL power. The DBR laser beam is then coupled in a optical fiber and used as probe in the detection zone of the fountain apparatus.
In this scheme problems are the long term instabilities of the laser power and the poor fiber coupling efficiency (about 30%). The former is mainly ascribed to the beam drift when the temperature variations slightly change the relative positions of the optical components.
Furthermore the coupling efficiency to the fiber is strongly correlated to the beam positions. Even if we used a double pass configuration for the AOM devices the laser beam intensities still showed a sensitivity to the operating RF frequencies driving the AOM.
Fig. 4 shows the laser intensities variations at the fibers output when the frequencies of the AOM devices are swept with a span of 10MHz: only the laser beam #4 shows a symmetric and bandpass-like characteristic unlike the beams #3 and #6.

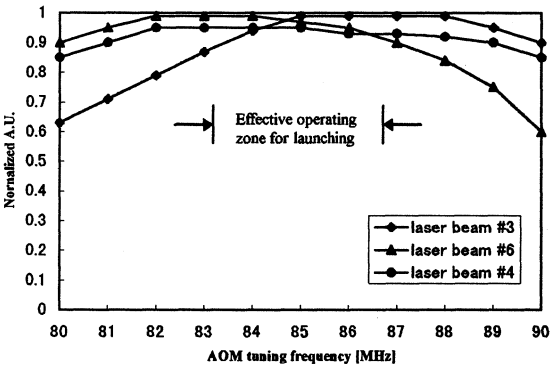


Fig. 4 : Laser beam intensity when AOM frequency is swept around the 85 MHz working point (for an easier caption only 3 of the 6 laser beams characteristics are shown).

In order to get more stable laser beam conditions we stabilized the power at the exit of the fibers controlling the RF power driving the AOM devices.
For this purpose we used optical beam samplers placed in front of the end of each fiber. A relatively low percentage of the laser power is detected by a photodiode and then fed in a control loop filter useful for the long term stabilization of the laser power.
This solution ensure a constant level of the laser power even if the AOM devices are tuned at different frequencies during the launching sequence. With the active control it is also possible to reduce the amplitude noise which is induced

when the optical fibers undergoes some unwanted micro bend.

5.EXPERIMENTAL RESULTS ON THE COLD ATOMIC SOURCE

With the optical set-up described in the previous chapter we performed some basic experiments with the aim to investigate on the performances of this new laser system which will be used for the fountain standard apparatus.
Firstly by a CCD camera we observed the characteristics of the 3D σ^+ - σ^- MOT: about 10^8-10^9 atoms are trapped when the six laser beams intensities are of the order of $2mW/cm^2$ and the laser frequency detuning is -20MHz. A stable shape of the atomic cloud is kept for long time.
Then the time of flight (TOF) method was used to measure the molasses temperature: after a 200ms loading time the MOT coils are switched off in few milliseconds and the molasses are further cooled by decreasing the intensity of the trapping beams to a value of about $0.3mW/cm^2$ and increasing the detuning. This cooling period can be adjusted by computer in order to get the best TOF signal. In this manner a final temperature of several microkelvin can be achieved.
After the ballistic flight the cold Cs atoms are detected in the lower part of the vacuum chamber (not shown in Fig.1) where a wide area photodiode collects the fluorescence signal obtained when the probe laser excites the (F=4-F'=5) cycling transition. A diaphragm is used to reduce the laser beam size to 2mm.
When the laser detuning is about -20MHz the detected TOF signal has a S/N of 65dB (1Hz of bandwidth) limited by the electronic noise and the scattered light. From integration of the TOF signal still the number of detected atoms is consistent with the number roughly evaluated from the MOT observation by CCD camera.
Moreover we can detect pure Cs molasses directly trapped with MOT coils switched off. In this case the number of loaded atoms is about ten times lower.
With this simple experiment it is possible to measure the molasses temperature by comparisons of the spread of the TOF signal with the related temperature of a maxwellian velocity distribution which best fits the TOF signal.
Fig. 5 shows the results obtained from these measurements at different laser frequencies detuning and laser beam intensities used during the cooling phase occurring after the MOT coils are switched off.

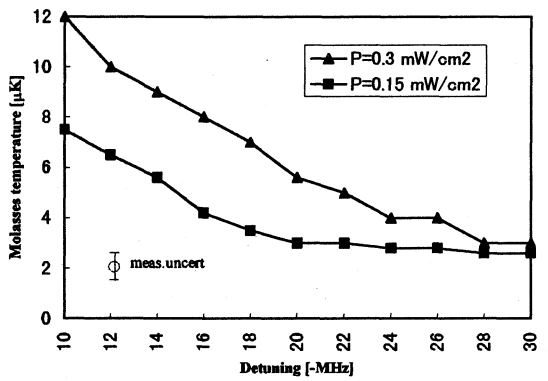


Fig. 5 : Temperature of the molasses at different laser detuning.

In Fig. 5 the TOF broadening due to the geometrical size of the probe laser and molasses is not removed and it limits the resolution of the measurement. As expected the molasses temperatures decrease with the laser detuning and the intensities of the lasers beams during the cooling phase. A saturation level of a couple of microkelvin is achieved for detuning larger than -20MHz.
Finally the TOF signal can be used to check the stability of the cold atomic source which is strongly related to the laser beams parameters (intensity, frequency stability, frequency detuning).
When laser power at the output of the fibers is not stabilized the fluctuations of the TOF signal are of the order of 10% peak to peak. Moreover we noticed an evident jitter of the arrival time of the molasses clouds.
These instabilities are notably reduced when the laser power control loop is operating: the fluctuations of the TOF signal become not exceeding 1% peak to peak even for long measurement time.
At this moment this value is probably limited by the fluctuations of background stray atoms and the Cs partial vapour pressure which change the trapping efficiency and affect the ballistic flight of the cold Cs atomic source.

6. SUMMARY

In this paper we presented the new optical set-up which will be used in the Cs fountain frequency standard at the National Research Laboratory of Metrology (Japan).
The high complexity of the lasers system was analyzed and we described some technical solutions as the offset locking and the injection locking of lasers. Moreover we used optical fibers for laser beam distribution in order to reduce the position fluctuations of the lasers beams. The main goal was to obtain a better lasers configuration which is necessary to increase the efficiency and the stability of the cold Cs atomic source.
With the aim to investigate and confirm the performances of this new optical set-up we also showed basic experimental results concerning the temperature measurements of the molasses by time of flight method. Finally the improved stability of the cold atomic source was here reported.

7. ACKNOWLEDGEMENTS

G.A. Costanzo wish to express his gratitude for the funding support from the Japanese Science and Technology Agency (STA) through the grant which permit his fruitful and agreeable stay in the NRLM Quantum Standards Section.

8. REFERENCES

[1] S. Ohshima, T. Kurosu, T. Ikegami, Y. Nakadan: "Multi pulse operation of Cesium atomic fountain", Proceeding of the 5th Symp. on Freq. Standard and Metrology, pp.60-65, Woods Hole (1995)
[2]see Journal of the Optical Society of America B, special issue on laser cooling and trapping of atoms, vol.6, n.11 (1989)
[3] W.D. Lee, J.H. Shirley, J.L. Lowe, R.E. Drullinger: "The evaluation of NIST-7", IEEE Trans. on I.&M., vol.44, n.2, pp. 120-123 (1995)
[4] A. De Marchi: "The high C-field concept for an accurate Cs beam resonator, Proc. 7th EFTF, Neuchatel (1993)

[5] A. Clairon, S. Ghezali, G. Santarelli, Ph. Laurent, S.N. Lea, M. Bahoura, E. Simon, S. Weyers, K. Szymaniec: "Preliminary accuracy evaluation of a Cesium fountain frequency standard", Proceeding of the 5th Symposium on Frequency Standard and Metrology, Woods Hole (1995)
[6] S. Ohshima, Y. Nakadan, Y. Koga: "Development of an optically pumped Cs frequency standard at the NRLM", IEEE Trans. on I.&M., vol.37, n.3, pp. 409-413 (1988)
[7] S. Ohshima, Y. Nakadan, T. Ikegami, Y. Koga, R. Drullinger, L. Hollberg: "Characteristics of an optically pumped Cs frequency standard at the NRLM", IEEE Trans. on I.&M., vol.38, n.2, pp. 533-536 (1989)
[8] Y. Koga, S. Ohshima, Y. Nakadan, T. Ikegami: "Performance of an optically pumped Cs frequency standard", Proc. Int. Symp. On Electr. Metrology, Beijing (1989)
[9] S. Ohshima, T. Kurosu, T. Ikegami, Y. Nakadan: "Cs atomic fountain with two dimensional moving molasses", Jpn. Journal of Applied Physics, vol.34, pp. L1170-L1173 (1995)
[10] T. Ikegami, S. Ohshima, Y. Nakadan: "Laser cooling of Cs atomic beam at the NRLM", Proc. 5th EFTF, Besancon (1991)
[11] J. Bouyer, C. Breant, P. Schanne: "Injection-locking mechanism in semiconductor lasers", SPIE, vol. 1837, pp.324-335 (1992)
[12] S. Kobayashi, T. Kimura: "Injection locking in AlGaAs semiconductor laser", IEEE J. of Q.E., vol QE-17, n.5, 1981, pp. 681-689 (1981)
[13] M. Prevedelli, T. Freegarde, T.W. Hansch: "Phase locking of grating-tuned diode lasers", Appl.Phys. B, vol.60, pp.S241-S248 (1995)
[14] G. Santarelli, A. Clairon, S.N. Lea, G.M. Tino: "Heterodyne optical phase-locking of extended-cavity semiconductor lasers at 9GHz", Optics communications vol.104, pp.339-344 (1994)

3D COOLING OF CESIUM ATOMS WITH ISOTROPIC LIGHT

C. Guillemot, Gh. Vareille, C. Valentin and N. Dimarcq

Laboratoire de l'Horloge Atomique
Unité Propre de Recherche du CNRS associée à l'Université Paris-Sud
Bâtiment 221, Université Paris-Sud, 91 405 ORSAY Cedex, France

1. ABSTRACT

We present a new 3 dimensional cooling configuration with isotropic light. We demonstrate that this cooling technique is very efficient and easy to operate. Various simple clocks configurations using a source of atoms cooled in isotropic light are described.

Keywords : Cold cesium clocks, miniature clocks

2. INTRODUCTION

The use of cold atoms has already demonstrated the great potentiality to improve both frequency stability and accuracy of frequency standards. At the present time, all the realizations are primary standards mainly based on the fountain principle. Such a device has been shown to exhibit high frequency performances. Primary standards using cold atoms will be obviously of great importance for high precision metrology and for precise tests in fundamental physics. Nevertheless, at the present time, the users of atomic clocks are generally satisfied by the frequency performances of existing clocks but they wish a reduction of the size of these clocks without any degradation of their performances.

The use of cold atoms is an answer to that wish. As a matter of fact, a common reduction - by the same factor - of the atomic velocity and of the microwave interaction length leads to the same interrogation time. The performances are expected to be roughly unchanged whereas the volume of the clock is reduced. But radiative cooling techniques will only become very attractive for the design of portable clocks when their operation will be easy. Unfortunately, the good operation of the existing generation of cold atoms clocks requires a great number of optical components for the control of the laser parameters (frequency, polarization, etc...). Moreover, a precise alignment of the different cooling laser beams is needed and mechanical misalignments may prevent from a long term reliable operation of such clocks.

We propose a new cooling configuration with isotropic light which has been demonstrated to be efficient and easy to operate. In section 3, we describe the experimental set-up and compare in section 4 the performances of this new cooling scheme to those obtained with the conventional cooling using collimated laser beams. First experimental results obtained on a preliminary prototype are shown in section 5. We present in section 6 two kinds of clock designs using a source of atoms cooled in isotropic light. The corresponding realizations should be very compact and should lead to high frequency performances, in the 10^{-15} level for applications on Earth and in the 10^{-16} level for space ones.

3. COOLING IN ISOTROPIC LIGHT EXPERIMENTAL SET-UP

The experimental set-up is depicted in Figure 1. The cesium atoms are contained in a pyrex cell (diameter = 4 cm). The isotropic light is created in an integrating sphere (diameter = 3 inches) made with SPECTRALON (produced by Labsphere). This Spectralon material exhibits a perfect Lambertian diffusion function and reflects more than 99% of the incident light at the considered wavelength λ = 852 nm.

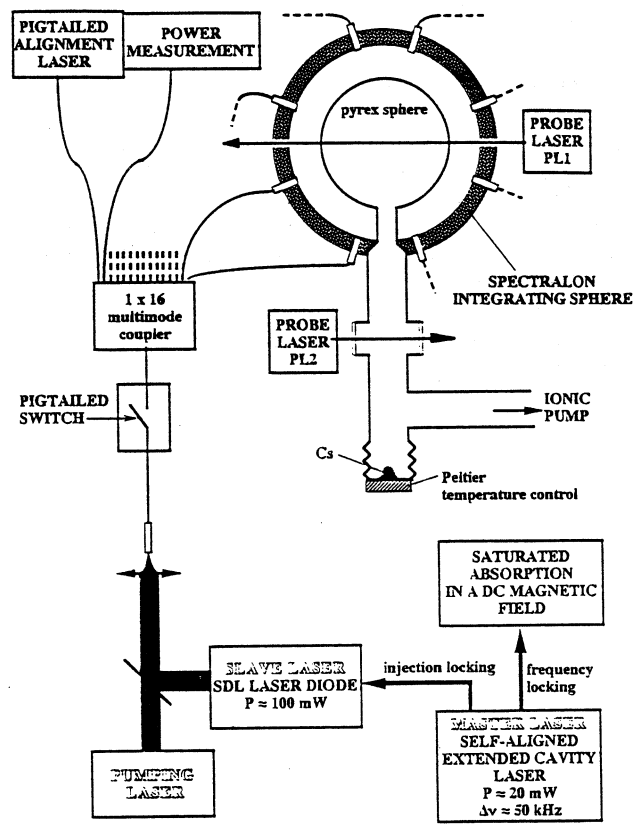


Fig.1 Experimental set-up

The laser light is injected inside the integrated sphere with 14 multimode fibers (50 / 125 µm) symmetrically disposed around the sphere. This number corresponds to fiber positions every 60° along the latitude and the longitude. These fibers are connected to a 1 x 16 Corning multimode coupler. The fifteenth output fiber allows the measurement of the power of the coupled laser light. For alignment purpose, a laser beam, issued from a pigtailed laser diode connected to the sixteenth output fiber, emerges from the input fiber. The superposition of

the cooling laser beams on this alignment beam allows easy and rapid settings of the lens which couples the light into the input fiber.

The cooling light comes from a SDL laser diode which delivers about 100 mW and which is spectrally purified with an injection locking technique. The master laser is a SC extended cavity laser. Its cavity exhibits high mechanical stability thanks to self-alignment techniques (1). Its frequency f_L is tuned below the resonance frequency f_0 of the $4 \rightarrow 5$ transition of the cesium D_2 line. The pumping laser, tuned to the $3 \rightarrow 4$ transition, avoids the trapping of the atoms in the F=3 fundamental state. About 30 mW cooling light and 1mW pumping light are injected in the integrating sphere.

No acousto-optical modulator is used. The cooling laser frequency is locked on the frequency of the saturated absorption signal for the $4 \rightarrow 5$ transition. The resonance of the absorbing atoms is shifted by Zeeman effect. The frequency detuning is then simply controlled by adjusting the magnitude of the applied static magnetic field.

For temperature measurements or clock operation, it is imperative to switch off the cooling and pumping lights with a high efficiency. For this purpose, we used two pigtailed switches in series with a global extinction ratio of 60 dB. The switching duration is about 1 ms.

Two probe laser beams are used (see Figure 1). We may observe the linear absorption of the PL1 laser beam to detect the presence of cold atoms inside the cell. The PL2 laser beam is used for measuring the cold atoms temperature with a time of flight technique (see section 5).

We have tested two kinds of pyrex cell. The first one has been frosted in order to avoid spurious specular light reflections at its surface. The second one was not frosted but we deposit on its inner surface a dry film made with OTS (2) in order to prevent the sticking of atoms on the cell surface (see section 5).

4. ADVANTAGES OF THE COOLING IN ISOTROPIC LIGHT

4.1 Basic advantages

In the conventional cooling configuration, the atoms undergo a radiative friction force in the intersection region of the collimated laser beams. The capture velocity, which is the maximal velocity of the thermal atoms which may be cooled, is then directly related to the laser beams diameter. A high number of cold atoms requires a high capture velocity and consequently a large beam diameter which must be associated to large and expensive optical components.

In the presented cooling scheme, the whole cell is irradiated by isotropic light. As a consequence, all the atoms contained in the cell undergo the radiative friction force.

The capture velocity is determined by Doppler cooling processes relying of the radiation pressure. These processes are all the more efficient as the laser light is resonant in the atomic frame. This implies that the laser frequency f_L shifted due to Doppler effect must be equal to the atomic resonance frequency f_0 :

$$f_0 = f_L + \frac{v \cos \theta}{\lambda} \qquad (1)$$

where λ is the laser wavelength, v is the atomic speed and θ is the angle between the atomic velocity and the laser beam direction. In conventional cooling schemes with collimated laser beams, this relation cannot be always fulfilled because the speed v changes when the atom is slowed down whereas f_0, f_L and λ are fixed and θ is random. In the cooling configuration with isotropic light, there is always a light cone (centered on the atomic velocity) which is resonant in the atomic frame. The angle θ of such a cone is fixed by the resonance condition (1). During the slowing down of the atom, this condition is always fulfilled because the variation of the speed v is balanced by the variation of the angle θ of the resonant light cone. This self-adaptation effect lead to a noticeable improvement of the cooling processes efficiency. It results an enhancement of the capture velocity.

The increase of both the capture velocity and the capture volume leads to a larger number of cooled atoms.

4.2 Experimental advantages

The experimental set-up is very easy to operate. Only the injection of the light into the multimode fiber requires precise settings. The integrating sphere plays various roles :

(i) it creates the isotropic light ;

(ii) thanks to the great reflectivity ($\approx$ 99 %) of Spectralon at 852 nm, the cooling light is stored inside the sphere. It results a noticeable enhancement of the effective intensity of the light seen by the atoms ;

(iii) the light diffusion processes inside the integrating sphere drastically filter spurious imbalances between the intensities of the beams emerging from the various fibers. Moreover, there is no need to control the polarizations of the cooling beams since the light inside the sphere is randomly polarized.

Furthermore, we have observed a greater insensitivity of the cooling configuration to spurious magnetic fields. Cold atoms have been observed without any compensation of Earth or instrumental magnetic fields. Nevertheless, it is only possible to obtain very low temperatures with sub-Doppler cooling mechanisms if the residual magnetic field is reduced.

5. EXPERIMENTAL RESULTS

5.1 Demonstration of the cooling

To demonstrate the feasibility of the cooling, we observe the linear absorption of the laser beam PL1 (see figure 1). Figure 2 points out the narrow absorption peaks of cold atoms superimposed on the broadened absorption signal of residual thermal atoms. The three peaks correspond to the three allowed transitions $4 \rightarrow 3'$, $4 \rightarrow 4'$

and 4→5'. The amplitude of the spurious broadened signal is directly related to the density of thermal atoms. We have decreased this density by cooling the Cs reservoir with a Peltier element. The number of cold atoms remained roughly constant whereas the background almost disappeared. Nevertheless, the loading duration of the optical molasse is increased when the density of thermal atoms is too low.

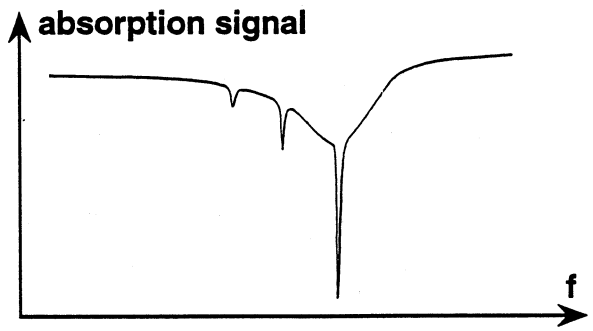


Fig.2 Linear absorption of the PL1 beam

The number of cold atoms is a few 10^9. We have observed a noticeable enhancement of this number when a dry-film is deposited on the inner surface of the pyrex sphere. Nevertheless, this is only a qualitative result and further detailed experiments must be undertaken to study this effect.

5.2 Temperature measurement

The easier way to measure the cold atoms temperature is the Time Of Flight (TOF) technique. At a given instant, cooling and pumping beams are switched off. The cold atoms fall and emit fluorescence photons when they cross the resonant probe laser PL2. A typical TOF signal is depicted in Figure 3.

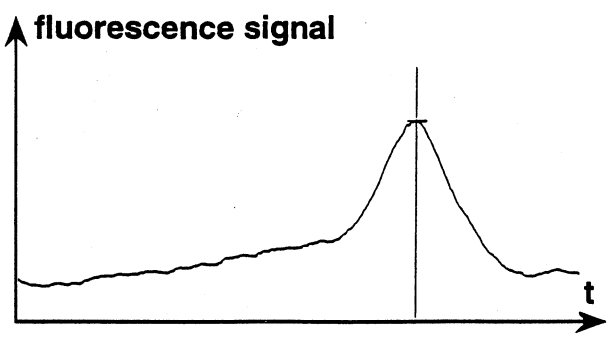


Fig.3 Typical Time Of Flight signal

The dispersion of this TOF signal results both from the dispersion of the initial atomic velocities and from the initial size of the cold atoms cloud. In conventional cooling, the fall height is generally far larger than the initial cloud size and the temperature can be directly deduced from the width of the TOF signal. In the isotropic cooling configuration, the initial locations of the cold atoms in the glass sphere are unknown. The fall height and the sphere diameter have similar values. As a consequence, the temperature cannot be precisely deduced from the TOF signal.

Moreover, for different experimental parameters, we have also observed TOF signals with 2 peaks (Figure 4) which point out strange locations of cold atoms inside the sphere. Nevertheless, the estimated temperatures are in the 30 μK - 300 μK range. The obtained sub-Doppler temperatures demonstrate the presence of intensity or polarization gradients in the isotropic light. A better control of the laser parameters (detuning, intensity) should allow us to reach the 1 μK level.

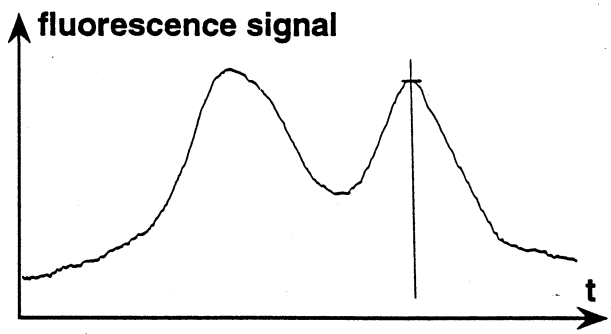


Fig.4 Double peak TOF signal

In order to measure more precisely the temperature, two techniques may be used :

(i) the double TOF technique where the cold atoms are detected at two different heights during their fall.

(ii) the stimulated optical Compton scattering technique (3).

6. CLOCK CONFIGURATIONS

As previously mentioned, it is important to design a cold atoms clock which is compact and which does not require neither numerous optical components nor critical settings. We present two basic configurations which take advantage of the isotropic cooling :

(i) the simplest clock where the cold atoms are launched using gravity

(ii) a miniature clock where the cold atoms are not launched

6.1 First configuration : launching with gravity

The set-up is schematically depicted in figure 5. The clock has a sequential operation. In a first step, the atoms are initially cooled with isotropic light. The lasers are then switched off. During their fall, the cold atoms undergo the microwave interrogation in a cylindrical cavity operating in the TE_{011} mode. The clock signal is optically detected below with a resonant laser.

The time sequence is shown in figure 6. If the length of the microwave cavity is 5 cm and the fall height is about 10 cm, the microwave interrogation duration is

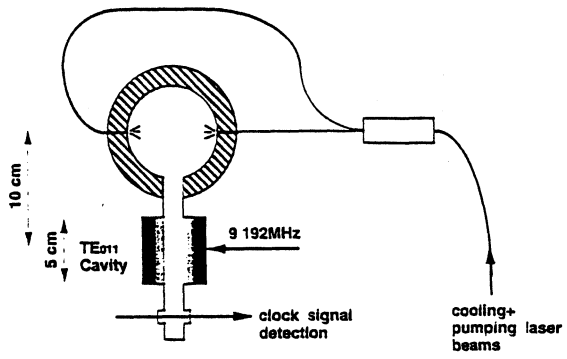


Fig.5 Clock using launching with gravity

about 30 ms. The durations of the various steps are the following :

- initial Doppler cooling : 100 --> 500 ms
- sub-Doppler cooling to reach very low temperatures : a few ms
- preparation of the atomic state : 10 μs
- microwave interrogation : 30 ms for example
- optical detection of the clock signal : 1 ms

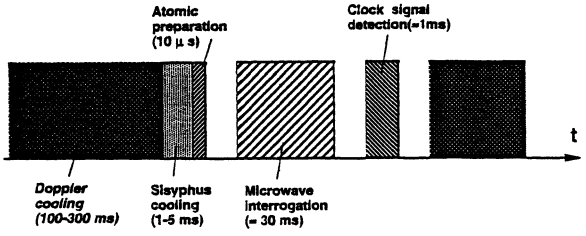


Fig.6 Time sequence

The clock signal obtained with this simple scheme is expected to have a linewidth about 30 Hz which should lead to frequency performances in the 10^{-15} level.

6.1 Second configuration : no launching

In this very different configuration, the atoms successively undergo the state preparation, the microwave interrogation (in a cylindrical cavity) and the clock signal detection at the same place where they have been cooled. The set-up is schematically depicted in figure 7.

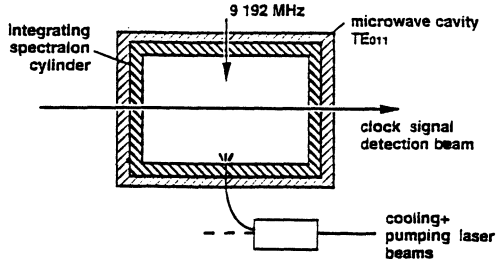


Fig.7 Miniature clock with no launching

The time sequence of the operation of this clock is shown in figure 8. The durations of the various steps in the first cycle are the following :

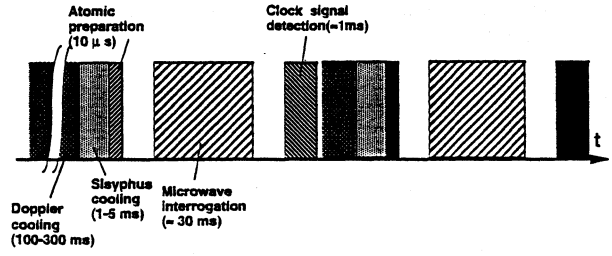
- initial Doppler cooling : 100 --> 500 ms
- sub-Doppler cooling to reach very low temperatures : a few ms
- preparation of the atomic state : 10 μs
- microwave interrogation : 30 ms
- optical detection of the clock signal : 1 ms

Two kinds of microwave interrogations may be chosen :

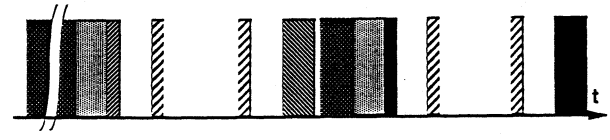
(i) the Rabi scheme : a single π pulse with a total duration T (figure 8-a)

(ii) the Ramsey scheme : two short π/2 pulses separated by a dead time T (figure 8-b)

For a same total duration, the Ramsey scheme leads to a linewidth twice lower than the one obtained with the Rabi scheme.



8-a Rabi scheme



8-b Ramsey scheme

Fig.8 Time sequence

Since the microwave interrogation duration is not too long, the cold atoms have not enough time to be warmed up again. As a consequence, the Doppler cooling step in the following cycles may be performed during a shorter time. This is one of the main advantages of this configuration compared to all the other clocks : the cold atoms are not lost at each cycle. As a result, the cycle duration may be consequently shortened and the duty cycle (interrogation time compared to the cycle period) is reduced.

The microwave interrogation duration T is limited by the displacement of the cold atoms during this time. Figure 9 shows the variations of this displacement as a function of T on Earth and in a microgravity environment. The corresponding linewidth, obtained in the Rabi scheme, is depicted in the same figure.

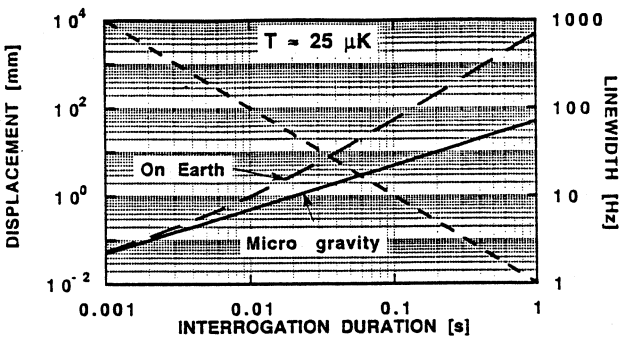


Fig.9 Cold atoms displacement and linewidth as a function of the interrogation duration

This miniature configuration is very flexible since the duration can be easily modified. This leads to a large range of frequency performances. If we accept a maximal displacement of 1 cm during the interrogation, the expected performances are the following :

(i) On Earth, the clock signal obtained is expected to have a linewidth as low as 20 Hz which should lead to frequency performances in the 10^{-15} level.

(ii) In a microgravity environment, these performances should be potentially improved by one order of magnitude.

7. CONCLUSION

The cooling configuration with isotropic light appears to be efficient, easy to operate and well suited to a compact clock with high frequency performances. Of course, an important work remains to do in order to reach such performances. In particular the following points must be studied :

- the interrogation of an atom with a accelerated motion in a cylindrical microwave cavity.
- the influence of collisions : between cold atoms, between an atom and the wall of the cell, etc...
- the determination of the exact locations of the cold atoms inside the sphere

ACKNOWLEDGMENTS

The authors want to thank the Direction des Recherches, Etudes et Techniques (Paris) and the Pôle de Transfert de Technologie Temps Fréquence (Besançon) for their financial support. They are also very grateful to Jacques Dupont and Pierre Petit in LHA for their efficient technical support.

REFERENCES

(1) " New extended cavity SC lasers structures using self-alignment techniques "
N. Dimarcq, M. Houssin, E. Aucouturier et M. de Labachelerie
Annual meeting of the Optical Society of America (OSA Annual), 10th Interdisciplinary Laser Science Conference (ILX-X), Dallas (USA), octobre 1994

(2) " Study of wall coatings for vapor-cell laser traps"
M. Stephens, R. Rhodes and C.E. Wieman
J. Appl. Phys., 76, pp. 3479-3488 (1994)

(3) " Method for velocimetry of cold atoms"
D.R. Meacher, D. Boiron, H. Metcalf, C. Salomon and G. Grynberg
Phys. Rev. A, 50, pp. R1992-R1994 (1994)

RESONATORS II

Chairman: R. Besson

THEORETICAL AND EXPERIMENTAL STUDY OF GRAVIMETRIC SENSITIVITY OF TRANSVERSE WAVES ON THIN QUARTZ PLATES

S. BALLANDRAS, J.B. BRIOT, G. MARTIN

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32 avenue de l'Observatoire - 25044 Besançon Cedex - France

ABSTRACT

Previous work has shown the interest of transverse wave propagation under metal strip grating deposited on thin Quartz plates. Particularly, thermal properties of these waves have been investigated from both theoretical and experimental points of view, emphasizing interesting temperature compensation effects on both AT and Z cut plate devices. High quality factors together with low insertion losses have also been measured on AT cut plate devices. As a consequence, the gravimetric sensitivity of such devices has been modeled and tested. A rigorous theoretical approach is then proposed to compute accurately the mechanical and electrical characteristics of the waves for different structures. This gives access to the frequency variation versus the thickness of the loading layer which appears strongly dependent on the elastic properties of the latter. Particularly, experiments have been performed on AT and Z cut plate devices loaded by Aluminium and Chromium layers. In the case of Aluminium, a decrease of resonance frequencies of both devices is predicted and measured. However, for chromium deposited on Z cut plate, an increase of lower mode resonance frequencies is observed. The proposed model has been used to explain this result. It is shown that due to a higher phase velocity of the wave in the chromium layer than in the Quartz substrate, first resonance frequencies of the loaded device are theoretically found higher than the ones corresponding to the unloaded device. Finally, theoretical and experimental gravimetric sensitivities are given and compared to state of the art.

1. INTRODUCTION

It is now well known that transverse horizontal waves, generated on any singly rotated Quartz cut (propagation along the rotated Z-axis), can be trapped on the surface by periodic grating, either thin metallic strip or shallow groove gratings [1-4]. Such trapped waves are called STW (Surface Transverse Waves) [1] and show great interest due to their high velocity (1.6 times more than of Rayleigh waves on (ST,X) Quartz) that allows to operate at higher frequencies without requiring any technological improvement.

A theoretical study and an accurate numerical modeling have been recently performed [5, 6], and resonators have been fabricated on thin AT-cut plates. Many modes can be generated on such thin plates, with insertion losses better than -10 dB and quality coefficient higher than 25000 measured at frequencies close to 130 MHz. A theoretical model based on Bloch-Floquet's series [1] has been written and numerical calculations have been done to predict the resonant frequencies of STW on AT and Z-cut thin Quartz plates. More accurate boundary conditions have been formulated to model the electrical and mechanical conditions on both lower and upper sides, allowing to predict the frequency shift due to mass loading on lower side. Experiments have been achieved on AT and Z-cut Quartz plates of 128 µm thick supporting metallic gratings of 20 µm spatial period on the top side. Aluminium (Aℓ) and Chromium (Cr) layers have been deposited on the bottom side of the plate and mass sensitivity has been then measured. A good agreement is found between theoretical predictions and experimental measurement.

Moreover, the elastic properties of the loading layer appear to have a strong influence on the sensing properties of the devices from both experimental and theoretical points of view. The first part of the paper is devoted to a description of the analytical model used to predict mass sensitivity of the devices. The results provided by this model are exposed and discussed in the following section. Finally, a comparison between theoretical and experimental data is reported, emphasizing the efficiency of the preposed theoretical analysis.

2. STW ANALYTICAL MODEL / MECHANICAL AND ELECTRICAL BOUNDARY'S CONDITION FORMULATION

2.1 Propagation equations

Figure 1 shows the theoretically studied device and experimentally tested. A definition of the axis system used to develop the analytical equations of STW propagation on thin Quartz plates is also reported on this figure. The polarization of the mechanical displacement is parallel to the X axis called a_1, and the STW propagates along the axis a_3 perpendicular to a_1. The axis a_2 is normal to the plate surfaces, pointing outwards the upper side. The mechanical displacement u_1 and the electrical potential ϕ assumed to be only depend on a_2 and a_3, allowing the development of a 2D model using scalar fields.

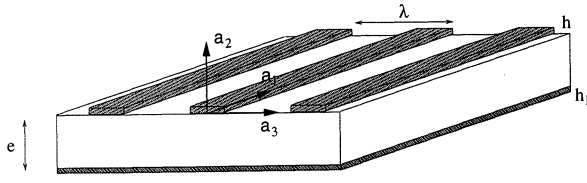


Fig. 1 : *Periodic metal-strip gratings on thin plate device and the coordinate system ($e = 128\,\mu m$, $p = 20\,\mu m$)*

Assuming a periodic grating infinite along a_3 allows to expend u_1 and ϕ in Floquet's series [6] consisting of an infinite sum of space harmonics which describes the spatial periodic modulation. Moreover, due to the finite dimensions of the plate along a_2, the system is correctly described by considering four partial waves instead of two as in the classical STW model [7]. As a consequence, u_1 and ϕ can be written as follows :

$$u_1 = \sum_{n=-\infty}^{+\infty} \sum_{r=1}^{4} (a_{nr} e^{-\alpha_{nr} a_2}) e^{-j(2\pi n/p) a_3} e^{j(\omega t - \beta_o a_3)} \quad (1)$$

$$\Phi = \sum_{n=-\infty}^{+\infty} \sum_{r=1}^{4} (b_{nr} e^{-\alpha_{nr} a_2}) e^{-j(2\pi n/p) a_3} e^{j(\omega t - \beta_o a_3)}$$

where a_{nr} and b_{nr} are respectively the amplitudes of the acoustic displacement and electrical potential, α_{nr} is the attenuation along a_2, β_o is the wave number and p the spatial period of the gratings. This first Floquet mode depending on $e^{-j\beta_o a_3}$ globally corresponds to a traveling wave along positive values of a_3. The second Floquet mode depending on $e^{j\beta_o a_3}$ is also solution of the system. The general coupled equations of motion, governing the STW behavior in a Quartz crystal are given by :

$$\left\{\begin{aligned}
&\rho_o \frac{\partial^2 u_1}{\partial t^2} = C_{66}\frac{\partial^2 u_1}{\partial a_2^2} + C_{55}\frac{\partial^2 u_1}{\partial a_3^2} + 2C_{56}\frac{\partial^2 u_1}{\partial a_2\,\partial a_3} \\
&\quad + e_{26}\frac{\partial^2 \Phi}{\partial a_2^2} + (e_{25}+e_{36})\frac{\partial^2 \Phi}{\partial a_2\,\partial a_3} + e_{35}\frac{\partial^2 \Phi}{\partial a_3^2} \\
&(e_{25}+e_{36})\frac{\partial^2 u_1}{\partial a_2\,\partial a_3} + e_{26}\frac{\partial^2 u_1}{\partial a_2^2} + e_{35}\frac{\partial^2 u_1}{\partial a_3^2} \\
&\quad - \varepsilon_{22}\frac{\partial^2 \Phi}{\partial a_2^2} + \varepsilon_{33}\frac{\partial^2 \Phi}{\partial a_3^2} + 2\varepsilon_{23}\frac{\partial^2 \Phi}{\partial a_2\,\partial a_3} = 0
\end{aligned}\right. \tag{2}$$

where ε_{ij}, e_{ij}, C_{ij}, and ρ_0 are respectively the dielectric, direct piezoelectric and elastic coefficients and the density of Quartz. Equation (2) allows one to determine the amplitudes a_{nr} and the attenuation coefficients α_{nr} both depending on the propagation constant β_o.

2.2 Mechanical and electrical boundary conditions

A set of boundary condition equations must be defined to numerically calculate the transverse wave characteristics. Because of the periodicity of the problem, one can reduce the analysis to only one spatial period. Accurate electrical and mechanical conditions are taken into account on both sides.

On the top side, a stress free condition is assumed on the upper surface of the metal strips. Considering the Quartz symmetries, the wave polarization and assuming a strip thickness h very smaller than the acoustic wavelength [6] provide the following form of this boundary condition :

$$T_{12}^{'}(e, a_3) + hf(a_3)\frac{\partial T_{12}^{'}(e, a_3)}{\partial a_2} - h\frac{\partial f(a_3)}{\partial a_3}T_{13}^{'}(e, a_3) = 0 \tag{3}$$

where $f(a_3)$ is a function describing the metal strip profile.

Because of the periodicity of the grating, periodic electrical boundary conditions must be calculated on the upper side of the plate. On the non metallized surface of the substrate, both electrical potential and electrical induction are continuous through the substrate-vacuum interface. This condition is then written for $-p/2 < a_3 < -p/4$ and $p/4 < a_3 < p/2$ as follows :

$$\left\{\begin{aligned}
&\Phi^S(e, a_3) = \sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4} b_{nr}\, e^{-\alpha_{nr}e}\, e^{-j\beta_n a_3}\, e^{j\omega t} \\
&\quad = \Phi^V(e, a_3) = \sum_{n=-\infty}^{+\infty} \Phi_n^{'}\, e^{-\gamma_n e}\, e^{-j\beta_n a_3}\, e^{j\omega t} \\
&D_2^S(e, a_3) = e_{21\ell}\frac{\partial u_1}{\partial a_\ell} - \varepsilon_{2\ell}\frac{\partial \Phi^S}{\partial a_\ell} \\
&\quad = D_2^V(e, a_3) = -\varepsilon_o\frac{\partial \Phi^V}{\partial a_2}
\end{aligned}\right. \tag{4}$$

where $\beta_n = \beta_0 + 2\pi n/p$, γ_n is the attenuation coefficient of the wave in the vacuum ($\gamma_n = |\beta_n|$), ϕ_n is the amplitude of the potential in the vacuum and D_2 is the electrical induction normal to the crystal surface. The upperscripts S and V respectively hold for the substrate and the vacuum.

At the interface between metal and substrate, the following equations are considered :

$$D^S(e, a_3) - D^V(e, a_3) = \sigma(a_3) \qquad \text{for } -p/4 < a_3 < p/4 \tag{5}$$
$$\Phi^S(e, a_3) = 0$$

where $\sigma(a_3)$ represents a charge distribution deduced from the propagation characteristics of the wave. The assumption of an infinitely thin strip has been performed to establish Eq. (5) which correspond to a shorted electrode condition.

On the lower side, the stress free condition is verified at the top of the loading layer :

$$T_{12}^{'}(-e - h_l, a_3) = 0 \tag{6}$$

where h_ℓ is the metal layer thickness.

However, it is necessary to calculate the acoustic displacement inside the layer to establish the stress and displacement continuity at $a_2 = -e$. This displacement $u_1^{'}$ can be rigorously calculated assuming that the metal is isotropic or cubic and considering that the waves exhibit the same polarization in the metal than in the substrate.

$$u_1^{'}(a_2, a_3) = \sum_{n=-\infty}^{+\infty} \Gamma_n(a_2)\, e^{-j\beta_n a_3}\, e^{j\omega t} \tag{7}$$

Using the propagation equation of the transverse waves in the layer allows to obtain the following expression of $u_1^{'}$:

$$u_1^{'}(a_2, a_3) = \sum_{n=-\infty}^{+\infty}\left(\Gamma_{n1}\, e^{-\gamma_n a_2} + \Gamma_{n2}\, e^{-\gamma_n a_2}\right) e^{-j\beta_n a_3}\, e^{j\omega t} \tag{8}$$

where

$$\gamma_n = \sqrt{(\beta_n)^2 - \frac{\rho^{'}\omega^2}{C_{66}^{'}}} = \frac{\omega\sqrt{V_{\ell a}^2 - V_S^2}}{V_{\ell a}\, V_S} \tag{9}$$

$V_{\ell a}$ is the velocity of the shear wave in the metal layer and V_S is an effective phase velocity of transverse waves in the substrate.
$\Gamma_{n1,2}$ are determined by the continuity relations of the acoustic displacement and of the stress T_{12} at the metal-Quartz interface. Then,

$$T_{12}^{'}(-e - h\ell, a_3) = \sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4} -C_{66}^{'}\gamma_n\, th(\gamma_n h_\ell) + \left(C_{56}j\beta_n + C_{66}\alpha_{nr} + (e_{26}\alpha_{nr} + e_{36}j\beta_n)d_{nr}\right) a_{nr}\, e^{-\alpha_{nr}e}\, e^{-j\beta_n a_3}\, e^{j\omega t} \tag{10}$$

In the case of a metal layer loading the lower side, the surface is assumed grounded. The electrical boundary condition can be expressed as Eq. (5) :

$$\Phi^S(-e, a_3) = \Phi^M(-e, a_3) = 0 \tag{11}$$

The relation of continuity gives, in the unloaded case :

$$\Phi^S(-e, a_3) = \Phi^V(-e, a_3) \tag{12}$$

Using this electrical and mechanical boundary conditions allows to obtain a set of linear algebraic equations whose determinant must be zero for the existence of a non trivial solution. Consequently the propagation coefficient β_o can be numerically evaluated by the resolution of this determinant.

The resonance frequency is then calculated using a transmission line model applied to a reflector-cavity-reflector structure. Since only synchronous devices are considered, the resonance generally occurs at the beginning of the stop-band of the resonator. The resonance condition of constructive interference at the mechanical ports of the cavity accurately determines the resonant frequency which is evaluated by numeric computation too.

3. NUMERICAL RESULTS

The model presented in Section 2 has been used to determine the influence of a loading layer on the dispersion relation of trasverse waves on thin plate. It has been shown [6] that this relation is usually characterized by many stop-bands, the first one taking place at a value of $\beta_0 = \pi/p$, p being the spatial period of the top side metal grating. Real and imaginary parts of β_0 are both plotted on the following curves.

The case of transverse waves on AT cut Quartz plate devices loaded by Aℓ layer has been investigated first. The evolution of the dispersion relation versus the thickness of the loading layer is reported in Fig. 2. This figure indicates a shift down of the frequency stopband along the increase of the loading layer thickness h_ℓ. This phenomena appears linear only for small values of h_ℓ (up to 0.1 µm). This shift gives the gravimetric sensitivity of the device (theoretically predicted values of sensitivity are resumed in Table I). The sensitivity S_m is defined by :

$$S_m = \lim_{\Delta\rho_s \to 0} \frac{\Delta f / f}{\Delta \rho_s} \qquad (13)$$

where Δf is the frequency shift and $\Delta\rho_s$ the surfacic mass variation.

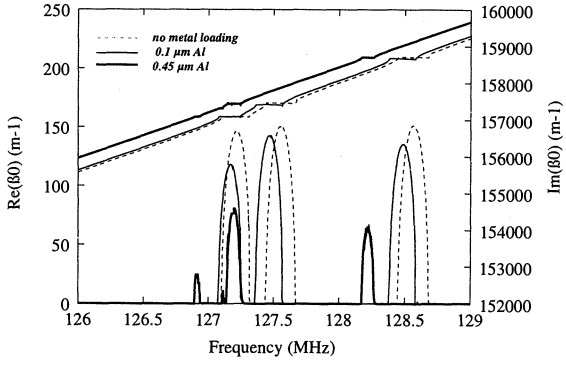


Fig. 2 : *Theoretical influence of Aℓ loading on transverse wave propagation*

For thick loading layers (>1 µm), the width and magnitude of the stop-bands strongly decrease. The reflection coefficient of the reflector strongly depends on this amplitude, and is expressed in a two spatial harmonic approximation [8] as follows :

$$\Gamma = \frac{x'_{-1}\, x_{-1}\, e^{2j(\pi/p)d} \left(e^{j\delta\beta_L} - e^{-j\delta\beta_L}\right)}{\left(x'_{-1}\, e^{j\delta\beta_L} - x_{-1}\, e^{j\delta\beta_L}\right)} \qquad (14)$$

where x_{-1} and x'_{-1} are the proportionality coefficient between the amplitude of modes $n=-1$ and $n=0$, for first and second Floquet modes respectively. L is the length of the reflector and d is the phase origin. $\delta\beta = \beta_0 - \pi/p$ represents the amplitude of the imaginary part of β in the stop band. Equation (14) shows that Γ is proportional to $\delta\beta$ for small values of $\delta\beta$, and then decrease as it. The quality factor of the resonator is then assumed to decrease for thick Aℓ loading, and even de-trapped the STW.

Figure 3 shows the Chromium loading effect on a Z-cut quartz plate device. In that case, the phase velocities of the first modes in the substrate are lower than shear waves in the layer ($V_{\ell a} \simeq 4\,000$ m.s^{-1} ; $V_S \simeq 3\,600$ to $4\,000$ m.s^{-1}). As a consequence, these first modes exhibit an evanescent behaviour in the layer and the electroacoustic energy trapping of these waves in the substrate is improved. Moreover, according to Fig. 3, the corresponding stopbands are shifted up, denoting a slight acceleration of the wave. The amplitude of the stopband is also increased, inducing higher reflection coefficient Γ according to Eq. (14). It should be noted that for Chromium thicknesses larger than an acoustic wavelength, no more variations of the dispersion relation are predicted. For higher modes, the apparent phase velocities of the waves in the substrate become larger than the layer shear wave one, and a behaviour similar to Aℓ on AT Quartz is emphasized. However, since the phase velocity difference is smaller in the present case than for Aℓ on AT Quartz, the frequency shifts and the decrease of the stopband amplitudes are smaller too.

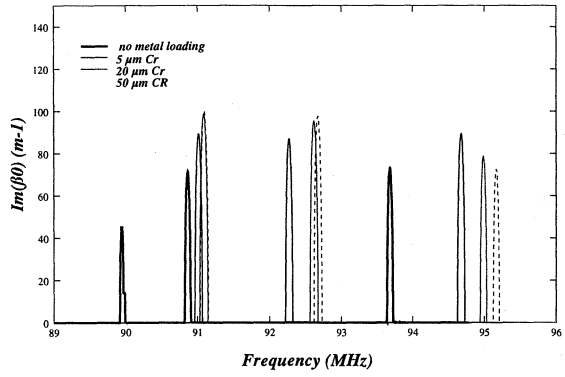


Fig. 3 : *Theoretical influence of Cr loading on transverse wave propagation*

4. EXPERIMENTAL RESULTS

Experiments have been done to validate the proposed model. Transverse wave resonators have been built on Z-cut plate with a 20 µm periodic grating and IDT's period is 20 µm so that there is no discontinuity in the device which may be considered as infinitely periodic. The electrical response of the tested resonator shown in Fig. 3 presents small insertion losses and quality coefficients

$$Q_{-3dB} = \frac{f}{\Delta f} \qquad (15)$$

larger than 10000 are measured in the frequency range 90 MHz-120 MHz. It can be remarked that the theoretical predictions for this device are in good agreement with experimental date, except for one mode which was predicted and not found experimentally.

A Chromium layer has been then deposited by e-beam evaporation on the lower side of the Z-cut plate. Although the thickness of the Cr layer was not accurately measured (estimated thickness 200 nm), the measured frequency shifts are found to correspond to a theoretical value of Cr thickness in the vicinity of 300 nm. More precise measurements are now required to provide better comparisons for Z-cut devices loaded by Chromium. However, it must be emphasized that the lower mode frequencies are slightly increased when Chromium present on the lower side, accordingly to the numerical results of Section 3. Also, the higher mode frequencies are shifted down, but the general aspect of the electrical response is not strongly affected in comparison with the unloaded device response (cf. Fig. 3).

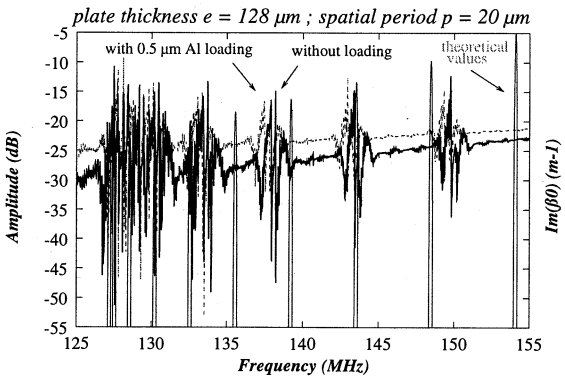


Fig. 4 : *Experimental influence of Aℓ lodading on the electrical transmission spectrum of transverse wave resonator on AT-cut quartz thin plate. Comparison with theoreticalvalue of wavve number β_0.*

Aℓ deposition on an AT-cut thin Quartz plate device allows more accurate comparisons between theory and experiments. The geometry of the device is the same as before but the transverse waves are more efficiently generated on AT-cut and electrical response looks much better (higher Q factors, smaller insertion losses). Aℓ is also deposited by e-beam evaporation and thickness of the layer is measured during deposition by a Quartz micro-balance : 0.5 μm of Aℓ were deposited by this way. Comparison between electrical responses of unloaded and loaded devices is shown in Fig. 4. Experimental frequency shifts correspond to the predicted ones and the sensitivity measured from this experiment is shown in Table I. A good agreement is found between the value of sensitivity of each modes, though it was not possible to measured it for the first modes because of the difficulty to identify the peaks corresponding to the theoretically predicted modes. A systematic difference of 10 % between experiments and theory is emphasized by Table I. This can be explained by the difference between fundamental elastic constants used for Aℓ in the proposed theoretical analysis and the effective constants of the layer due to the deposition operation. Moreover, the presence of internal stresses induced by the deposition technique can also modify the effective elastic properties of the metal.

Experimental				130.47	133.73	137.9	143.38	149.78	157.03
values				−2.1	−2.08	−2.20	−2.27	−2.25	−2.37
Predicted	127.31	127.66	128.66	130.35	135.69	139.34	143.64	148.6	154.21
values	−1.43	−2.49	−2.17	−2.24	−2.41	−2.43	−2.54	−2.74	−2.91

State of the art [8]

Acoustic plate modes		BAW μbalance on Quartz
1st mode	Higher order modes	
−1	−2	−3

Table I : *Comparison between predicted, experimental and state of the art values of gravimetric sensitivity (frequency in MHz, S_m in ppm/mg)*

5. CONCLUSION

A new model has been developed to predict the influence of metal layer loading on the resonance frequency of transverse waves resonators built on thin Quartz plates. This model based on a Bloch-Floquet's analysis of transverse waves propagation on thin plate has been validated by experimental measurements. The strong influence of the elastic constants of the loading layer on the wave properties has been theoretically and experimentally demonstrated. The ratio between phase velocity in the substrate and in the layer appears to strongly condition the sensing properties of the device.

The measured gravimetric sensitivities for AT cut devices are slightly smaller than those of classical SAW and slightly higher than classical Acoustic Plate modes. However, the proposed devices benefit from thermal compensation on an extended frequency range (127 to 150 MHz for AT cut devices) and from high quality electrical response which allow an accurate sensing of mass loading. Moreover, a clever choice of crystal cut allows to optimize the device for a given type of loading layer.

6. REFERENCES

[1] B.A. Auld, J.J. Gagnepain, M. Tan, "Horizontal shear waves on corrugated surfaces", Elect. Lett., vol. 12, n° 24, pp. 650-651 (1976).

[2] Yu.V. Gulyaev and V.P. Plesskii, Sov. Tech. Phys. Lett., vol. 3, n° 3, pp. 87-88, (march 1977).

[3] D.F. Thompson, "Temperature compensation of microwave acoustic resonators", Ph. D. Dissertation, Ginzton Laboratory, Stanford University, (june 1986).

[4] C.A. Flory and R.L. Baer, "Surface transverse wave mode analysis and coupling to interdigital transducers", Proc. 1987 Ultrasonics Symp., IEEE cat.n° 87CH2492-7, pp. 313-318.

[5] S. Ballandras, E. Gavignet, E. Bigler, G. Martin, "Experimental measurements of STW properties on quartz plates of finite thickness", Electronics Letters, vol. 32, n° 4, pp. 413-414 (feb. 1996).

[6] S. Ballandras, J.B. Briot, E. Gavignet, E. Bigler, "Propagation of transverse waves on piezoelectric plates supporting single or double side metal strip gratings", to be published in J. of Acoustical Society of America.

[7] E. Gavignet, S. Ballandras, E. Bigler, "Theoretical analysis of surface transverse waves propagating on a piezoelectric substrate under shallow-groove or thin metal strip grating", J. Appl. Phys., vol. 77, n° 12, pp. 6228-33 (1995).

[8] D.S. Ballantine, R.M. White, S.J. Martin, A.J. Ricco, F.T. Zellers, G.C. Frye, H. Wohltjen, *Acoustic wave sensors : theory, design and physico-chemical applications*, Academic Press, 1997.

INFLUENCE OF ELECTRODES ARRANGEMENT ON THE QUARTZ RESONATOR EQUIVALENT CIRCUIT

J. Zelenka * and J. Suchánek **

* Faculty of Mechatronics, Technical University in Liberec,
** Krystaly a.s., Hradec Králové,
Czech Republic

The analyses of the influence of electrodes on the frequency-temperature characteristic in AT-cut quartz resonators has been presented in [1]. The results of the study of the influence of electrodes dimensions on the parameters of the electrical equivalent circuit have been given by Tiersten [2] and some remarks were published also in Proc. EFTF96 [3]. The planparallel quartz plates with the same shape and dimensions of electrodes deposited on the both sides of the main planes of the plate vibrating in thickness-shear mode have been considered in all these cases.

The influence of the tabs of electrodes on the electrical equivalent cicuit parameters of AT-cut quartz resonators was shortly described by Suchanek and Zelenka [4].

In this paper the quartz resonators in the form of the circular or rectangular planparallel plate with the electrodes deposited in different arrangement on each side of the main planes of the plate so that only part of the electrodes overlaps is considered. The aim of the paper is analyzed the influence of electrodes on the electrical equivalent circuit and the resonant frequancy temperature dependence. The rotated Y-cuts and SC-cut quartz plates vibrated in the thickness-shear mode are considered by the analysis.

REFERENCES

[1] J. Zelenka: The calculation of the resonant frequency temperature dependence of quartz bulk resonators. Proc. 8th Piezoelectric Conference PIEZO'94, Zakopané 5-7 October 1994, pp. 49 - 60.

[2] H.F. Tiersten: Analysis of trapped-energy resonators operating in overtones of coupled thickness shear and thickness twist. J. Acoust. Soc. Am. 59 (1976), pp. 879 - 888.

[3] J. Zelenka: Influence of electrode mass-loading on the electrical equivalent circuit of the trapped-energy AT-cut quartz resonators. European Frequency Time Forum, Brighton 5-7 March 1996, pp. 234 - 237.

[4] J. Suchánek and J. Zelenka: Influence of the tabs of electrodes on the parameters of AT-cut quartz resonators. Proc. 9th Piezoelectric Conference PIEZO'96, Waplewo 2-4 October 1996,

CHARACTERIZATION OF SC-CUT RESONATORS OPERATING IN ANHARMONIC MODE

R. BOURQUIN, B. DULMET, J.J. BOY

ECOLE NATIONALE SUPERIEURE DE MECANIQUE ET DES MICROTECHNIQUES
Laboratoire de Chronométrie, Electronique et Piézoélectricité
26 Chemin de l'Epitaphe
25030 BESANCON CEDEX FRANCE

ABSTRACT

Resonators operating on the first anharmonic modes (n01 or n10) can be drived by means of two pairs of cross connected electrodes. The width of the gap between the electrodes and its orientation are the main parameters that enable to obtain a better electrical response of the C-mode. The computation of the electrical parameters is based on the Lewis formula and takes into account the geometry of the electric field in the gap. We give here examples of the variation of the motional resistance versus the gap width and the shift of the mode pattern with respect to the electrodes.

Keywords : Bulk wave resonator, contoured resonator

1 - INTRODUCTION

It is possible to design an SC-cut resonator for which the motional resistance of the C-mode is sufficiently lower than that of the B-mode so that its use in an oscillator does not require any selecting circuit. Such resonator operates on an anharmonic mode, having two vibrating areas with opposite phases, instead of the classical mode presenting an amplitude of vibration with a maximum at the center of the resonator **[1]**.

The reduction of the B-mode response is possible because, for some modes, the orientations or the shapes of the modes of the two families are sufficiently different. The driving of the selected mode is obtained by using two pairs of cross connected electrodes with a geometry giving a better drive of the C-mode (Fig. 1).

In a previous paper **[1]**, we have presented a simplified anlaysis of such a resonator. This analysis assumes that the electric field is strictly homogeneous between the electrodes and zero outside. Futhermore, only the derivative of the mechanical displacement with respect to the thickness of the resonator is taken into account. The here presented analysis removes these limitations without any unnecessary refinements.

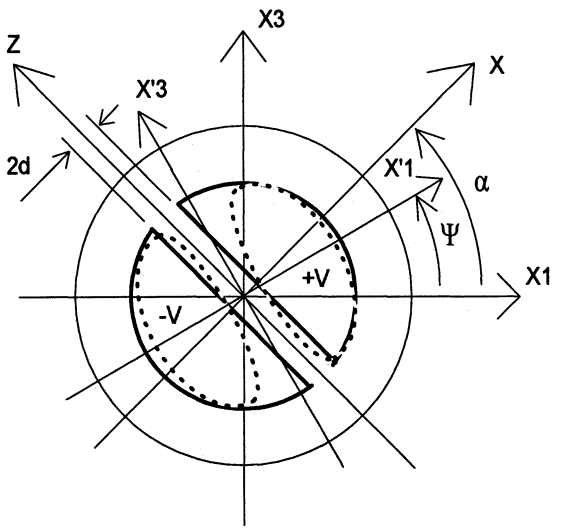


figure 1

2 - MOTIONAL PARAMETERS

The capacitance ratio C_1/C_o between the motional capacitance C_1 and the static capacitance C_o is computed from the Lewis formula **[2]**

$$\frac{C_1}{C_o} = \frac{\left(\int_v e_{mi} E_m S_i \, dV \right)^2}{\int_v \varepsilon_{ij} E_i E_j \, dV \int_v C_{ij} S_i S_j \, dV} \qquad (1)$$

where E_k are the components of the static electric field in the resonator, S_i the strain components due to the vibration and V the volume of the resonator.

Static electric field

We are mainly interested by the electric field in the gap between the electrodes. The non homogeneity of the field at the periphery of the electrodes can be neglected because of the smallness of the strain in this region. We also neglect the component of the electric field in the z-direction. The electric field is calculated by using the Schwartz-Christoffel conformal mapping **[3]**.

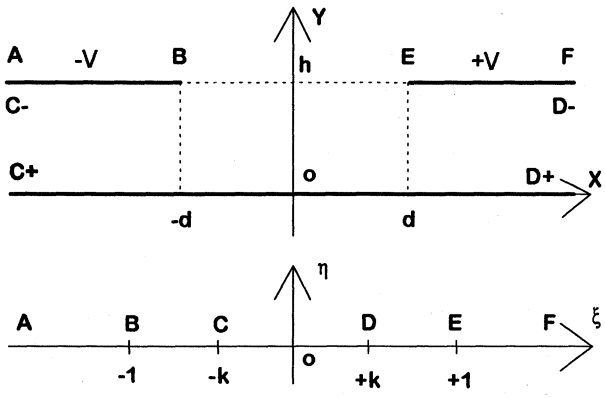


figure 2

The figure 2 shows a half section of the resonator and the corresponding transformed plane. Between the coordinate

$$z = x + iy$$

of a point in this section and the corresponding coordinate

$$\zeta = \xi + i\eta$$

in the transformed plane, we have the relation :

$$z = \frac{h}{\pi}\left(i\pi + \frac{2k}{1-k^2}\,\zeta - \mathrm{Log}\frac{\zeta - k}{\zeta + k}\right) \tag{2}$$

k is a parameter that depends of the ratio d/h, and is the solution of :

$$\pi\frac{d}{h} = \frac{2k}{1-k^2} - \log\frac{1-k}{1+k} \tag{3}$$

Since the potential is known along the ξ-axis, (-V between A - C, +V between D - F and zero elsewhere), the complex potential U+iF can be calculated in the transformed plane. So ζ can be expressed as a function of the potential U (-V<U<V and the flux F (- ∞ <F<+ ∞) at the corresponding point by :

$$\zeta^2 = k^2 + e^{\pi\frac{F}{V}}\left[\cos\left[\pi\left(1-\frac{u}{v}\right)\right]\right] + i\sin\left[\pi\left(1-\frac{u}{v}\right)\right] \tag{4}$$

The electric field is given by :

$$E_x = \frac{k^2-1}{kh}\,\frac{(1+\xi^2+\eta^2)\,\eta}{(1-\xi^2+\eta^2)^2+4\xi^2\eta^2}\,V$$

$$E_y = \frac{k^2-1}{kh}\,\frac{(1-\xi^2-\eta^2)\,\xi}{(1-\xi^2+\eta^2)^2+4\xi^2\eta^2}\,V \tag{5}$$

Dynamic strain

In the (x'_1, x_2, x'_3) frame, the mechanical displacement is given by :

$$u'_i = A_i\, u_{nmp} \tag{6}$$

A_i being the components of the displacement and u_{nmp} its spatial dependence [**1, 3**]

$$u_{nmp} = \sin\frac{n\pi x_2}{2h}\,H_m\left(x'_1\sqrt{\alpha_n}\right)Hp\left(x'_3\sqrt{\beta_n}\right)e^{-\frac{\alpha_n x'^2_1+\beta_n x'^2_3}{2}} \tag{7}$$

H_m,H_p are Hermite functions and α_n,β_n scale factors that are defined below. The two integrals corresponding to the electric energies are numerically evaluated. The mesh is made in the (x, y, z) frame, the cells in the (xy) plane are obtained by incrementing the potentiel U and the flux F. This way offers two advantages. First one, the electric field is then directly computed from (4) and (5) without having to solve the equation (2) as in the case where the mesh is made directly from the (x, y) coordinates. Second one, the size of the elements depends on the electric field intensity, a small element corresponding to a high electric field.

For the precision needed here, it is not necessary to evaluate all the terms in the second integral in the denominator of (1) corresponding to the elastic energy. Labeling F_2' the derivative of u_{nmp} with respect to x_2, ,this integral can be directly calculated by the following result when the indices m and p are equal to 0 or 1 :

$$\int_V C_{ij}S_iS_j\,dV = \int_V \left(C_{66}A_1^2 + C_{22}A_2^2 + C_{44}A_3^2\right)F_2'^2 dV = \hat{c}\frac{n^2\pi^3}{4h\sqrt{\alpha_n\beta_n}} \tag{8}$$

the integration being performed for a resonator of infinite diameter because of the trapping of the vibration.

The static capacitance of the resonator is given by

$$C_o = \frac{1}{4V^2}\int_v \varepsilon_{ij}E_iE_j\,dV \tag{9}$$

and the motional resistance R_1 by

$$R_1 = \frac{1}{QC_1\omega} \tag{10}$$

3 - RESULTS

The motional resistance R_1 depends on the size of the electrodes, especially the gap value, and on their orientations with respect to the mode pattern. The more different the mode patterns, the better the possibility to improve the C-mode response. The orientation of the mode patterns must be experimentally determined by varying the orientation of the electrodes [1] or by using X-Ray topography [5].

The curves given below have been obtained for 5 MHz (n, 0, 0 mode) SC-cut resonators having the following parameters:

Diameter of electrodes : 6 mm
3rd overtone : thickness : 2h = 0.542 mm
radius of curvature : R = 200 mm
5th overtone : thickness : 2h = 0.919 mm
radius of curvature : R = 115 mm

All curves, as other numerical data, are drawn assuming a Q factor equal to 1.10^6. In practice, one must remember that the Q factor is not the same for both mode families, the B- mode exhibiting a slightly higher value.

The mode shapes are governed by the parameters α_n and β_n which are related to the dispersion constants M_n and P_n and the effective elastic coefficient $\bar{C}$ of a mode family [4] :

$$\alpha_n = n\pi\sqrt{\frac{\bar{c}}{8Rh^3M_n}} \quad ; \quad \beta_n = n\pi\sqrt{\frac{\bar{c}}{8Rh^3P_n}} \qquad (11)$$

The values of M_n and P_n we have used are those deduced from the resonant frequencies of a resonator [1]. The adopted values, as well as the orientation ψ_n of the modes, are summarized in the table I.

Mode	Overtone	M_n	P_n	Ψn
B	3	74	38	43°
	5	85	97	- 56°
C	3	57	67	48°
	5	69	80	51°

Table I : Dispersion constants (in 10^9 N/m^2) and orientation angles of mode-shapes.

The figures 3 to 8 give the variations of the motional resistance R_1 as a function of the gap between the half electrodes for various orientations of the electrodes with respect to the mode pattern comprised between $\psi - \alpha = -20°$ to $\psi - \alpha = +20°$. We have noticed that the curves are not very sensitive to the value of the angle ψ of the mode.

We may notice that for the 301-modes the motional resistance R_1 of the B-mode ,(fig. 4) increases more quickly with the gap than that of the C-mode (fig. 3). This fact is due to the high trapping of the vibration of the B-mode in the x'_3 direction (the P_3 constant is about half that of M_3). The C-mode is slightly more sensitive to a variation of ψ-α than the B-mode.

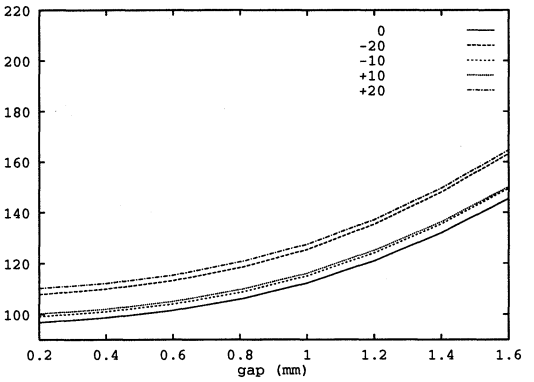


figure 3 C-301 Mode : Motional resistance R1 versus gap for various orientations of electrodes

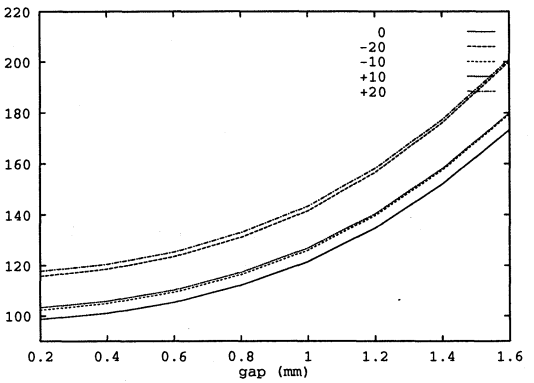


figure 4 B-301 Mode : Motional resistance R1 versus gap for various orientations of electrodes

For the 310-modes (fig. 5 and 6), on the other hand, both modes exhibit quite the same feature. Since the orientations of the B and C-modes of the third overtone are very close, the improvement of the C-mode response is only possible by using the C-301 mode orientation with electrodes having a wide gap with the same orientation as the mode pattern.

In the case of the fifth overtone, the modes of both families exhibit similar behaviours. The figures 7

and 8 are drawn for the 501 modes. Curves for the 510 modes are very similar. Since the orientation ψ of the B and C modes are different, it is possible to improve the C-mode response by using an orientation for the electrodes that gives a better driving of this mode than for the B-mode, but this is a less efficient way.

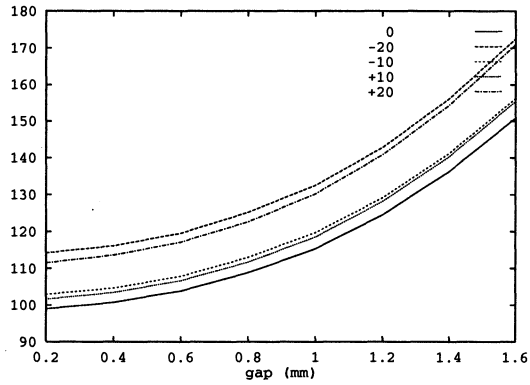


figure 5 C-310 Mode : Motional resistance R1 versus gap for various orientations of electrodes

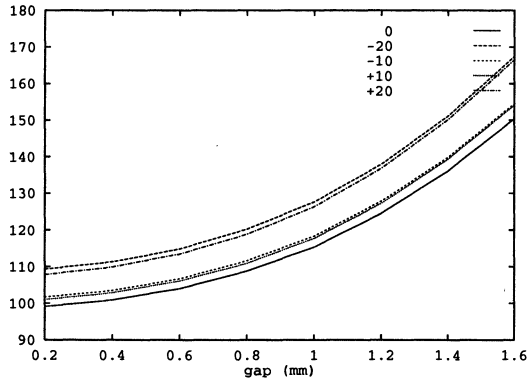


figure 6 B-310 Mode : Motional resistance R1 versus gap for various orientations of electrodes

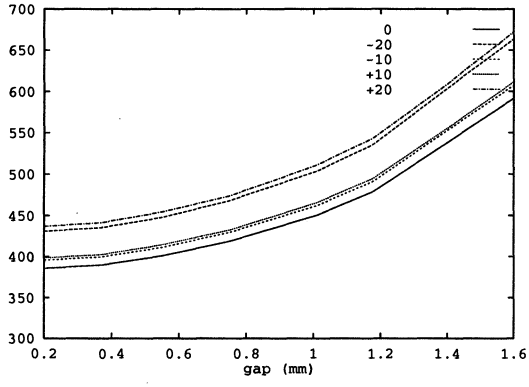


figure 7 C-501 Mode : Motional resistance R1 versus gap for various orientations of electrodes

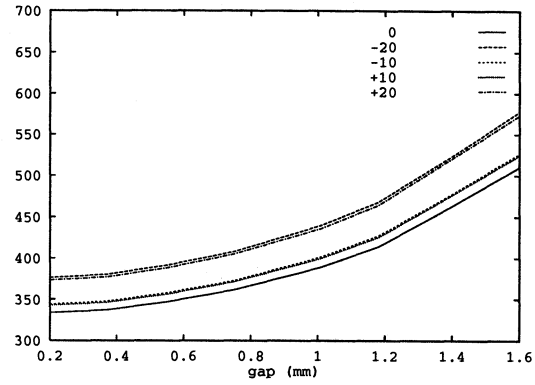


figure 8 B-501 Mode : Motional resistance R1 versus gap for various orientations of electrodes

The motional resistance is very sensitive to a shift of the mode in the direction perpendicular to the gap. This fact is illustrated by the figure 9 for the 301-modes when the width of the gap is equal to 1.2 mm. This effect can explain higher values of the observed motional resistance than those expected from the calculus. We may notice that the B-mode is most sensitive to this effect than the C-mode, which is a positive point since such a defect tends to reduce its electrical response.

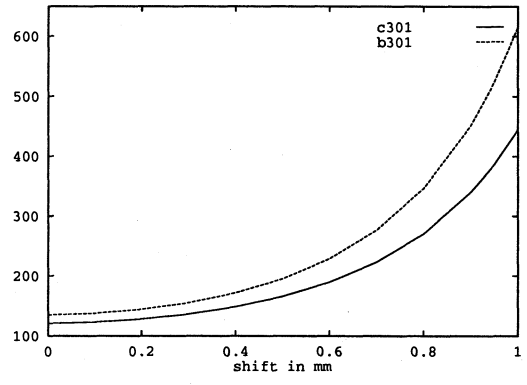


figure 9 Variation of R1 when the mode-shape is shifted in the direction perpendicular to the gap (301 modes)

Motional parameters of a resonator operating in 301 modes are indicated in table II for several values of gap-width. If the motional capacitance C_1 and the static capacitance C_0 decrease when the gap increases, the capacitance ratio remains quite constant: it varies from 6.70 to 6.82 10^{-5} in the case of the C-mode and from 6.58 to 6.11. 10^{-5} for the B mode

mode	gap(mm)	C_1(fF)	C_0(pF)	R_1(Ω)
C 301	0.2	0.165	2.46	96.8
	0.8	0.150	2.12	105.9
	1.2	0.131	1.92	120.9
B 301	0.2	0.161	2.45	98.6
	0.8	0.142	2.12	112.1
	1.2	0.118	1.93	134.6

Table II: Motional parameters.

A resonator with a width of gap of 1.2 mm has a B-mode resistance which is about 20% higher than that of the C-mode. This difference is sufficient to allow to build an oscillator without any selective circuit. The short term stability (Allan variance) obtained at this time is reported on the table III.

τ(s)	0,1	1	2	4	8	16	32
$\sigma_y(\tau)$ (10^{-13})	9,2	6,7	4,8	4,3	4	5,7	7,2

Table III : Short term stability of a resonator operating in C-301 mode.

ACKNOWLEDGEMENTS

The authors would like to thank M. Mourey for his contribution in performing short term stability measurements.

REFERENCES

1 - **R. BOURQUIN, B. DULMET, J.J. BOY:** "SC-Cut resonator operating in anharmonic modes with B-mode reduction", 10th EFTF - Brighton (1996), p 239-243

2 - **J.A. LEWIS** : "The effect of driving electrode shape on the piezoelectric properties of piezoelectric crystals", BSTJ, Sept. 1961, p 1259-1280.

3 - **E. DURAND :** "Electrostatique" TII Masson - Paris - 1966

4 - **D.S. STEVENS, H.F. TIERSTEN** : "An analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variation", J.A.S.A., vol. 79, 6, p 1811-1826

5 - **B. DULMET, R. BOURQUIN** : "3D analysis of thickness vibrations in doubly rotated contoured quartz plates", IEEE 1991 Ultrasonics Symposium, p 463-468.

A PRELIMINARY INVESTIGATION INTO A NOVEL APPROACH FOR THE REMOVAL OF DRIVE LEVEL DEPENDENCE IN QUARTZ CRYSTAL RESONATORS

R.C. Chittick
C-MAC Quartz Crystals Ltd., UK

D.M. Hyland
EPAK Electronics Ltd., UK

1. ABSTRACT

A unique solid carbon dioxide based cleaning system has been assessed as a process for removing drive level dependence (DLD) in quartz crystal resonators.

Trials have been conducted on mass base plated blanks selected for high levels of DLD. Initial results have shown significant improvements over untreated controls, in some cases reducing the number of failures in a batch from 47% to 0%.

The procedure has not yet been optimised and this preliminary investigation has been carried out using a simple manual process employing a hand held applicator. However, the system easily lends itself to automation, thus removing the uncontrolled variables associated with manual operation.

Key words: drive level dependence, solid carbon dioxide.

2. INTRODUCTION

The drive level dependence (DLD) or second level of drive at low applied power, observed in some quartz crystal resonators, has been attributed to the effects of surface defects and contamination by sub-micron particles (Ref. 1,2,3). The problem can be alleviated by using finer abrasives for lapping the blanks, followed by chemical etching and subsequent processing employing good clean room practice to avoid contamination.

However, the problem has never been entirely eliminated and it has become common to include an additional procedure to remove DLD. This usually involves driving the units at much higher than the normal working power (Ref. 1) in order to dislodge any particulate contamination or other material such as quartz fragments and badly adherent metallisation. While this treatment has proved effective, there has always been some concern over possible damage to the resonators and the permanence of the effect.

We have been assessing an alternative technique for reducing DLD using equipment on loan from EPAK Electronics Ltd. This is the Micro-Sno™ cleaning system (Ref. 4), manufactured by Deflex Corporation which employs a combination of Thermal-Ionised-Gas (TIG), which may be dry nitrogen, carbon dioxide or air, and solid carbon dioxide (Snow), to produce a variable high velocity jet that can dislodge sub-micron contaminants.

This paper describes the TIG-Snow technology and the results of a preliminary investigation into its use in reducing DLD in quartz crystal resonators.

3. THE TIG-SNOW SYSTEM

The TIG-Snow cleaning technology was developed by the Deflex Corporation, largely in response to the increasing need to move away from traditional CFC solvent based cleaning systems. It differs from conventional solid carbon dioxide based systems in that the snow is combined with a filtered, heated and ionised, inert gas at the applicator nozzle which can accelerate the snow to supersonic velocities.

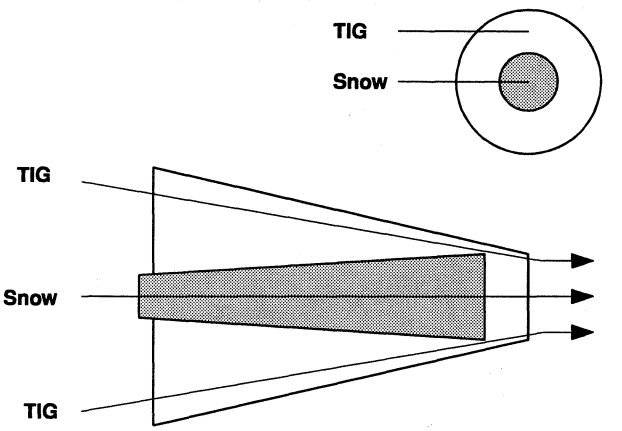


Figure 1. TIG-Snow spray nozzle.

Figure 1 shows the nozzle assembly. The TIG, which may be air, nitrogen or carbon dioxide gas (for our tests high purity cylinder nitrogen was used), flows in an outer flexible tube. Liquid carbon dioxide is expanded into gas and solid state and transported through an inner tube to the nozzle where it is combined with the TIG and ejected at high velocity. The spray pattern may be either convergent or divergent and a number of nozzles are supplied that produce different patterns and spray velocities. The pressure of the snow spray is boosted by the TIG by an amount dependent on the TIG pressure as shown in Figure 2. This is for a 3:1 TIG:Snow nozzle ratio.

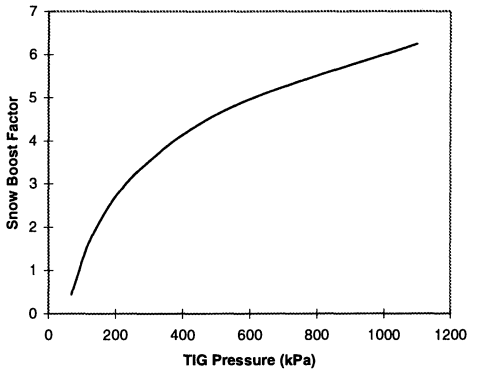


Figure 2. Snow pressure boost by energy transfer from TIG

The snow concentration is controlled by a timer operated valve that allows pulses of variable length to be applied and this in conjunction with the heated TIG allows the user to control the temperature of the surface to be cleaned.

The cleaning action of this system is both physical and chemical. The very high snow velocities achieved can overcome electrostatic and Van der Waal's forces that bind small particles to the surface. There may also be an effect resulting from charged carbon dioxide particles colliding and combining with small charged particles on the surface which are then swept away by the snow jet. Although the mechanism is not fully understood, organic residues on a surface can also be removed by the snow (Ref. 4). This may possibly be due to dissolution of the of the organic material in liquid phase carbon dioxide at the surface.

4. APPLICATIONS IN QUARTZ CRYSTAL MANUFACTURE

It is clear that the TIG-Snow technology may be applied to quartz crystal manufacture in its intended role as a cleaning agent, for example, as an alternative to wet processes for cleaning crystal blanks. The purpose of this investigation, however, was to assess the potential of the system for removing contamination likely to cause drive level dependence.

Some preliminary tests were carried out to establish a suitable operating window for the TIG-Snow parameters (pressure, TIG temperature, Snow pulse time etc.), taking into account the fragility of the crystals, material usage and effectiveness of the treatment. Since a minimum pressure of 240 kPa is required to activate the TIG heater, the operating pressure was kept within the range 240 kPa to 275 kPa. The TIG temperature was set at 60° C and Snow pulse times used were between 0.2 and 1 second. The applicator was held at a distance of about 50 mm from the crystal and at an angle of 45°. No attempt was made to optimise the TIG-Snow parameters in this exercise, since the manual application procedure precludes the possibility of precise control of position and direction of the jet.

The method used for assessing DLD is that described by P. E. Morley et al. (Ref. 5), using a Hewlett Packard HP4195A Network/Spectrum Analyzer which displays a plot of resistance (ESR) against applied power over the range -56 dBm to +4 dBm into 50 ohms. Figure 3 is a plot of resistance over the range -27 dBm to -1 dBm for a 1 MHz fundamental unsealed crystal, showing the behaviour typical of DLD due to surface contamination. The maximum drive level has been restricted, so as not to remove any particulate contamination by the high drive effect. Figure 4 is the resulting plot for the same crystal after a one second application of the TIG-Snow jet to each side of the crystal. All trace of DLD has been removed and the minimum resistance has been reduced from 192.7 ohms to 143.2 ohms.

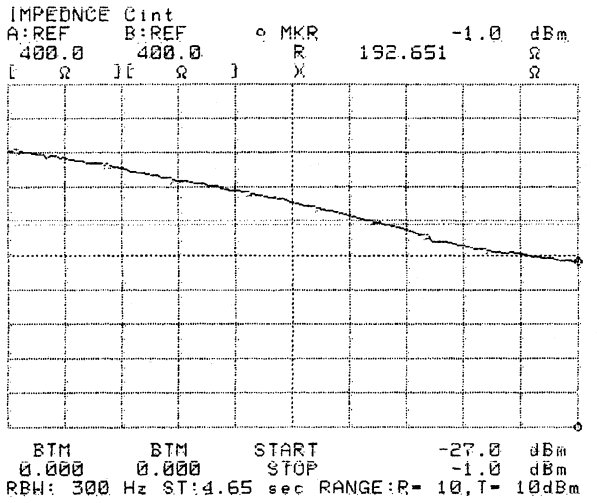


Figure 3. Plot from HP4195A showing drive level dependence of ESR in a bad unit.

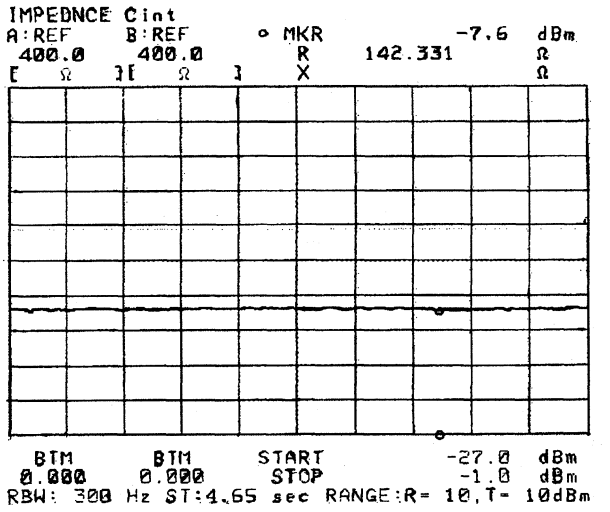


Figure 4. Plot of ESR of above unit after TIG-Snow treatment.

Trials have been carried out at two stages of manufacture, immediately after mass base plating and after the crystals have been mounted in their bases. There are advantages and disadvantages at both of these stages. After plating, the crystals, still in the plating masks, are supported peripherally and are therefore less likely to be broken on application of the TIG-Snow jet than is the case for mounted crystals that are supported only at two edges. However, as the plating stage occurs early in the manufacturing process subsequent re-contamination could occur. This is less likely to happen if the treatment is carried out on mounted crystals before frequency

adjusting and sealing. Of course, in the case of some product types, such as flat-pack resonators and Dual-in-Line (DIL) oscillators, in which only one side of the mounted crystal is available for TIG-Snow treatment, the only option is to carry out the procedure after mass base plating.

4.1 Treatment After Mass Base Plating

Two types of crystal were used in this test, flat blanks intended for use in DIL oscillators and plano-convex blanks for flat-pack resonators. Both groups had a high incidence of DLD. In the former, this had been the result of a faulty plating procedure. The plano-convex blanks had been lapped with a coarse abrasive resulting in sub-surface damage.

Figures 5 to 7 and Tables 1 to 3, compare untreated controls with TIG-Snow treated flat crystals, measured for DLD after mass base plating. DLD is expressed as the difference between the maximum and minimum resistance over the power range. In this case the failure criterion is a DLD of 7 ohms and the mean DLD and ESR given are the values for those that passed the test. In all other tests reported here the failure levels are those recommended by CECC 68 000 Issue 1. The 16 MHz crystals (Figure 6 and Table 2) gave the best result with 100% recovery. The 6% no activity (NA) failures in the 20 MHz crystals in Figure 7 may indicate that the TIG pressure was excessive for this thickness and size (7.5 mm) of crystal and consequently some damage has been done to these units. It is encouraging to note that both the mean DLD and ESR for all three groups have been reduced by the TIG-Snow treatment.

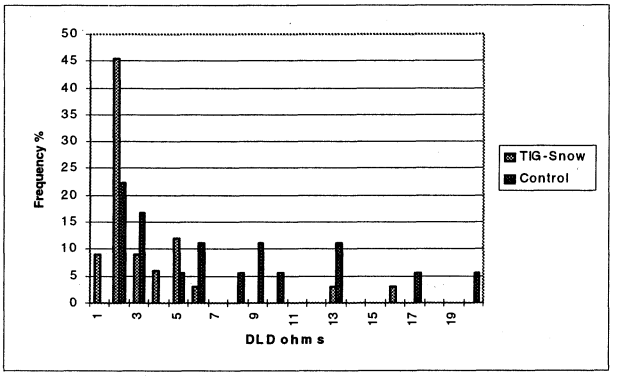


Figure 5. Distribution of DLD for treated and untreated 11 MHz DIL crystals.

	Control	TIG-Snow
UNITS	18	33
DLD Failures %	44	9
ESR Failures %	22 (all DLD)	3
No Activity %	0	3
Total Failures %	44	15
Mean DLD ohms	2.9 +/-1.6	2.2 +/-1.3
Mean ESR ohms	13.2 +/-5.4	10.4 +/-3.4

Table 1. Failure statistics for treated and untreated 11 MHz DIL crystals.

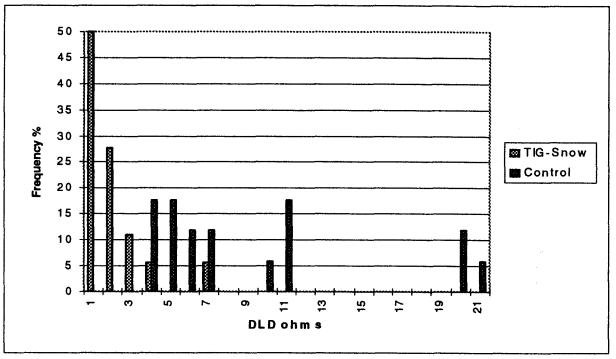


Figure 6. Distribution of DLD for treated and untreated 16 MHz DIL crystals.

	Control	TIG-Snow
UNITS	17	18
DLD Failures %	47	0
ESR Failures %	0	0
No Activity %	0	0
Total Failures %	47	0
Mean DLD ohms	4.7 +/-1.1	1.6 +/-1.3
Mean ESR ohms	8.9 +/-0.9	7.4 +/-1.0

Table 2. Failure statistics for treated and untreated 16 MHz DIL crystals.

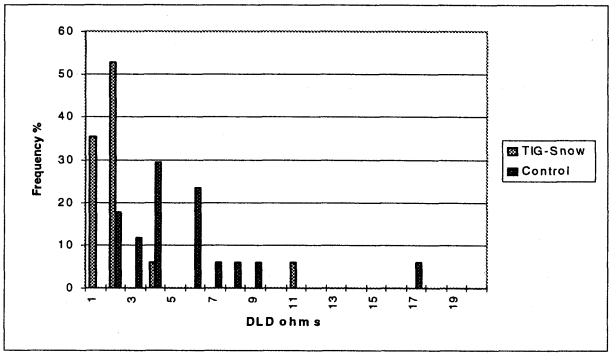


Figure 7. Distribution of DLD for treated and untreated 20 MHz DIL crystals.

	Control	TIG-Snow
UNITS	18	18
DLD Failures %	17	6
ESR Failures %	0	0
No Activity %	0	6
Total Failures %	17	12
Mean DLD ohms	3.7 +/-1.6	1.3 +/-0.7
Mean ESR ohms	7.6 +/-1.0	6.9 +/-1.2

Table 3. Failure statistics for treated and untreated 20 MHz DIL crystals.

The results of the second group are given in Figure 8. This was the largest trial, involving 749 units (12.8 MHz). Two slightly different TIG-Snow conditions were used here.

a) 0.1 second Snow pulse.
b) 0.2 second Snow pulse.

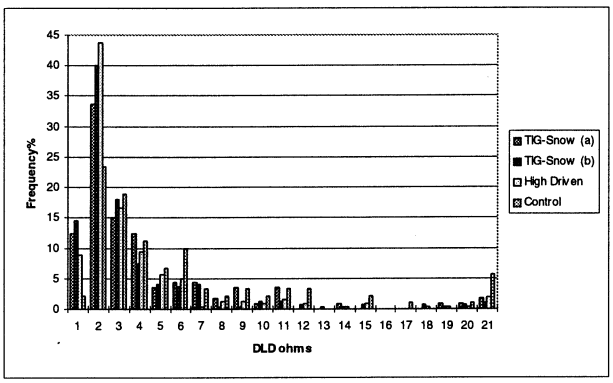


Figure 8. Distribution of DLD for treated and untreated 12.8 MHz contoured crystals, showing comparison with high driven units.

	Control	TIG-Snow (a)	TIG-Snow (b)	High Driven
UNITS	85	113	305	246
DLD Fail%	18.8	6.2	5.6	5.7
ESR Fail%	2.4	0.9	1.0	0.4
No Activity%	0	0	0.6	0.8
Total Fail%	21.2	7.1	7.2	6.9
Mean DLD ohms	3.5 +/-2.2	2.8 +/-2.2	2.5 +/-1.9	2.6 +/-1.9
Mean ESR ohms	10 +/-2.9	11.2 +/-1.6	10.8 +/-1.2	11.2 +/-1.6

Table 4. Failure statistics for treated and untreated 12.8 MHz contoured crystals, showing comparison with high driven units.

The DLD measurement has been carried out on completed resonator units and these are compared with untreated controls and units that have been high driven before final adjust and sealing. There is very little difference between the effectiveness of the two TIG-Snow conditions. The results of the high drive treatment are also similar but it is worth remembering that the TIG-Snow procedure was carried out after mass base plating and consequently many opportunities exist for re-contamination of the crystals between this and the stage at which the high drive was carried out.

4.2 Treatment After Mounting

Treatment after mounting the crystals in their bases was restricted to the more robust lower frequency crystals. Tables 5 and 6 give the results of tests on 6.4 MHz and 13 MHz plano-convex units with a poor surface finish. The DLD measurements were carried out immediately after TIG-Snow treatment which in this case consisted of a 2 second Snow pulse on the 6.4 MHz units and a 1 second pulse on the 13 MHz units.

	Control	TIG-Snow
UNITS	10	35
DLD Failures %	10	0
ESR Failures %	0	0
No Activity %	0	0
Total Failures %	0	0
Mean DLD ohms	3.1 +/-0.8	2.1 +/-0.4
Mean ESR ohms	19.2 +/-2.5	21.7 +/-2.5

Table 5. Failure statistics of treated and untreated 6.4 MHz contoured crystals (mounted).

	Control	TIG-Snow
UNITS	10	10
DLD Failures %	40	0
ESR Failures %	0	0
No Activity %	0	0
Total Failures %	40	0
Mean DLD ohms	3.2 +/-1.1	1.3 +/-0.2
Mean ESR ohms	15.4 +/-1.0	15.5 +/-1.3

Table 6. Failure statistics of treated and untreated 13 MHz contoured crystals (mounted).

4.3 Stability of TIG-Snow Treated Crystals

The effectiveness of TIG-Snow technology in reducing DLD in quartz crystal resonators has been clearly demonstrated. However, of more importance, is the stability of the effect and whether treated crystals will recover their drive level dependence after a period of inactivity, or worse become "sleeping" crystals. Table 7 gives the results of a group of TIG-Snow treated 13 MHz contoured crystals, measured immediately after sealing and after 51 days. There is no observed degradation, however the number on test is small and this needs to be repeated with a much larger batch.

	After Sealing	After 51 days
UNITS	10	10
DLD Failures %	0	0
ESR Failures %	0	0
No Activity %	0	0
Total Failures %	0	0
Mean DLD ohms	1.3 +/-0.4	1.0 +/-0.2
Mean ESR ohms	7.9 +/-1.3	7.0 +/-1.2

Table 7. Stability of sealed contoured 13 MHz crystals.

4.4 Mechanisms of DLD Removal

The physical mechanisms involved in the removal of DLD by TIG-Snow are not fully understood and are dependent on the causes of the DLD. For example, it would seem reasonable to suppose that if the DLD was the result of sub-micron particulate contamination on the active area of the crystal, then the mechanism would simply be the removal of this by momentum transfer and shear stress.

The case of contoured crystals is not so clear. These exhibit a much higher incidence of DLD than flat crystals and it has been suggested that this may be related to the surface finish. Sub-surface damage resulting from lapping with a coarse abrasive, which is not completely removed by etching, can be in the form of micro-cracks that cause small regions of the surface to be almost detached (Ref. 2). The removal of these regions by the TIG-Snow would reduce DLD. This would also involve the removal of the plating at these weak points. Badly adherent plating also contributes to DLD. Thermal stress resulting from the rapid cooling effect of the Snow would weaken these areas allowing the snow jet to remove the faulty plating. The plano-convex crystals described in section 4.1. exhibit a frequency shift of about +10 ppm after treatment, which could be accounted for by the removal of 0.3% of the plating. Small areas, where the plating has been removed, can in fact be seen under the microscope as shown in Figure 9.

(a)

(b)

Figure 9. A treated (a) and untreated (b) crystal with plating on one side only, showing part of the electrode area in transmitted light.

This shows part of the electrode area of a contoured crystal, plated only on the contoured side and view in transmitted light a) before TIG-Snow treatment and b) after treatment. It is unlikely that this plating has been removed by abrasive action since solid carbon dioxide is extremely soft, having a hardness on the Moh scale of only 2. Therefore it must be assumed that the plating on these regions had a reduced adhesion.

5. CONCLUSION

Although the system has not been optimised for quartz crystal treatment, the TIG-Snow system has been shown to be effective in reducing the incidence of drive level dependence in resonator and DIL crystals. The simple manual implementation is not ideal as there are too many uncontrolled process variations such as distance and angle between the applicator and the crystal surface but even so, it has been shown possible to reduce the failure rate in one group of crystals with 47% failures to 0%. Further improvements would be expected by automating the system and synchronised pulses from two applicators, one on each side of the crystal would minimise stresses and enable higher frequency crystals to be treated without breakage.

The system is capable of removing very small particles from the crystal surface that may give rise to DLD and has been shown to reduce DLD in crystals with sub-surface damage. Although the effects of the latter type of fault have been minimised, long term stability tests will have to be done on a much larger scale than in this exercise, to be certain that the DLD will not return. The real solution to this, of course, is to make the blanks correctly in the first place.

6. REFERENCES

1. S. Nonaka, T. Yuuki and K. Hara, "The Current Dependency of Crystal Unit Resistance at Low Drive Level", Proc. 25th Ann. Freq. Cont. Symp., April 1971, pp 139-142.

2. J.E. Knowles, "On the Origin of 'Second Level of Drive' Effect in Quartz Oscillators", Proc. Ann. Freq. Cont. Symp., May 1975, pp 230-236.

3. V.E. Bottom, "The Behaviour of Quartz Resonators at Low Drive Level", Proc. 5th Quartz Crystal Conf., August 1983, pp 245-249.

4. D.P. Jackson, "Precision Cleaning Using Thermal-Ionized-Gas (TIG) Snow Cleaning Technology", Deflex Corporation Technical Paper.

5. P.E.Morley, R.J.T. Marshall and R.J. Williamson, "Crystal Parameter Measurement Using a Hewlett Packard HP4195A Network/Spectrum Analyzer", Proc. 3rd European Time and Frequency Forum, March 1989, pp 133-138.

SOME PRACTICAL CONSIDERATIONS ON THE PROBLEM OF INTERMODULATION DISTORTION AND POWER HANDLING CAPABILITY IN VHF AND MONOLITHIC CRYSTAL FILTERS

Branislav Milić, Branislav Martinović, Milka Samardžija

Institute "Mihajlo Pupin"
Quartz Crystal Division, Belgrade, Yugoslavia

ABSTRACT

The influence of several factors contributing to the improvement of In-Band and Out-Band Intermodulation Distortion and Power Handling Capability has been considered.
The effects were analyzed on 38 MHz, 40.032 MHz, 81.4 MHz, 97.8 MHz , and 109.35 MHz . V.H.F. crystal filters, as well as on 21.4 MHz Monolithic Crystal Filters.
The results that are presented fully justify the introduction on new and improved activities within technology.
Some problems regarding the measurement of Intermodulation Distortion were also considered.

1. INTRODUCTION

Over the past two decades, the constructors of new telecommunication devices faced to the fact that crystal filters behave as non-linear systems as results of non-linear effects in the crystal resonators. Non-linearities in crystal resonators can be considered as being of four types as it given in (Ref. 3). Historical background of the problems caused by non-linearity and drive power level sensitivity is presented very interesting in (Ref.2).
It is necessary for purposes of this paper to cite the definitions of intermodulation and intermodulation ratio (IMR). The intermodulation is mixing of two or more signals in non-linear device which results in the generation of new frequencies. These new frequencies are sums or differences integer multiples of original frequencies. If some of those IM-generated frequencies fall in the passband region of a filter, IM distortion occurs. Intermodulation distortion can be caused by input signals inside the filter passband (In-Band IM Distortion) or by signal outside the passband (Out-Band IM Distortion).
The level of intermodulation distortion is usually described by Intermodulation Ratio. The Intermodulation Ratio (IMR) is defined as the ratio of available power at the filter input (PTT), or either test tone to the power in specified intermodulation products, measured at the filter load (PIM).

IMR = PTT/PIM

At the consideration of intermodulation it is very important to define the value of drive input level (strain level). Generally two cases can be analyzed:
a) Low Strain Level
b) High Strain Level
By our paper which follows we will consider the influence of low strain level and high strain level on the level of intermodulation distortion of 3rd order.
On the basis of our experience we think that it is possible to divide the intermodulation distortion requirements into three groups:
I. Low level requirements.
II. Normal level requirement
III. High level requirements

We also tried to set the borders for those tree levels and for the cases of In-Band and Out-Band intermodulation distortion. For better definitions we suggest to adhere to the Intercept - Point (IP).

IP = Pin +IMD/2
Pin ⇒ Input Power (dB).
IMD⇒ Intermodulation Distortion (dB)

Low Level Requirements		Normal Level Requirements		High Level Requirements	
In-Band I.P. (dB)	Out-Band I.P. (dB)	In-Band I.P. (dB)	Out-Band I.P. (dB)	In-Band I.P. (dB)	Out-Band I.P. (dB)
0-10	0-20	10-20	20-35	20-30	35-50

Table 1

The enclosed table has the aim to give approximate orientation in the analysis of the requirement and achieved results. The discrete or monolithic realization, as well as LF or VHF crystal filter are not taken into consideration particularly. Over the last decade in Institute Mihajlo Pupin - Quartz Crystal Division a big effort was done in order to improve the quality of crystal filter due to intermodulation distortion. We had not ambition toward the theoretical considerations of non-linear characteristics of the crystal resonators. Our intention was, that on the basis of experiences of others, numerous references which consider the phenomenon of intermodulation, to give our own contribution to the practical realizations, presenting the achieved results and the way how we achieved them. Particularly, it was done in the field of improvement of the technology of crystal filters and design and technology of crystal resonators. Due to requirements of market, the main efforts were in the field of:
a) VHF crystal filters 50 - 120 MHz, In-Band and Out-Band IMD at low strain level and high strain level.
b) Monolithic filters 21.4 MHz. In-Band IMD at low strain level and high strain level.
c) Power handling was solved by design of crystal resonators. (Ref 3) give exelent results using SC, BT or IT cuts.

2. EXSPERIMENTAL RESULTS

In further review of our achieved ressults we will adopt, because of clarity, an approach to the analyssis of In-Band and Out-Band Intermodulation Disstortion, separatly.

2.1. In-Band Intermodulation Distortion In VHF Crystal Filters.

From four mentioned mechanisms (Ref. 5) which make non-linear effects in crystal resonator, the influence of surface effects has been observed. It has been observed the connection between IMD and defectives of the surface before or after evaporation, bad adhesion of the film, effect of evapora-

tion, damage of film by scratch. At the mechanical processing of crystal blanks on the surface the micro crack are made, which are expanded, then it come to breaking of the pieces of quartz which stay on the surface or it is easy connected whit it. In order to avoid or to alleviate this moment we have used the polished blanks, only.
For good adhesion of film the cleaning of the blanks before deposition of electrode is very important.
In that aim the U.V. cleaning has been introduced.
It is known that the particles of the dust are easily connected to the polished surface. To proceed from an assumption that in working room there are enough particles of the dust the stabilization of the crystal resonators has been introduced in vacuum chamber for drying after final adjustment in order to avoid the possibility of the soil. To proceed from the fact that on the surface of units there is the dust which generally contains the oil, the washing of crystal units has been used. Very good results have been obtained by adding Zonil (surface active substance) to the IDEX- solution for chemical adjustment of frequency (ref 13). The results are verified on the crystal filters 40.032 MHz, 81.4 MHz, 97.8 MHz, 109.35 MHz and many others.
Some of those results are shown on Fig. 1 to 4.

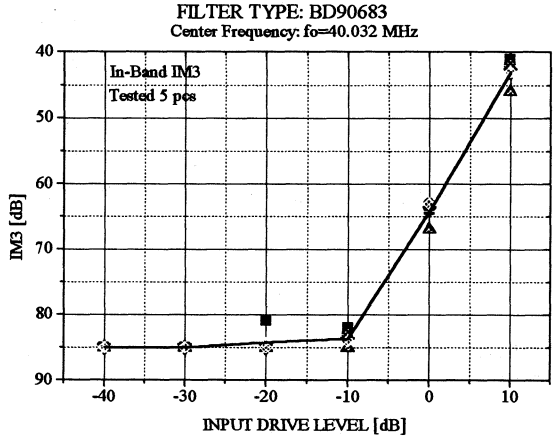


Figure 1

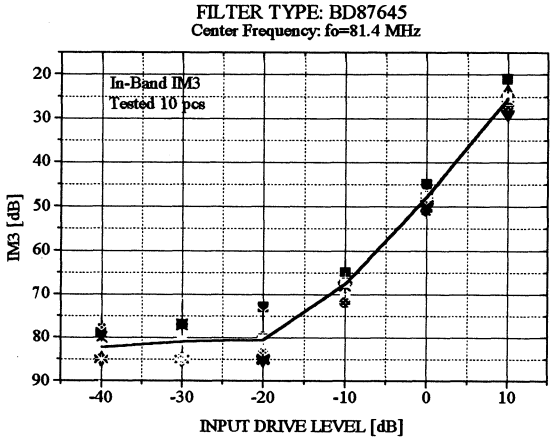


Figure 2

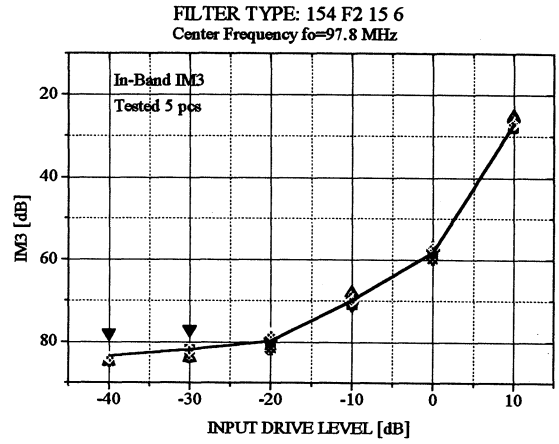


Figure 3

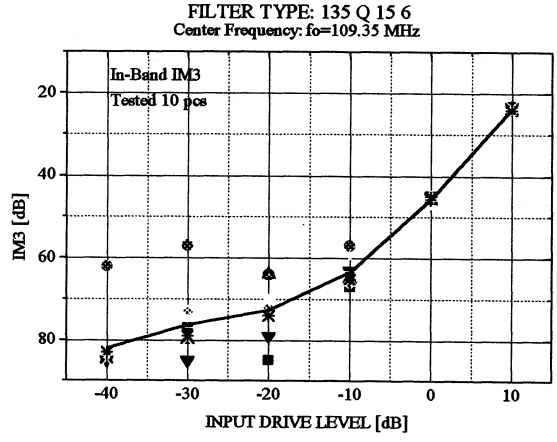


Figure 4

2.2 In-Band Intermodulation Distortion in Monolithic Crystal Filters

The monolithic crystal filter are since 1966 in very wide and numerous applications. From the origin solution on 10.7 MHz, the technology of production monolithic dual resonators has made a big progress and today the application of monolithic crystal filters is in the range of 5-160 MHz. The relative technological complexity of one elementary 2-pole monolithic crystal filter imposes the more severe technological requirements in its realization, and particularly relating than requirements of intermodulation distortion. The problem of intermodulation distortion in monolithic filter has been considered by several authors (Refs. 2, 3, 5, 8) where the 2-pole realozation of monolithic crystal filterhave been considered. To our opinion, the statements given in the section 2.1. can be applied completely also at monolithic filters. The reasons of genesis of the intermodulation distortion, as well as the way of their elimination are the same.
To achieve the results in the category of high requirement (Table 1) is more difficult according to our experience. The obtained results are, on the basis of our experience, few reliable and so-called "sleeping sickness" is more expressed then in the case of VHF crystal filters.
In IMP-QCD we analyzed the monolithic crystal filter 21.4 MHz 4 poles, 8 poles and 10 poles.
The obtained results are given in Fig.5 to 7.

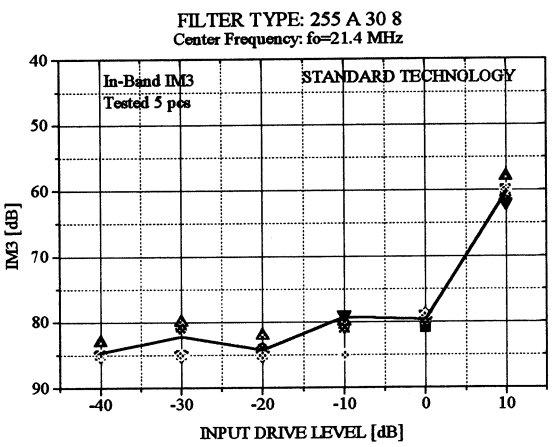


Figure 5

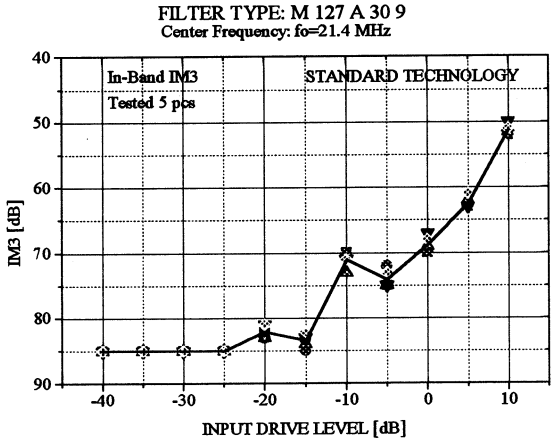


Figure 6

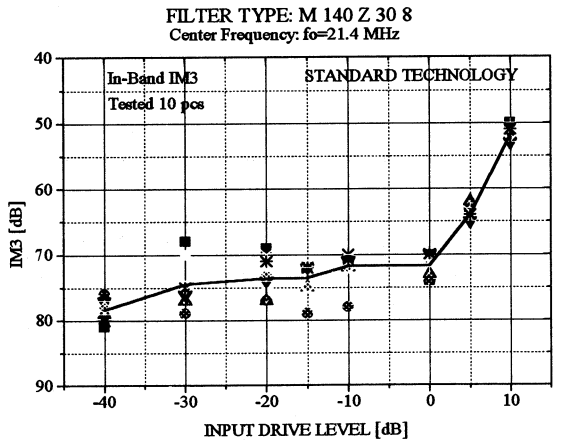


Figure 7

Monolithic filter 10-pole, type M 140 Z 30 8 has very high requierments at low strain level .By using standard technology for production we could not achive good yeald. It forced us to introduce longer each time then usualy. The obtained results are given at fig. 8 to 9.

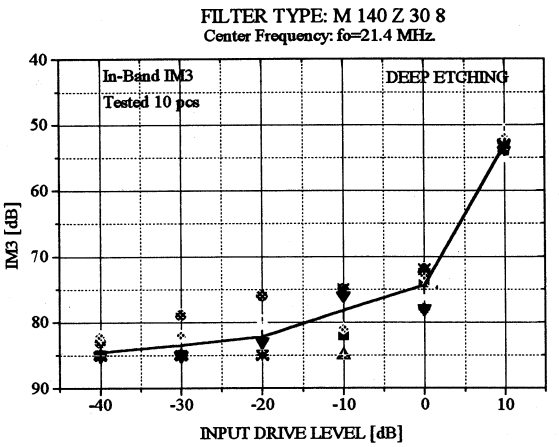


Figure 8

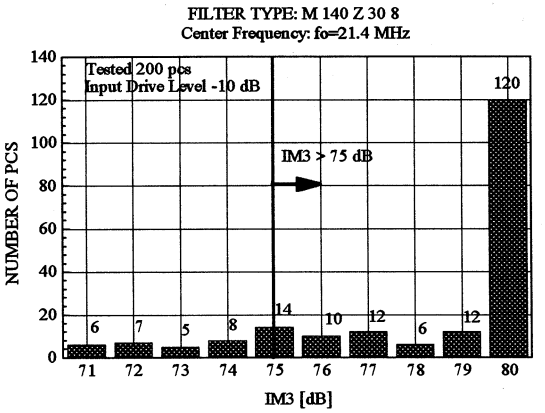


Figure 9

2.3. Out-Band Intermodulation Distortion in VHF Crystal Resonators and Filters

In the cited literature (Refs. 2, 5) doas not stand out the essential difference of the reason of intermodulation distortion for the case of Out-Band . Generally can be said: At low levels of strain the non-linearity causing IM3 , whether In-Band and Out-Band are predominantly related to surface defects. The obtained results are given in Fig. 10 to 11.

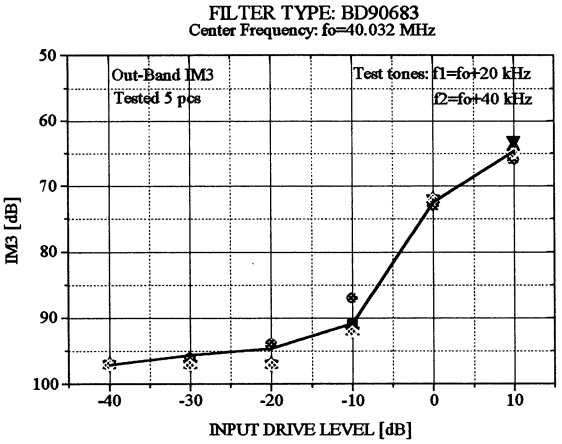


Figure 10

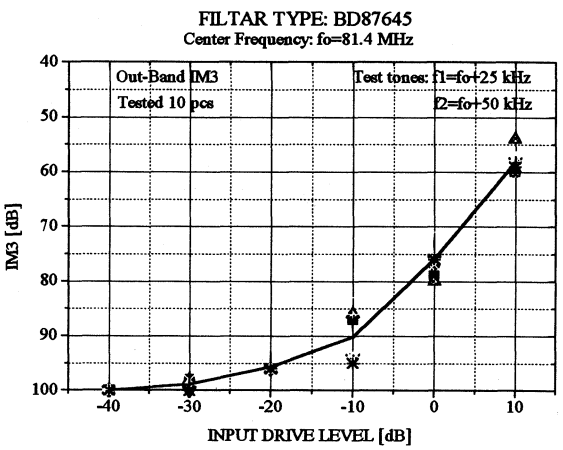


Figure 11

2.4. Out-Band Intermodulation Distortion in Monolithic Crystal Filters

The obtained results are given in Fig. 12 to 13.

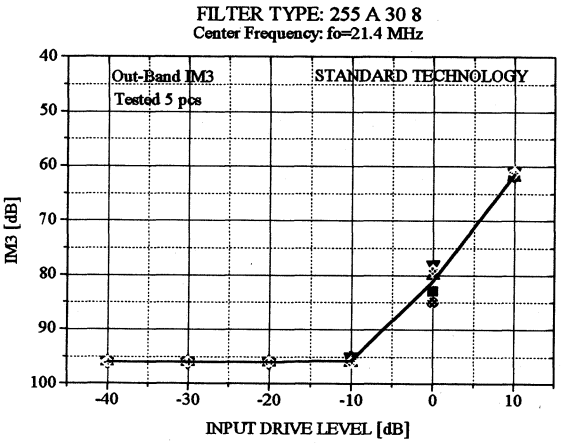


Figure 12

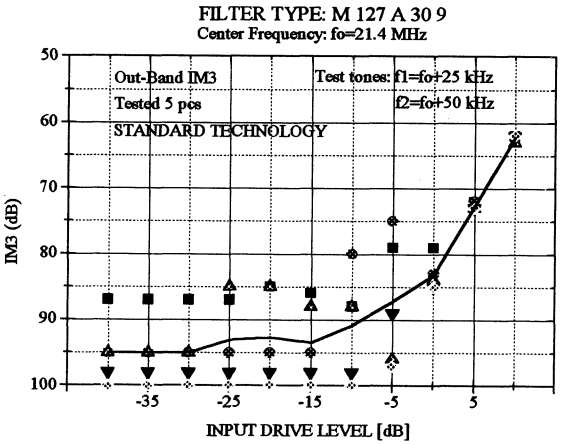


Figure 13

2.5. Conclusions

a) Accepting the hypothesis that low levels non-linear effects were related surface defects and depending upon the electrode film as well as the quartz, in IMP-QCD are placed the technological processes which have given in practice trough serial production of VHF crystal filters satisfactory results and very good yield.
b) The same hypothesis is applied also in the technological procedure of the realization of monolithic filter and results have justified this attitude. The complexity of the construction of dual resonators in the cases of the extreme requirements for intermodulation distortion at low strain level has forced us to introduce longer etch time then usually.
c) Experimentally, then , it appears at low strain level the non-linearity causing intermodulation whether In-Band or Out-Band, are predominantly related to surface defects of one kind or another .
d) The causes of intermodulation distortion of high strain levels are less easily determined, although the experience indicates the decreasing importance of surface conditions as strain level amplitude increases. Same experience indicates a highly reproducible mechanisms. The regularity of square function appears.
e) Also, same experience indicates the reduction of intermodulation distortion by increasing of the surface of electrode in the case of VHF crystal resonators.
f) In the case of monolithic filters there are the indication that the influence of coupling coefficient is minor.

3. MEASURING METHODS AND PROBLEMS

Measurement of intermodulation distortion is really very complex measurement. The basic instructions for the measurement of IM3 are given in (Ref. 9). We use the test arrangement as shown on fig. 14.

Test circuit for IM3 In-Band measurement

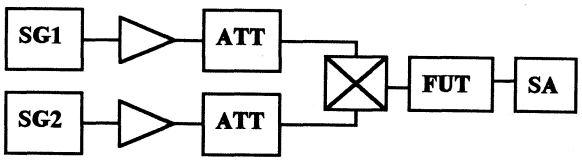


Figure 14

Test circuit for IM3 Out-Band measurement

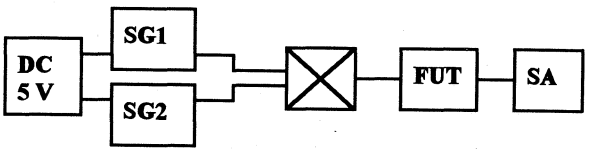


Figure 15

Where are:

SG1 and SG2	HP 8640B Signal Generators
▷	Amplifires
ATT	HP 355D Attenuators
⊠	HP 8721A (6dB) Splitter
FUT	Filter Under Test
SA	HP 8568A Spectrum Analyzer

(Refs. 10, 11) give the useful suggestions for improvement of the measuring methods. In the cases when exist the requirements for high level input in IMP-QCD, we use the specially constructed amplifiers. By using such test-set up the measurement is enabled:
a) In-Band IM3 up to 85 dBm.
b) Out-Band IM3 up to 100 dBm.
The tolerances of measurement is ± 5 dBm.
For the purpose of control of the production of quartz resonators and dual resonators we introduced the testing of IM3 requirements for resonators and dual resonators according to the suggestion in (Ref. 8). For each type of filters the correlation of requirements is placed. For IM3 for filter and corresponding IM3 requirements for resonators or dual resonators can be set. It was shown that such correlation contributes exceptionally to the high percentage of crystal filters yield.

4. POWER HANDLING

The new generations of Military H.F. radios is expected to demand even more stringent IM3 performance from crystal filters. Also, in order to improve exiter signal-to-noise ratios it is desirable to have as much signal amplification as possible before filtering. To achieve these objectives filters capable of handling higher power levels while still operating in a linear manner are needed. In (Ref. 3), for the first time was given the review of the solution of this problem by using SC, BT or IT cut and for drive levels between 0 dB and 25 dB.
In IMP-QCD this problem was considered in concrete cases, being solved exceptionally by special design of crystal resonator. It was shown that the influence of the surface of electrode is very important, dominant for the solution of stated problem. Very good results are obtained in the field of VHF resonators and filters for up-conversion for 75 MHz, 81.4 MHz and 38 MHz.
Conclusion: The problem of power handling at crystal filters for up-conversion can be solved by increasing of surface of electrode achieving of necessary compromise for spurious response requirements.

5. SUMMARY AND DISCUSSION

In IMP-QCD the problem of intermodulation distortion and power handling capability has been considered exclusively through the experiments. For the purpose of IM requirements for In-Band and Out-Band , the technological procedure of the production of VHF crystal resonators and dual resonators for 21.4 MHz has been changed . We introduced in this treatment a few new steps which have considerably contributed to the improvement of IM3, as well as to very good yield of the realized resonators.
In Table 2 is given the review of the obtained results in IMP-QCD for VHF and monolithic crystal filter. The review is based on the results obtained from serial production.
In our previous work, first of all experimental, unfortunately, besides all undertaken measurements, the following appearances were noticed and up to now they were not solved:
a) " Sleeping sickness". This appearances has been also noticed at VHF crystal filter, as well as 21.4 MHz monolithic filters. We can say that at monolithic filter is more outstanding.
b) The change of IM3, unfortunately worse, after thermal cycling of filter or after final temperature test. By checking of dual resonators , we stated that usually one resonator has considerably changed its IM3 characteristic.
c) The presence of histeresis effect, and at the change of testing course this appearance only occurs when the drive level is increasing.

No.	Center Freq. (MHz)	Type	Band-With (kHz)	Pol.	*In-Band* Input Level (dB)	*In-Band* IM3 (dB)	*In-Band* I.P. (dB)	*Out-Band* Input Level (dB)	*Out-Band* IM3 (dB)	*Out-Band* I.P. (dB)
Low Frequency Crystal Filters										
1	1.4	BD 89353	6	8	-20	-60	10	*	*	*
2	1.4	BD 90250	5.7	8	-25	-66	8	*	*	*
HF and VHF Crystal Filters										
1	8.4	SSB	2.5	8	-30	-60	0	*	*	*
2	40.032	BD 90683	12	4	-5	-70	30	-5	-88	39
3	41.4	161-1-12-6	12	6	-5	-60	25	-7	-82	34
4	70	251-Z-15-7	15	4	*	*	*	-7	-105	46
5	75	171-3-10-5	10	6	-13	-50	12	*	*	*
6	80.196	250-Z-5-5	5	4	-23	-65	9.5	*	*	*
7	81.4	166-1-12-6	12	6	-11	-65	22	-7	-82	34
8	97.8	154-F2-15-6	6.4	6	-6.5	-60	24	*	*	*
9	99	156-Q1-12- 6	12	4	-20	-70	15	-4	-75	34
10	109.35	165-K1-10-5	10	4	-7	-50	18	*	*	*
11	109.35	135-Q-15-6	15	4	-13	-50	12	*	*	*
12	115	89-K-14-6	14	4	-13	-45	9.5	*	*	*
13	144.456	208-13-20-7	20	4	-10	-60	20	0	-75	38
Monolithic Crystal Filters										
1	21.4	255-A-30-8	30	4	-23	-65	9.5	-5	-68	29
2	21.4	M103-A-15-9	15	8	-10	-60	20	-10	-80	30
3	21.4	M127-A-30-9	30	8	-10	-60	20	-10	-80	30
4	21.4	M140-Z-30-8	30	10	-10	-75	28	*	*	*

Table 2

In the (Refs. 3, 4) the mathematics expression for IM3 in function of central frequency, passband width, overtone and surface of electrode and input level, are given. We think that areas for future experiments, for the purpose of improvement of crystal filter quality, must be:
a) Optimum choice of the electrode surface.
b) Influence of "overtone" at IM3.
c) Testing of the influence of base heating during the group evaporation.
d) Influence of deep etching at IM3.

6. ACKNOWLEDGMENT

We would like to express our appreciation to Mrs. Gordana Ivanović for her contributions to the design and production of Crystal and monolithic resonators and to Mr. Ćamil Zabeljaj for his talented design and construction of crystal filters.

7. LITERATURE

[1] Malinowski, S. & Smith, C. "Intermodulation in Crystal Filters", Proc. 26th Annual Frequency Control Symposium, pp 180-186, 1972.

[2] Dworsky, L. & Kinsman, G. "A Simple Single Model for Quartz Crystal Resonator Low Level Drive Sensitivity and Monolithic Filter Intermodulation" IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 41, No2, March 1994. pp 261-269

[3] Howard M. D., Smythe R. C & Morley "Monolithic Crystal Filters Having Improved Intermodulation and Power Handling Capability"

[4] Tiersten H. F., "Analysis of Intermodulation in Thickness-shear & Trapped Energy Resonators" J. Acoust. Soc. Am., Vol 57, No3, pp 667-681, March 1975.

[5] Horton W. H, & Smythe R. C. "Experimental Investigations in Monolithic Crystal Filters", Proc. 27th Ann. Freq. Control Symp., pp 243-245, 1973.

[6] Smythe R. C. "Intermodulation in Thickness-shear Resonators", Proc. 28th Annula Frequency Control Symposium, pp 5-7, 1974.

[7] Smythe, R. C. "Some Recent Asvances in Integrated Crystal Filters", Proc. 32th Annual Frequency Control Symposium, pp m220-230, 1978.

[8] Gordon - Smith D & Almond D. P. "Anomalous Nonlinearity in Quartz Crystal Filters", Electronics Letters, Vol 17, No 5, pp 207-208, March 1981.

[9] IEC - Standards
IEC-368-1 Piezoelectric Filters,
Part 1, General Information, standard values and test conditions

[10] R. Kinsman " Crystal Filters", J. Wiley & Sons, 1978.

[11] Third Order Intermodulation Characteristics
Signal Generator HP 8640/B
Application Note AN 170-2, Jan. 1975.

[12] Third Order Intermodulation Distortion Measurements and Downloadable Procedure for HP 8568 B Spectrum Analyzers, Product Note 8566B/8568B-1

[13] R. Brandmayr & J. Vig "Chemical Polishing in Enching Solutions That Contain Surfactants", Proc. 39th Annual Frequency Control Symposium, pp 276-281, 1985.

TIME AND FREQUENCY COMPARISONS

Chairman: P. Eskelinen

GLONASS TIME TRANSFER AND ITS COMPARISON WITH GPS

W. Lewandowski[1], J. Azoubib[1], M. Weiss[2], V. Zhang[2], V. Hanns[2], A. Gevorkyan[3], P. Bogdanov[3], N. Tutolmin[3], J. Danaher[4], G. de Jong[5], J. Hahn[6], M. Miranian[7]

[1]Bureau International des Poids et Mesures, Sevres, France
[2]National Institute of Standards and Technology , Boulder, Colorado, USA
[3]Russian Institute of Radionavigation and Time, St Petersbourg, Russia
[4]3S Navigation, Laguna Hills, California, USA
[5]NMi Van Swinden Laboratory, Delft, The Netherlands
[6]Deutche Forschungsanstalt fuer Luft und Raumfahrt
[7] United States Naval Observatory, Washington DC, USA

Abstract

Currently the most popular method of comparing remote clocks is to use the GPS and GLONASS satellite navigation systems. Comparisons via GLONASS signals were suspended for many years because full deployment of the system was delayed and there were no commercial time receivers. Currently, however, nine major timing centres are equipped with twelve GLONASS or GPS/GLONASS time receivers, and their number is increasing. This paper presents the results from GLONASS common-view time comparisons, obtained using a GLONASS receiver of type ASN-16 from the Russian Institute of Radionavigation and Time (RIRT) and a R-100 type of 3S Navigation, in four European and three US time laboratories, while following a BIPM international tracking schedule. The paper continues with the analysis of a comparison with GPS common views.

INTRODUCTION

This paper provides a tentative estimation of the precision of time comparisons by GLONASS common views, describes the main characteristics of the ASN-16 and R-100 receivers and presents the results of time comparisons between several laboratories in Europe and North America according to the BIPM international GLONASS schedule. Procedures for data processing and standards for GLONASS international common-view time transfer are also discussed. Since most laboratories participating in this exercise are also equipped with GPS, or GPS/GLONASS time receivers, we are able also to report comparisons of GLONASS and GPS common-view time transfer. GPS common-view time transfer is accomplished with an accuracy of several nanoseconds and gives an excellent reference with which to evaluate the ultimate performance of GLONASS common-view time transfer, if we exclude some systematic errors like those in receiver calibration. The present comparison shows a slightly lower performance for GLONASS time transfer. The causes of this, and possible solutions, are presented.

The use of GLONASS signals, which for time synchronization have characteristics similar to those of GPS, was restricted for a long time because there were no commercial time receivers. In late 1993, the Russian Institute of Radionavigation and Time (RIRT) completed the development of a GLONASS time receiver, satisfying BIPM requirements and based on its own airborne ASN-16 receiver. To obtain and process GLONASS time measurements automatically, an interface between the ASN-16 and a personal computer was built. Receivers to this specification are already in operation at the Russian State Time/Frequency Reference in VNIIFTRI, Mendeleevo, at the other Russian Time Service laboratories, and at the BIRM, Beijing, China. In mid-1995, 3S Navigation commercialized the GLONASS R-100 receiver in accordance with BIPM requirements. In 1996, a dual-system was produced by this company, the first commercial GPS/GLONASS time receiver. These receivers were installed at the BIPM, USNO, VSL, NIST, DLR and other laboratories, see Table 1. After the appearance of these special timing receivers, the BIPM published the first tracking schedule for international time and frequency comparisons by GLONASS common views. Regular measurements and data exchange between laboratories began on 4 January 1996. At present the third BIPM international GLONASS tracking schedule is being implemented. There are some differences between the two systems which have an impact on time transfer. These differences and possible ways of overcoming them are described below. The need for GLONASS and GPS/GLONASS time transfer data format standardization is also emphasized.

Table 1. Laboratories observing GLONASS in common-view and showing interest.

Laboratory	Equipment	Est. unct. of GLO ant. coord. in the ITRF /m
1. Laboratories observing GLONASS in common-view:		
BIPM (Sevres, France)	R-100/10 R-100/30 GPS opt	0,3
USNO (Washington D.C., USA)	R-100/10	0,1
NIST (Boulder, Colorado, USA)	R-100/30 GPS opt	
3S (California, USA)	R-100/10 R-100/30 GPS opt	10,0
RIRT (St.Petersburg, Russia)	ASN-16-01	10,0
VSL (Delft, Netherlands)	R-100/30 GPS opt	4,0
DLR (Oberpfaffenhofen, Germ.)	R-100/30 GPS opt	3,0
BIRM (Beijing, China)	ASN-16-02 R-100/30 GPS opt	
LDS (Leeds, UK)	Spot GPS opt.	
2. Laboratories in preparation or showing interest:		
VNIIFTRI (Mendeleevo, Russia)	ASN-16-02	
TL (Chung-Li, Taiwan)	R-100/30	
NPLI (New Delhi, India)	R-100/10	
IFAG (Wettzell, Germany)	R-100	
CSIR (Pretoria, South Africa)	R-100/30	
BNM-LPTF (Paris, France)	R-100/10	
NPL (Teddington, UK)	R-100/30	

TIME REFERENCES

One of the major differences between GPS and GLONASS is that they use different references for time and space. For time reference, GPS relies for its GPS time on UTC(USNO), Coordinated Universal Time (UTC) as realized by the USNO. GLONASS relies for its GLONASS time on UTC(SU), UTC as realized by the Russian Federation. UTC is produced by the BIPM and is the internationally recognized time reference for the whole Earth..

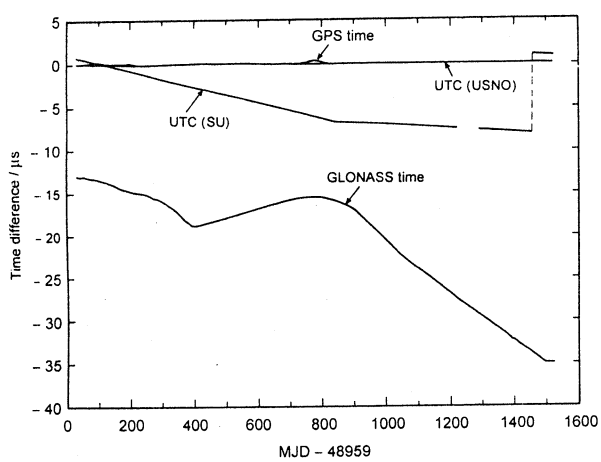


Figure 1: Deviation of UTC(USNO), UTC(SU), GPS time and GLONASS time from UTC from 2 January 1993 to 31 January 1997.

The deviation of UTC(USNO) from the UTC generally remains within 20 ns. In the Russian Federation, however, UTC(SU) was going away from UTC. The deviation, [*UTC-UTC(SU)*] approached minus 8000 ns in November 1996, see Figure 1. The GPS operators keep GPS Time within 100 ns from the UTC(USNO), and it has generally performed much better than this except for a period of about two weeks in December 1994 when, due to a malfunction, GPS time made an excursion from UTC(USNO) of about 270 ns. For GLONASS, the difference between GLONASS Time and UTC in January 1997 was around 35000 ns, and was increasing steadily. The 13th Meeting of the CCDS (Comité Consultatif pour la Définition de la Seconde) held on 12-13 March 1996, recommended (Recommendation S 4 1996, Ref. 1):

- that the reference times (modulo 1 second) of satellite navigation systems with global coverage* be synchronised as close as possible to UTC,
- that the reference frames for these systems be transformed to be in conformity with the terrestrial reference frame maintained by the International Earth Rotation Service (ITRF),
- that both GPS and GLONASS receivers be used at timing centres.

This Recommendation specifies a basis for harmonizing GPS and GLONASS, which does not make GLONASS depend on GPS, or GPS on GLONASS, but requires that both systems maintain their time and space references in agreement with international standards. Following this Recommendation, the Russian Federation agreed to improve the synchronization of its time scales with UTC. On 27 November 1996 a time step of 9000 ns was applied to UTC(SU) in order to make it approach UTC, see Figure 1. Next, on 10 January 1997 a frequency step was applied to GLONASS time to adjust its frequency to be close to that of UTC(SU). Further adjustments of these two time scales with respect to UTC are expected. This development is a sign of good will and understanding.

The values of [*UTC - GLONASS time*] reported on Figure 1 were published in BIPM Circular T until 31 December 1996, for the standard dates at 5-day intervals, according to observations of [*GPS time - GLONASS time*] performed at University of Leeds with a GPS/GLONASS receiver, 'Bart', designed and built in house. The same differences, derived from GLONASS observation at the BIPM with a R-100/10 receiver, differ by a constant of about 1250 ns, see Table 2. Similar differences were observed using other R-100 receivers at the VSL and the DLR. We know that R-100 receivers are not calibrated absolutely. We would prefer to trust the Leeds values, because there the primary observations are [*GPS time - GLONASS time*].

* Such as Global Positioning System (GPS), Global Navigation Satellite System (GLONASS), International Maritime Satellite Organization (INMARSAT), Global Navigation Satellite System 1 (GNSS1), Global Navigation Satellite System 2 (GNSS2).

Table 2. [*UTC - GLONASS time*] by BIPM Circular T and by a R-100/10.

[*UTC - GLONASS time*]			
Date 1995	by Circ. T /ns	by R-100/10 /ns	*Cir.T - R-100/10* /ns
Aug 10	-19812	-21055	1243
Aug 20	-20188	-21421	1233
Aug 30	-20549	-21813	1264

The hardware delay should be the same for GPS and GLONASS signals and should cancel in the difference. However the values of [*GPS time - GLONASS time*] provided by the new GPS/GLONASS receiver 'Slot' built at the Leeds University differ from that provided by 'Bart' by 200 ns (Ref. 2). Data from a similar comparison with GLONASS data recorded with a ASN-16-02 receiver show another shift. Russian receivers are not calibrated absolutely, so also do not provide an independent check.

Because GLONASS receivers are not calibrated absolutely we know [*UTC - GLONASS time*] with an accuracy no better that 1 microsecond. GPS receivers are absolutely calibrated and [*UTC - GPS time*] is known with an accuracy of a tens of nanoseconds, mainly because of SA. It follows that GLONASS provides an average user with world-wide real-time dissemination of UTC, as produced by the BIPM, to an uncertainty no better than 1 microsecond after the recent improvement of the synchronization between UTC(SU) and UTC. GPS does the same with uncertainty lower than 100 ns.

Since 1 January 1997, BIPM Circular T has provided daily values of [*UTC - GLONASS time*] derived from the observations of a R-100/30 receiver located at the VSL, Delft, the Netherlands. Until GLONASS receivers are calibrated absolutely it was decided to apply a constant bias of 1285 ns to the GLONASS observations obtained at the VSL, to maintain continuity with the previously published observations from Leeds University.

Summing up, we note that persisting differences between Russian time scales broadcast by GLONASS and UTC affect real-time dissemination of UTC through GLONASS, and to some extend complicates the dual system GPS/GLONASS navigation solution. However, this discrepancy does not affect common-view time transfer as readings of the satellite clock vanish in the difference. Also lack of absolutely calibrated GLONASS receivers can be easily overcome for common-view time transfer by differential calibration of receivers.

REFERENCE FRAMES

It is now general practice for laboratories engaged in accurate GPS time transfer to express ground-antenna coordinates with decimetric uncertainties in the ITRF, the internationally recognized ultra-accurate terrestrial reference frame. Post-processed GPS precise ephemerides are also expressed in the ITRF. Although actual GPS broadcast ephemerides are expressed in the WGS 84, the newest realization of this reference frame agrees to within one decimetre with the ITRF. In these conditions the uncertainties in the GPS ground antenna coordinates can have an impact on the accuracy of GPS common-view links of no more than 1 nanosecond. To sum up: the use of coordinate reference frames for GPS time transfer is well defined, well practised and fulfils all recommendations and standards.

The situation is somewhat different for GLONASS. Its geodetic reference frame PZ-90 can differ by up to 20 m from ITRF on the surface of Earth and no accurate relationship between PZ-90 and ITRF is yet known. The simplest way to determine GLONASS antenna coordinates is to average a series of navigation solutions. But the uncertainties of such coordinates are no better than several metres, see Table 1, and can have an impact on the accuracy of the common-view link of a few tens of nanoseconds. First users of GLONASS time receivers decided ad-hoc to determine the GLONASS ground-antenna coordinates in the ITRF wherever this was possible (Ref. 3). This has the obvious advantage of immediately providing the best possible consistency between the GLONASS ground-antenna coordinates. At present the ITRF antenna coordinates introduced into R-100 type receivers are transformed to the PZ-90 reference frame to make them consistent with broadcast ephemerides. All R-100 type receivers now in operation use the same transformation formulae and parameters. Russian ASN-16 receivers do not transform antenna coordinates. In order to harmonize GLONASS with GPS and avoid operations on ultra-precise ITRF coordinates, another approach would be to keep the ITRF coordinates unchanged in the receiver and transform broadcast ephemerides from PZ-90 into ITRF according to a standardized set of formulae and parameters. This solution was submitted to the CCDS Sub-group on GPS and GLONASS Time Transfer Standards (CGGTTS) at its last meeting on 2 December 1996 (Ref. 4).

The uncertainty of GPS broadcast ephemerides ranges from 5 m to 15 m. It is not affected by Selective Availability (SA). The uncertainty of GLONASS broadcast ephemerides is estimated to be slightly larger and should range from 15 m to 20 m. Post-processed precise ephemerides for GPS satellites are publicly available with a delay of several days. Their uncertainty is currently several decimetres, but is improving progressively. At present, no information is available on the computation of GLONASS post-processed precise ephemerides, but some geodetic civil institutions are already considering international efforts to compute GLONASS precise ephemerides for public use. We should also expect that future GLONASS post-processed precise ephemerides will be expressed in ITRF coordinates. In the present study we did not apply GPS precise ephemerides. In this way the GPS time links were computed under the same conditions as GLONASS links.

IONOSPHERIC DELAY

GPS C/A-Code time receivers correct time observations for ionospheric delay using a model based on broadcast parameters. As the GLONASS navigation message does not contain ionospheric parameters, GLONASS C/A-Code time receivers use a model based on fixed parameters. This limits the precision of the long-distance GLONASS links studied in this paper, because these links necessarily use low satellite tracks. Ionospheric measurements using two GPS frequencies are performed at the BIPM, the NIST and the USNO. To perform GPS time links under the same conditions as GLONASS, GPS ionospheric measurements were not, therefore, used in this study. Commercial double-frequency GLONASS time receivers are now available and these are capable of measuring ionospheric delay. Their use should greatly improve the quality of GLONASS measurements.

BRIEF DESCRIPTION OF GLONASS TIME RECEIVERS

Table 1 lists laboratories which observe GLONASS according to the tracking schedule for international time and frequency comparisons by GLONASS common views and laboratories which have expressed interest in using GLONASS common-views. The ASN-16, designed by the RIRT, is a one-channel, one-frequency unit designed for airborne navigation (Ref. 5). When used for time determination, it provides, via one chosen satellite, an output of 1 Hz synchronized to GLONASS time. That is why, for time comparisons via GLONASS signals using the ASN-16 receiver, an additional time intervallometer is necessary. To eliminate the need for this instrument the ASN-16 receiver was redesigned to provide a time difference with an external signal of 1 Hz. In this form, the ASN-16 receiver is designated ASN-16-02 and, through an interface, it provides fully automated measurements to a PC. The uncertainty of time determination between the user clock and the satellite clock by this receiver should not be worse than 60 ns (rms) after its absolute calibration. Tests of two ASN-16-02 receivers at the RIRT show that uncertainty of GLONASS zero-baseline common-view time comparisons is not worse than 10 ns (rms) for averages including 15 or more tracks per day.

Receivers of type R-100 are manufactured by 3S Navigation. The R-100/10 receiver is also a one-channel, one-frequency, C/A-code unit. It provides time differences between the user clock and the satellite clock with an uncertainty not worse than 60 ns (rms) and common-view time comparisons with an accuracy of a few nanoseconds (rms) when calibrated relatively. The R-100/30 receiver is a multi-channel, two-frequency, two-system GPS/GLONASS, instrument which uses P-code for GLONASS and C/A-code for GPS. It provides independent measurements for each channel, and for GLONASS accounts for ionospheric delays by the two-frequency technique. The uncertainty of time determination between the user clock and satellite clock is not worse than 60 ns (rms), if we exclude errors in receiver calibration, and the accuracy of common-view time comparisons is a few nanoseconds (rms) for differentially calibrated receivers.

The ASN-16 and R-100 receivers are both controlled by a PC and use a standard format developed for GPS common-view technique by the CCDS Group on GPS Time Transfer standards (Ref. 6). The R-100 receivers also use the standard formulae and parameters adopted for GPS. The ASN-16-02 receiver does not yet follow these standards, but their full implementation is expected soon.

RESULTS OF GLONASS COMMON-VIEW TIME TRANSFER

Ten time links on baselines ranging from zero to 9000 km we consider for this study (Ref. 7). We show that the baseline length affects the precision and accuracy of satellite common-view time transfer. The greater the distance, the larger is the effect of uncertainties in the satellite ephemerides and ionospheric delay on time transfer. However, uncertainties of the antenna coordinates, see Table 1, may add a major contribution to the accuracy and precision of the common-view link even over a short baseline.

Table 3. Tentative theoretical uncertainty budgets for GLONASS common-view time comparisons.

Component	Uncertainty /ns (for 1 track)		Uncertainty /ns (for 30 tracks)	
	1000 km	9000 km	1000 km	9000 km
Satellite clock error (cancel in CV mode)	0	0	0	0
antenna coordinates	60	60	11	11
Satellite ephemerides	4	40	1	7
Ionosphere (day time, normal solar activity, elevation > 20 deg.)	2	30	1	5
Troposphere (elevation > 20 deg.)	2	2	2	2
Instrumental delay (relative)	2	2	2	2
Receiver software	2	2	2	2
Multipath propagation	5	5	1	1
Receiver noise (13-min average)	3	3	1	1
Total	61	78	12	14

Table 3 gives a tentative theoretical uncertainty budget for GLONASS time comparisons in common-view mode, at distance *d*, for C/A-code receivers, for one 13-minute track and for an average of 30 tracks over one day. In making these calculations it is assumed that: the noise of the laboratory clocks and the rise time of the reference pulses are negligible; ground antenna

coordinate uncertainties are of the order of 10 m; ephemeride uncertainties range from 15 m to 20 m (Ref. 8); a model with fixed parameters is used to determine the ionospheric delay.

Table 4. Precision determination of GLONASS common-view links.

Common-view links	Base - line	Estimated precision /ns Date 1996/97								
		M	J	J	A	S	O	N	D	J
BIPM(100/30)-BIPM(100/10)	0	4	4	4	4	4	-	-	-	-
BIPM - VSL	400	-	-	-	-	9	10	10	11	9
VSL - DLR	400	8	7	9	8	8	8	9	10	7
BIPM - DLR	500	9	9	8	9	8	8	8	8	7
RIRT - VSL	2100	13	13	16	15	12	12	13	11	10
RIRT-VSL*	2100	-	16	15	17	17	-	14	14	-
BIPM - RIRT	2200	-	-	-	-	7	7	7	6	7
BIPM -USNO	6000	-	-	-	-	-	6	8	10	11
BIPM - NIST	8000	-	-	-	-	7	8	10	10	-
BIPM - 3S	8400	-	-	-	-	8	-	-	12	13
RIRT - 3S	11000	-	11	-	18	16	-	-	-	-

* Computed by RIRT

Table 4 shows the results of precision determination of GLONASS common-view time comparisons between clocks in some laboratories noted above, for intervals of one month.

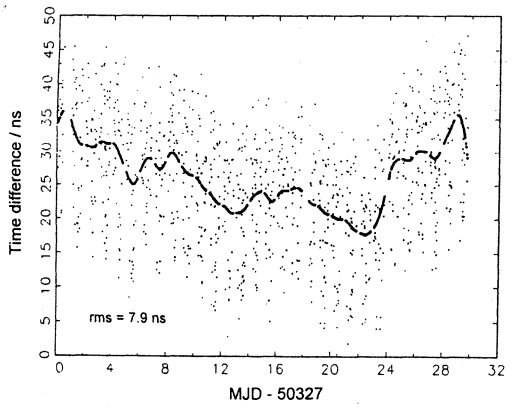


Figure 2. [*UTC(DLR) - UTC(VSL)*] plus a constant, by GLONASS common views.

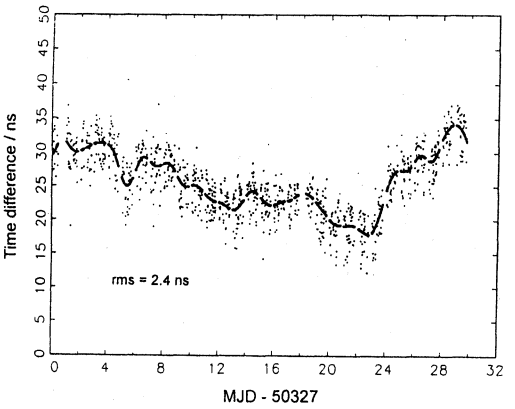


Figure 3. [*UTC(DLR) - UTC(VSL)*] plus a constant, by GLONASS common views after removal of the biases.

We speak here about precision and not accuracy, as many sources of systematic effects could not yet be resolved for GLONASS links. We have chosen to express the precision of a single 13-minute GLONASS common-view measurement in terms of the root mean square (rms) of the differences between raw and smoothed values. The data analysis covers the nine month period in which the 1st, 2nd and 3rd BIPM international GLONASS schedule were implemented. From 7 to 62 GLONASS common views were available daily. Vondrak smoothing (Ref. 9), which acts as a low-pass filter with cut-off periods ranging from about 1 day for a 0 km baseline to about 10 days for a 9000 km baseline, was performed on the raw GLONASS common-view values. This range of cut-off periods was chosen as representing, approximately, the limit between short time intervals, for which measurement noise is dominant, and longer intervals, for which clock noise prevails. The number of common views per link and cut-off periods are listed in Table 5. At the RIRT the method of least squares interpolation was employed, using a linear model for time differences and one day averaging. The link RIRT - VSL computed using the RIRT approach is also reported (marked * in Table 4). Estimating of precision derived from two methods are similar.

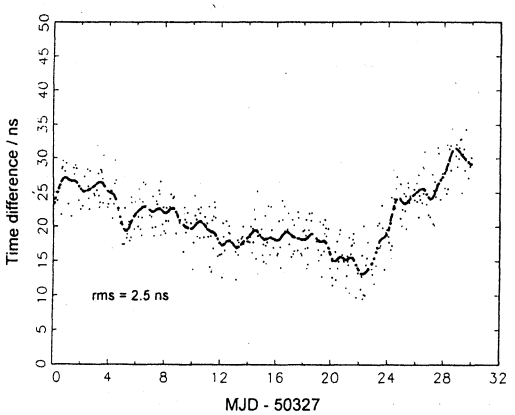


Figure 4. [*UTC(DLR) - UTC(VSL)*] plus a constant, by GPS common views.

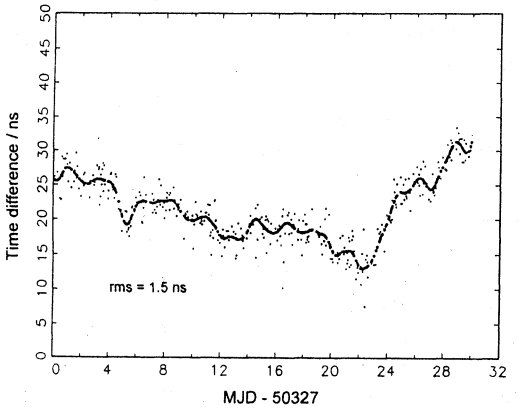


Figure 5. [*UTC(DLR) - UTC(VSL)*] plus a constant, by GPS common views after removal of the biases.

Table 5. Number of common views per link and cut-off periods for smoothing.

ommon-view links	Base-line /km	Number of CV per day	Cut-off period /day
BIPM(100/30)-BIPM(100/10)	0	41	1
BIPM - VSL	400	27	1 to 2
VSL - DLR	400	41	1
BIPM - DLR	500	38	1 to 2
RIRT - VSL	2100	16	5 to 6
RIRT-VSL*	2100	14	
BIPM - RIRT	2200	18	1
BIPM - USNO	6000	8	4 to 5
BIPM - NIST	8000	14	4 to 5
BIPM - 3S	8400	12	4 to 5
RIRT - 3S	11000	7	8 to 10

At the BIPM a procedure to remove constant biases between observations in different directions of the sky is used operationally for the treatment of GPS data. It has been shown for GPS common views that, for baselines up to 1000 km, these constant biases are mostly due to errors in the differential coordinates of the laboratories involved. We have chosen the link DLR - VSL to illustrate the use of this procedure for GLONASS common views. Figure 2 shows the common views before removal of the biases, and Figure 3 shows the same views after removal of the biases. The rms is reduced from 7,9 ns to 2,4 ns. This is a strong indication that the differential coordinates between these two laboratories have an error of several metres. In fact we already know (see Table 1) that the GLONASS antenna coordinates at the DLR and VSL have errors of several metres in the ITRF. The reasons of expressing GLONASS antenna coordinates in the ITRF reference frame are explained in detail above and in Ref. 3.

COMPARISON OF GLONASS AND GPS COMMON-VIEW TIME TRANSFERS

To evaluate the performance of the GLONASS common-view method we computed five time links for a period of five months using the GPS common-view method for baselines ranging from zero to 8000 km. To estimate the quality of GPS links we used the same statistical approach as for estimation of the precision of GLONASS links. However, in case of GPS we can speak about accuracy as all known sources of systematic effects were considered and as much as possible removed. The results are given in Table 6. We note that for a baseline of zero, the behaviour of GPS and GLONASS receivers is similar. We remark that the effect of baseline length on the accuracy of GPS common-view time transfer is similar that on GLONASS. The greater the distance, the larger is the effect on time transfer of uncertainties in the satellite ephemerides and ionospheric delay. We note, however, that the precision of GLONASS links is slightly lower than the GPS ones. This is easily explained by problems with GLONASS antenna coordinates, and slightly lower quality of GLONASS broadcast ephemerides and ionospheric correction.

Table 6. Estimated accuracy of GPS common-view links.

Common-view links	Base -line /km	Estimated accuracy / ns Date 1996/97				
		Sep	Oct	Nov	Dec	Jan
BIPM(TTR5)-BIPM(TTR6)	0	3	3	3	3	3
BIPM - VSL	400	2	2	3	3	2
VSL - DLR	400	3	3	3	2	2
BIPM -USNO	6000	5	5	5	4	5
BIPM - NIST	8000	6	5	6	5	4

In Table 7 we show a comparison of GPS and GLONASS common views between DLR and VSL. There is a constant shift of 324 ns between the two methods, partly due to the use of uncalibrated GLONASS and GPS receivers and partly to the less accurate geodetic coordinates available for GLONASS. When constant shift is removed from the difference between the GPS and GLONASS results, the values obtained are strikingly low, generally 1 ns. Figures 4 and 5 illustrate the removal of biases from GPS observations. The slight improvement, from 2,5 ns to 1,7 ns rms, is due to an error of about 0,5 m in differential coordinates between these two laboratories.

Table 7. Comparison of GPS and GLONASS common-view time transfer for August and September 1996 at five-day intervals.

[UTC(DLR)-UTC(VSL)]				
MJD	by GPS /ns	by GLONASS /ns	*GPS - GLONASS* /ns	*GPS - GLONASS* -324 /ns
50299,0	1862	1539	323	-1
50304,0	1860	1535	325	1
50309,0	1874	1549	325	1
50314,0	1887	1564	323	-1
50319,0	1892	1568	324	0
50324,0	1897	1572	325	1
50329,0	1905	1580	325	1
50334,0	1908	1581	327	2
50339,0	1912	1587	325	1
50344,0	1911	1587	324	-1
50349,0	1917	1591	326	1
50354,0	1906	1581	325	1

CONCLUSION

1) The appearance of commercial timing receivers of types ASN-16-02 from the RIRT (Russia) and R-100 from 3S Navigation (USA) has made it possible to begin regular international time transfer using GLONASS.

2) Persisting differences between the time scales broadcast by GLONASS and UTC and the absence of absolutely calibrated receivers, affect real-time dissemination of UTC through GLONASS but do not affect common-view time transfer as the readings of the satellite clock vanish in the difference, and differential calibration of receivers is easy to perform.

3) The results show that the precision of GLONASS common-view time comparisons is of the order of a few nanoseconds (rms) for distances of up to 1000 km, and

of the order of ten nanoseconds for intercontinental distances. This is comparable with the performance of GPS measurements.

4) The overall accuracy of GLONASS time links is inferior to that of GPS. Improvements will be made possible by:

- determination of accurate ground-antenna coordinates in the ITRF,
- use of post-processed precise ephemerides,
- double-frequency measurement of ionospheric delay,
- differential calibration of GLONASS receivers,
- adoption of standardized software,
- keeping the antennas in constant temperature enclosures (Ref. 10).

ACKNOWLEDGEMENTS

The authors are pleased to express their gratitude to their colleagues in the BIPM, RIRT, 3S Navigation, VSL, DLR and NIST for their participation in the execution of measurements and the processing and analysis of data.

REFERENCES

[1] Report of the 13th Meeting of the CCDS, BIPM, 1996.

[2] Daly P., Personal communication, 1997.

[3] Lewandowski W., Azoubib J., Gevorkyan, A.G., Bogdanov P.P., Klepczynski W.J., Miranian M., Danaher J., 1996, A Contribution To the Standardization of GPS and GLONASS Time Transfers, Proc. PTTI'96, in press.

[4] Report of the Open Forum on GPS and GLONASS Standardization, 5th CGGTTS Meeting, 2 December 1996, Reston, Virginia, USA.

[5] Gouzhva Y.G., Gevorkyan A.G., Bogdanov P.P., Ovchinnikov V.V., Full-Automated System for Receiving and Processing GLONASS Data, Proc. of ION GPS-94 7th Intern. Techn. Meeting, pp.313-316, 1994.

[6] Allan D.W. and Thomas C., 1994, Technical Directives for Standardization of GPS Time Receiver Software, Metrologia, **31**, pp. 67-79.

[7] Lewandowski W., Azoubib J., Gevorkyan A.G., Bogdanov P.P., Danaher J., de Jong G., Hahn J., 1996, First Results From GLONASS Common-View Time Comparisons Realized According to the BIPM International Schedule, Proc. PTTI'96, in press.

[8] GLONASS Interface Control Document, 1990, GLAVKOSMOS.

[9] Vondrak J., 1969, A Contribution to the Problem of Smoothing Observational Data, Bull. Astron. Inst. Czechoslovakia.

[10] Lewandowski W., Moussay P., Guerin P., Meyer F., Vincent M., 1996, Testing Motorola Oncore GPS Receiver and Temperature Protected Antennas for Time Metrology, Proc. PTTI'96, in press.

MODIFIED RECEIVING AND DATA PROCESSING SYSTEM FOR GLONASS COMMON-VIEW TIME COMPARISONS

A.Gevorkyan ,P.Bogdanov , O.Nechaeva, G.Travin, N.Tutolmin

Russian Institute of Radionavigation and Time (RIRT), Russia

In late 1993 in Russian Institute of Radionavigation and Time (RIRT) timing receiver named ASN-16-02 was developed on the base of its own serial GLONASS ASN-16 receiver. It meets the requirements of Bureau International des Poids and Mesures (BIPM) Time Section. To obtain the execution of measurements and their processing automatically the interface between receiver and personal computer was held. This set of devices was named as Receiving and Data Processing System (RDPS).

At the same time many platforms of the Russian Time Service (RTS) continue to use for GLONASS time comparisons the serial ASN-16 receiver with additional time interval measurer of the difference between receiver internal time scale and time scale of user. To provide the execution of measurements and their processing automatically in this case also the additional interface modul and develope of software was required.

Modified Receiving and Data Processing System on the basis of ASN-16 type receivers provides:
executing of GLONASS time measurements according to schedule, formed by BIPM, by RTS and/or independently for given user;
statistical processing of measurements for various variants and conditions of their execution;
data exchange with other users (almanac, tracking schedule and results of measurements in the BIPM format and RTS format) using E-mail;
forming an additional information about presence of satellites in visible zone for given user or network of users;
graphic displaying of measurements processing results and other information.

The developed system is successfully maintaining in the structure of the econdary Time/Frequency Reference of RIRT and provides GLONASS common-view time comparisons with accuracy about 15 ns.

IMPROVED CLOCK SYNCHRONISATION WITH LOW FREQUENCY REAL-TIME DIFFERENTIAL GPS

J. Hahn, H.Nau

DLR, Institut für Hochfrequenztechnik, Postfach 1116, D-82230 Oberpfaffenhofen,
E-mail: joerg.hahn@dlr.de, hartmut.nau@dlr.de

1. ABSTRACT

To support real-time precise navigation and positioning a low frequency real-time differential GPS service is provided by the German Telekom and the Institute for Applied Geodesy. The broadcasted from Telekom station in Mainflingen pseudo-range corrections are influenced by the reference receiver clock offset to GPS system time. A time user equipped with a low frequency real-time differential GPS time receiver can use this property for improved clock synchronisation. The paper presents DLR's first experimental results with LF-DGPS time receivers. The described method could be of interest for the timing community in the European and other regions.

Keywords: *Synchronisation, GPS, LF-DGPS, Pseudo-Range and Range Correction, Time Receiver*

2. INTRODUCTION

Low frequency real-time differential GPS (LF-DGPS) is provided by the German Telekom and the Institute for Applied Geodesy. Using the signals transmitted by GPS satellites the accuracy of the position measurement is about 100 meters. LF-DGPS delivers real time correction data and improves the accuracy up to better than 5 meters. Trial measurements all over Germany lead to errors of 1-3 meters in the position measurement. Of course, these errors are dependent on the distance of the user from the reference station, cf. [DID-95] and [GPS-97].

The mentioned LF-DGPS service is based on a reference station located at the Telekom broadcasting station in Mainflingen (near Frankfurt/M.) the coordinates of which are well known. The data are broadcast at 123.7 kHz in the Radio Data System format (RDS). The advantages of the low frequency are its large coverage (larger than Germany), small radio shadows, small and inexpensive receivers, the possibility of integration of the LF-receiver into the GPS receiver, and modern broadcasting und modulating methods resulting in high quality in reception and guarantee high reliability. , cf. [DID-95] and [GPS-97].

The broadcasted from Telekom station in Mainflingen pseudo-range corrections are influenced by an offset consisting of the difference between the GPS system time and reference receiver clock. A time user equipped with a LF-DGPS time receiver can use this property for improved synchronisation to the reference clock. If the generated pseudo-range corrections are compensated for the clock offset of the reference clock in Mainflingen with respect to GPS system time or an UTC(xx) (range-correction), these time scales could be disseminated in a very accurate manner. This has no effect to the navigation user.

3. THEORY

3.1 Formulation of the pseudo-range correction

The signal pseudo-propagation time $\rho_R^j(t_p)$ of an electromagnetic signal between a satellite j and a reference station R at time t_p can be formulated as:

$$\begin{aligned}\rho_R^j(t_p) &= \left\{\psi_R^j + dt^j - dt_R\right\}_{t_p} + \\ &+ \left\{\begin{array}{l} dt_{R,eph}^j + dt_{R,SAeph}^j + dt_{SAclock}^j + \\ + dt_{R,ion}^j + dt_{R,trop}^j + dt_{R,mp}^j + \\ + \varepsilon_R^j \end{array}\right\}_{t_p} \\ &= \left\{\psi_R^j + dt^j - dt_R\right\}_{t_p} + dt_{R,re}^j(t_p)\end{aligned} \tag{3.1}$$

with

ψ_R^j - signal propagation time due to the geometric distance between satellite j and reference station R in [s];

dt_R - clock offset of the reference receiver R with respect to system time in [s];

dt^j - clock offset of the satellite with respect to system time in [s];

$dt_{R,eph}^j$ - propagation delay caused by the ephemerides error of the satellite j with respect to reference station R in [s];

$dt^j_{R,SAeph}$ - error due to selective availability (SA) - satellite *j* ephemerides degradation with respect to reference receiver *R* in [s];

$dt^j_{SAclock}$ - error due to SA clock degradation of satellite *j* in [s];

$dt^j_{R,ion}$ - ionospheric delay on the propagation path between satellit *j* and reference station *R* in [s];

$dt^j_{R,trop}$ - tropospheric delay on the propagation path between satellite *j* and reference station *R* in [s];

$dt^j_{R,mp}$ - propagation delay due to multipath effects on the propation path between satellite *j* and reference station *R* in [s];

$dt^j_{R,re}$ - remaining delays on the propagation path satellite *j* - reference station *R* in [s];

ε^j_R - reference receiver *R* noise in [s].

The system time is concerned with a certain navigation system. In Eq.(3.1) the delay $\dot{\psi}^j_k dt_R$ due to the motion of the satellite and receiver during the signal propagation is not considered.

The geometric propagation delay can be easily determined as the difference of the known position vectors of the satellite and reference receiver:

$$\psi^j_R = \frac{\sqrt{(x^j - x_R)^2 + (y^j - y_R)^2 + (z^j - z_R)^2}}{c} \qquad (3.2)$$

The correction term $\Delta^j_R(t_p)$ as the difference between the pseudo- and geometric propagation time at time t_p can be written as

$$\begin{aligned} \Delta^j_R(t_p) &= \left\{\rho^j_R - \psi^j_R\right\}_{t_p} \\ &= \left\{dt^j - dt_R + dt^j_{R,re}\right\}_{t_p} \end{aligned} \qquad (3.3)$$

The term $c \cdot \Delta^j_R$ is called *pseudo-range correction* for the satellite *j* with *c* - speed of light in vacuum. Note that the value of this term depends on the certain reference station *R* used, because $dt_R(t_p)$ is included.

3.2 Using pseudo-range corrections in a time receiver

We assume that the above defined pseudo-range correction is available for a time receiver *k* at time t_{p+1}. It is subtracted from the determined pseudo propagation time $\psi^j_k(t_{p+1})$ of the time receiver *k* which is defined following Eq.(3.1) substituting the index *R* by *k*:

$$\begin{aligned} (\rho^j_k(t_{p+1}))_c &= \left\{\psi^j_k + dt^j - dt_k + dt^j_{k,re}\right\}_{t_{p+1}} - \Delta^j_R(t_p) \\ &= \psi^j_k(t_{p+1}) + dt^j(t_{p+1}) - dt^j(t_p) - \\ &\quad - dt_k(t_{p+1}) + dt_R(t_p) + dt^j_{k,re}(t_{p+1}) - \\ &\quad - dt^j_{R,re}(t_p) \end{aligned} \qquad (3.4)$$

The value of dt_k indicates the clock offset of the internal time receiver clock with respect to system time, $(\rho^j_k)_c$ is the corrected signal pseudo-propagation time between satellite *j* and receiver *k*.

It is assumed that $\Delta^j_R(t_p)$ can be corrected for $dt^j(t_{p+1})$. Thus, the clock offset *B'* as the 1 pps output of the time receiver can be calculated with

$$\begin{aligned} B'(t_{p+1}) &= dt_R(t_p) - dt_k(t_{p+1}) \\ &= \left\{(\rho^j_k)_c - \psi^j_k\right\}_{t_{p+1}} - \\ &\quad - dt^j_{k,re}(t_{p+1}) - dt^j_{R,re}(t_p) \end{aligned} \qquad (3.5)$$

and shows that a time user *k* using this output gets synchronised with the reference receiver clock.

The reference to the system time of the navigation system is lost now. For a navigation user this is not important.

3.3 Formulation of the range correction

If we assume that the reference receiver clock error $dt_R(t_p)$ is known we can define the so called *range correction* $(\Delta^j_R(t_p))_c$ with

$$(\Delta^j_R(t_p))_c = \left\{\Delta^j_R + dt^j_R\right\}_{t_p} \qquad (3.6)$$

The range correction is independent on the reference receiver clock offset. To generate this value, a highly stable frequency input and an accurate

clock modelling with respect to system time has to be realised on the reference station.

3.4 Using range corrections in a time receiver

Now we apply the range correction $(\Delta^j_R(t_p))_c$ to Eq.(3.4) and get the signal *corrected pseudo-propagation* time $(\rho^j_k(t_{p+1}))'_c$ with

$$\begin{aligned}(\rho^j_k(t_{p+1}))'_c &= \left\{\psi^j_k + dt^j - dt_k + dt^j_{k,re}\right\}_{t_{p+1}} - \\ &\quad -(\Delta^j_R(t_p))_c \\ &= \left\{\psi^j_k - dt_k + dt^j_{k,re}\right\}_{t_{p+1}} - \\ &\quad -dt^j_{R,re}(t_p)\end{aligned} \tag{3.7}$$

again assuming that $\Delta^j_R(t_p)$ can be corrected for $dt^j(t_{p+1})$. In this case the offset *B"* is equal to the clock offset of the 1 pps output of the time receiver with respect to system time

$$\begin{aligned}B''(t_{p+1}) &= -dt_k(t_{p+1}) \\ &= \left\{(\rho^j_k)'_c - \psi^j_k\right\}_{t_{p+1}} - \\ &\quad -dt^j_{k,re}(t_{p+1}) + dt^j_{R,re}(t_p)\end{aligned} \tag{3.8}$$

The time user gets again synchronised with the system time.

3.5 Resulting remaining delay

The *resulting remaining delay* $(dt^j_{R,k})_{re}$ for a clock synchronisation of a time receiver *k* using pseudo-range and range corrections from a reference station *R* for a satellite *j* will be briefly discussed here.

The delays for the time receiver *k* are defined in the same way as for the reference receiver *R* in Eq.(3.1).

It is assumed that the correction terms coming from the reference station *R* are updated after a certain time (i.e. 8 sec) and transmitted by a LF broadcasting station. Thus, we have a problem of the latency and age.

The value $(dt^j_{R,k})_{re}$ can be written as

$$\begin{aligned}(dt^j_{R,k})_{re} &= dt^j_{R,re}(t_p) - dt^j_{k,re}(t_{p+1}) \\ &= dt^j_{R,eph}(t_p) - dt^j_{k,eph}(t_{p+1}) + \\ &\quad + dt^j_{R,SAeph}(t_p) - dt^j_{k,SAeph}(t_{p+1}) + \\ &\quad + dt^j_{R,SAclock}(t_p) - dt^j_{k,SAclock}(t_{p+1}) + \\ &\quad + dt^j_{R,ion}(t_p) - dt^j_{k,ion}(t_{p+1}) + \\ &\quad + dt^j_{R,trop}(t_p) - dt^j_{k,trop}(t_{p+1}) + \\ &\quad + dt^j_{R,mp}(t_p) - dt^j_{k,mp}(t_{p+1}) + \\ &\quad + e^j_R(t_p) - e^j_k(t_{p+1})\end{aligned} \tag{3.9}$$

The characteristic we are interested in is the variance of $(dt^j_{R,k})_{re}$. For this we can neglect the variances of receiver noises, „normal" ephemerides and multipath effects. The main drivers for an increasing variance of $(dt^j_{R,k})_{re}$ are the variances of:

* reference and user receiver clock offset: the generation of the clock offset to system time has to be realised on the reference receiver and strongly depends on the modelling effort. For the pseudo-range correction an offset is also generated to bring the internal reference clock near to GPS time. But this value is only a raw estimate and thus can have a high variance. For the range correction this value has to be modelled as accurate as possible and is essential for the time user. Finally a GPS time restitution has to be realised, and this is not so easy to do under SA. A restitution strategy has to be formulated, and the model proposed must be adapted to the applied reference clock;
* satellite SA clock and ephemerides (if activated) effects mainly due to latency / age and increasing baseline);
* ionospheric, and tropospheric effects (mainly dependent on the baseline).

It also arises the question whether atmospheric delays are modelled or measured or considered on the user site.

3.6 Calibration and application aspects

Calibration and application aspects in this context mean hardware calibration in the conventional

sense and application of the dt_R^j worth at the receiver site.

The hardware calibration is a more or less understandable feature and will not be discussed here.

The main point for the calibration task is the application of the differential corrections, i.e. the reference clock offset in the time receiver. This is a large error source, because GPS receivers are usually designed for navigation and do often not fulfil the requirements of the timing user. Thus, it is very difficult to get hard- and software details from a dedicated manufacturer.

3.7 Theoretical conclusions

The transmission of additional information containing pseudo-range and range corrections includes the following properties:

* possible improved synchronisation for the time user with respect to the reference receiver clock (using pseudo-range corrections) or to the system time (using range corrections);
* if the reference receiver clock is synchronised to GPS-time or UTC(xx) the time user has an accurate direct access to this time scale;
* increasing the baseline between user and reference station receiver increases the variance of the resulting remaining delays and degrades the accuracy.

It should be noted here that the introduction of range corrections instead of pseudo-range corrections on the reference station will not influence the navigation user.

4. EXPERIMENTAL SETUP

In our measurement campaign pseudo-range corrections have been used broadcasted on 123.7 kHz form the Telekom in Mainflingen (near Frankfurt/M.) The reference receiver is a Trimble 4000SSI. In the past this station operated with a highly stable quartz oscillator, now it was replaced by a caesium atomic clock. At this point we want to mention that the improvement due to the atomic clock has not be considered in the clock offset monitoring presently, because the time user has not been considered as a possible customer of this service.

In DLR Oberpfaffenhofen an Efratom „DGPS FC" (S/N 001) six channel DGPS time receiver was installed. This receiver was fed with the differential data coming from a LF receiver (STAR·TRACK LWRX). The 1 pps output of that receiver was compared to

, the local clock - the hydrogen maser H1(DLR) of the Russian company Kvarz. The general setup can be obtained from fig. 4.1.

Now, over the measurement period of some weeks the difference between the 1 pps output and the hydrogen maser's 1 pps has been recorded each 5 seconds. These records were subjects for further analyses.

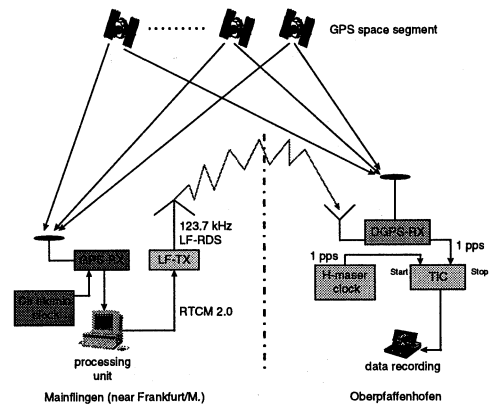


Fig. 4.1 Overview of the system's experimental setup.

5. POTENTIAL ACHIEVABLE ACCURACY

Let us estimate the resulting remaining error for a time user applying pseudo-range or range corrections of a reference station transmitted by a LF broadcasting station. The baseline is assumed to be about 300 km, the update rate is about 8 s (5 s latency, about 8 s age).

Resulting remaining error	uncertainty [ns]
ephemerides	3.0
SA clock under age concern	12.0
ionosphere	2.0
troposphere	1.0
multipath	2.0
receiver noise	2.0
calibration and application (mainly application of DGPS clock correction)	5.0
Total	≈ 14.0

Tab. 1 Potential achievable synchronisation accuracy for a time receiver using LF transmitted pseudo-range or range corrections

6. RESULTS

In this section we will present examples of two types of results: running the time receiver without and with DGPS data. In each case the local receiver position is fixed and hold constant during the measurement.

Fig. 5.1 shows the residuals of the 1 pps output from the time receiver with respect to the 1 pps of H1(DLR) over 1 day observation period (5.02.97). The measurement interval was 5 s. The standard deviation was calculated to be 35 ns (1σ).

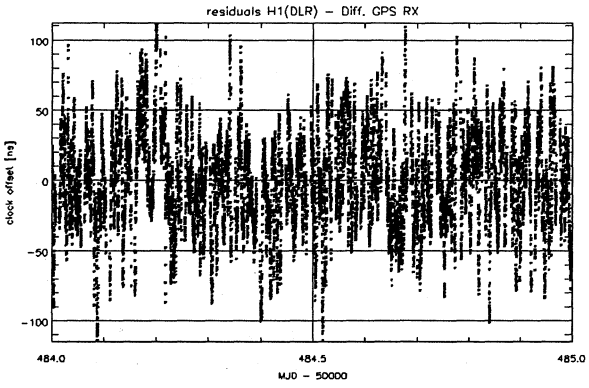


Fig. 5.1 Residuals of H1(DLR) - Efratom time receiver 1 pps comparison over one day (5.02.97) without differential data input.

With the same operation deals Fig. 5.2, but in this case the time receiver was fed with differential data. This measurement took place on the 28.01.97. The standard deviation was calculated to be 27 ns (1σ).

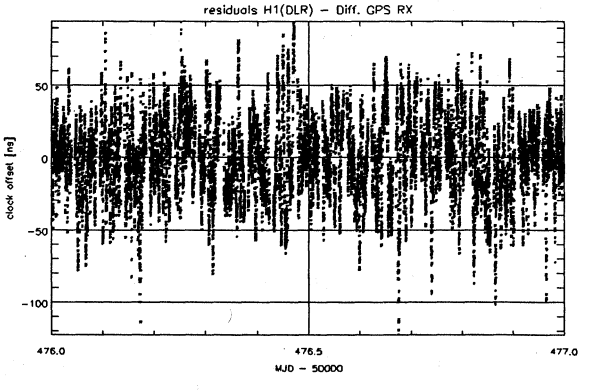


Fig. 5.2 Residuals of H1(DLR) - Efratom time receiver 1 pps comparison over one day (28.01.97) with differential data input.

7. CONCLUSION

We can obtain a significant difference between both preliminary estimated and achieved accuracy (note, in the experiment we only analysed the precision). The sources for such discrepancies have to be identified.

One significant error source is assumed to be located in the reference station: the clock offset modelling is not as accurate as possible for the time user and the new implemented atomic clock is not considered sufficiently. The clock model runs on the old quartz oscillator presently. The navigation user does not matter about this.

Another point is the unknown procedure in the time receiver to consider the pseudo-range and range corrections. DGPS receivers are usually designed for navigation purposes and do not meet all the requirements concerning the time user. Thus, we are not ready to describe the detailed algorithm for the DGPS processing inside the time receiver. On the other hand it is strongly recommended to know about this to mark important points.

We have also no accurate information about the latency and age intervals. This input is needed to estimate the error coming due to SA. We tried to find out realistic figures.

The DGPS time receiver is not capable to compensate for ionospheric effects. But, these effects are considered on the reference site. The differential effect can increase the error variance.

Further problems occured with the LF-receiver. The signal reception was interrupted repeatedly. This was indicated by a LED. What about the data handling of the DGPS-receiver in this case ?

The stand-alone solution for the DGPS-time receiver looks good. This is caused by the fact that „all in view processing" with 6 satellite channels has been made. Using only a typically one channel receiver the improvement would be significant. On the other hand multi-channel time receivers are inexpensive on the market now and have to be assumed as the standard equipment for the near future.

We conclude, that finally we should stress our investigations to sufficiently model the reference receiver clock offset. This leads to the restitution of GPS system time. The internal receiver processing has to be discussed assuming certain latency / age concerns.

REFERENCES

[DID-95] Dittrich J. and Kuehmstedt E., Real Time Differential GPS over long wave. Contribution of the Institut für Angewandte Geodäsie, Frankfurt am Main, to DGPS in GIBS, November 1995.

[GPS-97] GPS Information and Observation System (GIBS), Institut für Angewandte Geodäsie, February 1997

A New Method for Two-Way Common-View Synchronization of Ground-Based Atomic Clocks with PRARE

S. BEDRICH[1], J. HAHN[2]

[1]GeoForschungsZentrum Potsdam, Div. 1, Telegraphenberg A17, D–14473 Potsdam
(Stefan.Bedrich@dlr.de)
[2]DLR, Institut für Hochfrequenztechnik, P.O. Box 1116, D–82230 Oberpfaffenhofen
(Joerg.Hahn@dlr.de)

ABSTRACT

Most precise time transfer between remote atomic frequency standards is achievable with two-way signal transmission systems. This paper describes the results of a validation experiment that was realized by means of the PRARE system, which offers all necessary features to compare – in two-way Common View (CV) mode – time offset and time drift of up to four ground-based atomic clocks simultaneously. The obtained results confirm that the two-way PRARE CV technique is superior to standard GPS CV methods and should, therefore, be seriously considered for application in a future Global Navigation Satellite System (GNSS).

keywords: two-way time transfer, PRARE, Common View technique, GNSS

1 INTRODUCTION

When discussing new navigation satellite systems like the proposed European GNSS-2, not only adequate clocks for precise time keeping, but also appropriate methods for accurate time transfer have to be discussed [Ref. 1]. The time comparison method should allow to really control the stabilities of the involved clocks and should not have to deal with insufficient time signal transmission and data evaluation techniques. Two-way time transfer offers the possibility to strictly separate clock deviation and ground station position/satellite ephemeris errors. Their main advantage is that the basic one-way time interval measurements of space and ground clock can be reduced by quasi-simultaneous, nearly error-free two-way time interval measurements, if the signal-generating equipment is able to receive the time signal once again, after it has accomplished the round-trip travel, and compare it to the previously transmitted signal. By that approach, the clocks' deviation is truly *measured,* there is no need to *model* it.

The PRARE system (= Precise Range And Range-Rate Equipment), which has originally been developed for highly precise orbit determination and ground positioning purposes [Ref. 2], offers exactly these features. The PRARE measurement principle is based on two-way tracking of microwave signals in X-band (accuracy: sub-mm level) with an additional S-band downlink for ionospheric refraction determination. The signals are coherently modulated by a 10 respectively 1 Mcps/s pn-code for simultaneous two-way ranging (accuracy: sub-dm level). The complete PRARE system consists of a space segment, which is integrated onboard the target satellite (presently active onboard the European remote sensing satellite ERS-2), master control facilities, and about 30 transportable ground stations, which are distributed near-globally. A detailed description of the system and its main purposes can be found in [Refs. 3, 4].

2 EXPERIMENTAL SETUP

All PRARE tracking signals are referred strictly coherently to the central BVA-oscillator of the space segment. Moreover, the space segment is able to track up to four ground stations simultaneously. These important features have been taken advantage of in a dedicated time transfer experiment with PRARE that has been carried out between November 1995 and February 1996 [Ref. 5]. The standard PRARE tracking signals have been used to compare the time offset of three ground based frequency standards, which are located at the facilities of the

GeoForschungsZentrum (GFZ) Potsdam (Cs clock), at the Deutsche Forschungsanstalt für Luft- und Raumfahrt's (DLR) branch establishment in Weilheim (H-maser clock), and at the PRARE-Masterstation of GFZ in Oberpfaffenhofen (Rb clock, fig. 1).

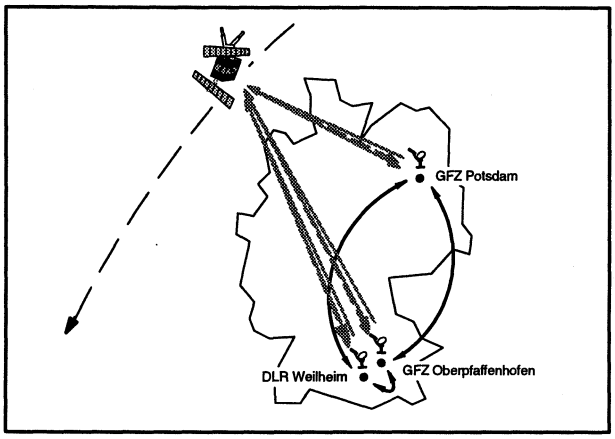


Fig. 1: Experimental two-way Common View comparison of a Cs (GFZ Potsdam), a Rb (GFZ Oberpfaffenhofen), and an H-maser frequency standard (DLR Weilheim) with PRARE on ERS-2.

At each site, a PRARE ground station and adequate time interval measurement equipment has been installed to measure during local passes of ERS-2 (five to six times a day lasting about 10 to 12 minutes each) the time offset and time drift between each of the ground clocks and the PRARE space segment oscillator. After reducing those individual one-way measurement data sets with the corresponding two-way measurement data sets, which are routinely preprocessed at the PRARE-Masterstation [Ref. 4], and correlating two of these scaled data sets respectively, all influences of the space segment and the space oscillator, which are common to both links, are cancelled, and the ground clocks' time scales can be inter-compared directly.

3 SYSTEM CALIBRATION

The fundamental data correlation technique including all corrections of the signal disturbances due to atmospheric refraction, internal hardware delays, phase centre variation, and the PRARE measurement principle, has been fully described in [Ref. 5]. In this paper, only the systematic experiment calibration efforts are focussed in detail.

For local time keeping, each of the clocks is driving a permanently operated GPS time receiver. The experiment layout was chosen such that the PRARE stations were linked to the clocks parallel to the GPS receivers at each site. This way, simultaneous clock comparison by means of the GPS Common View technique could be carried out, which was used for absolute calibration of the PRARE results. Internal calibration was fulfilled by careful measurement of the signal delays induced by the locally used equipment. Figs. 2 to 4 show the hardware setups and the relevant signal delays of each measurement site.

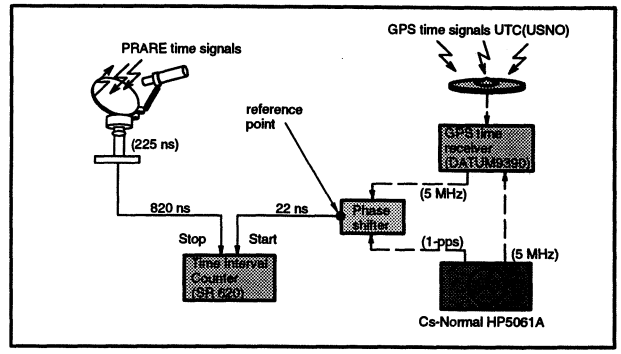


Fig. 2: Experiment setup and signal delays at GFZ Potsdam.

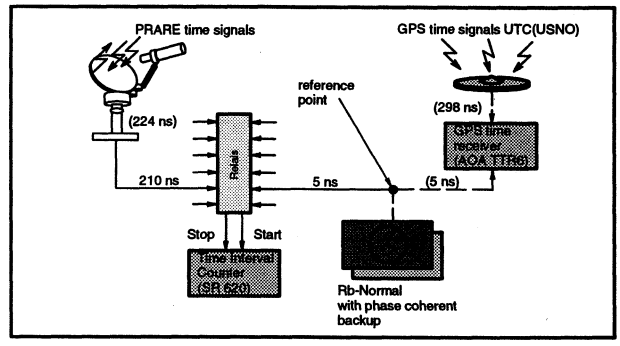


Fig. 3: Experiment setup and delays at GFZ Oberpfaffenhofen.

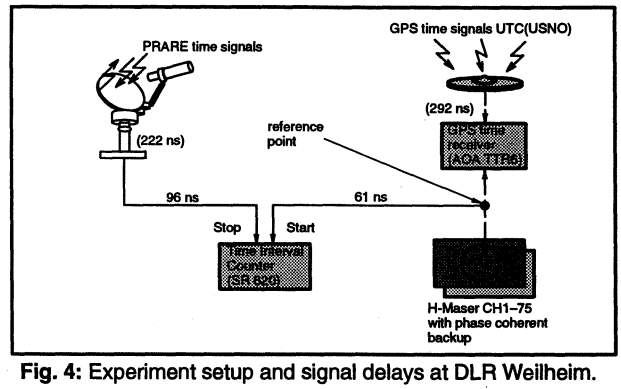


Fig. 4: Experiment setup and signal delays at DLR Weilheim.

4 PRARE DATA EVALUATION

The locally generated PRARE clock data sets, which consist of one value per second, have been smoothed after combining the data sets from two sites respectively in order to suppress short time noise.

This was done – in accordance to BIPM Common View data smoothing and to IGS data smooting procedures [Refs. 6, 7] – for 30 second and for 300 second intervals each. For the time transfer discussed in this paper, two representative PRARE measurement data sets of 14 December 1995, 19:30 h and 21:15 h, have been used. Fig. 5 shows the obtained clock correlation results for each of the couples H – Rb, Rb – Cs, and Cs – H including the respective standard deviation values, both for 30 s and 300 s smoothing. The local residual hardware delays have been taken into consideration.

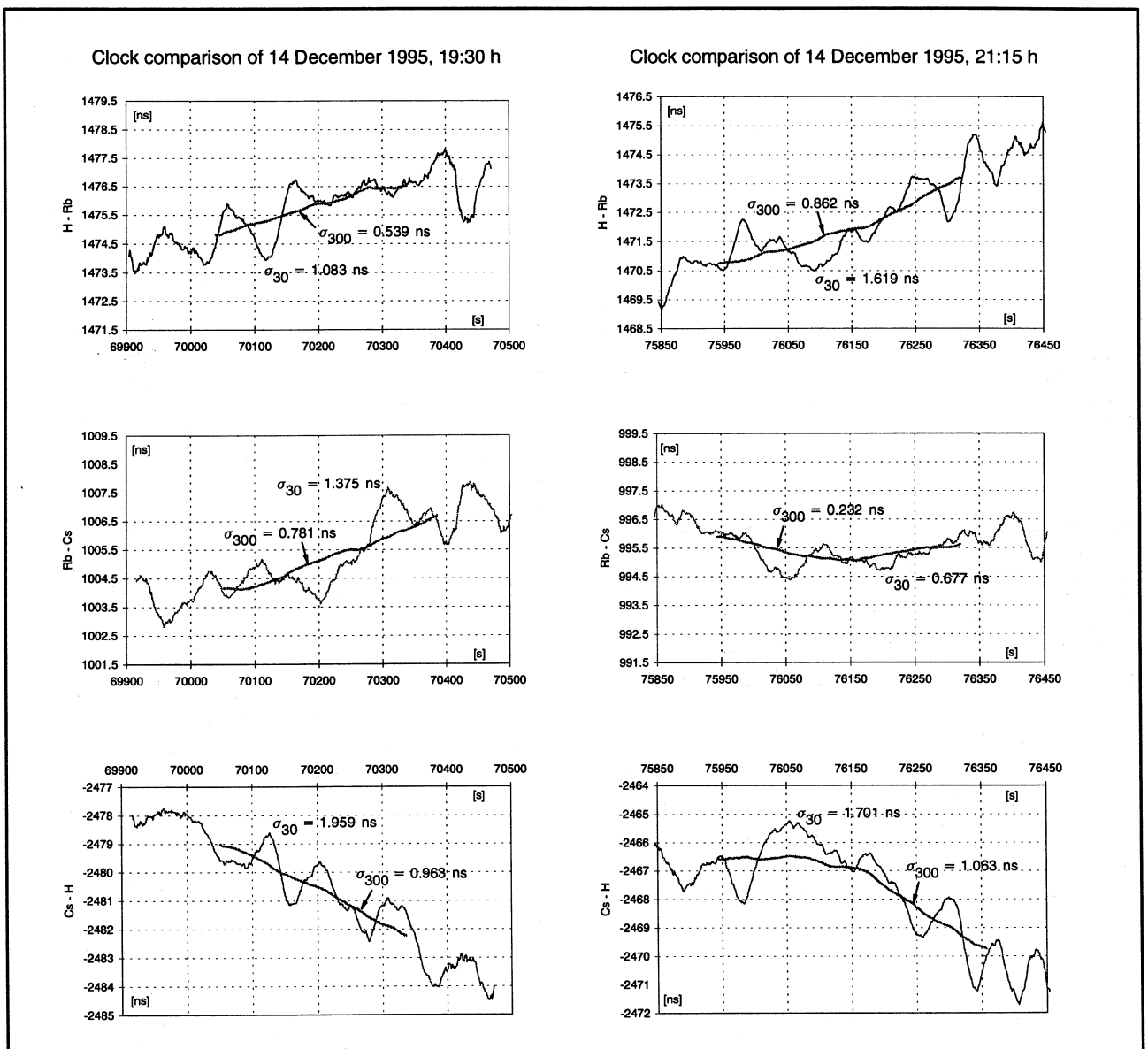


Fig. 5: Time offset and time drift between H-maser clock Weilheim and Rb clock Oberpfaffenhofen (top row), Rb clock Oberpfaffenhofen and Cs clock Potsdam (middle row), Cs clock Potsdam and H-maser clock Weilheim (bottom row) on 14 December 1995, 19:30 h (left column) and 21:15 h (right column) from PRARE two-way Common View observations. σ-values without drift removal.

It becomes obvious that the precision of the PRARE two-way Common View clock comparison technique is very high showing standard deviation values around 1 ns for 30 s smoothing. These values can be further reduced significantly by longer smoothing intervals (for example σ = 200 ... 800 ps for 300 s smoothing) and/or by linear adjustment of the clocks' relative drift. The latter method has not been made use of because of the short pass lengths, but both methods are standard techniques for clock data processing. By these approaches, routine precision values of around σ = 250 ps are easily achievable. Additionally, the results confirm that the scaled two-way data sets, which are used for reduction of

the one-way measurements, show no residual systematic distortions or deviations. This is due to the specific PRARE tracking data preprocessing routines which could be significantly improved during the last months following intensive internal calibration efforts [Ref. 4].

The baseline length between the clocks' locations seems not to play an important role, although the visibility zone of the ERS-2 satellite is not very large because of the low orbit height. Therefore, this question could not be investigated further. It should, however, be mentioned again that the quality of the obtained results is basically driven both by the high signal frequencies, which offer low signal noise, and the fact that all correction parameters are measured, not modelled within the PRARE system.

5 ACCURACY DETERMINATION

For system accuracy verification, three GPS time receivers (2 x AOA TTR6, 1 x DATUM 9390) were used during the experiment. By implementing the up-to-date BIPM Common View tracking schedule for Europe, simultaneous observation of identical GPS satellites was assured. Unfortunately, the tracking schedule could not be programmed at the GFZ Potsdam time receiver due to diverse reasons. Therefore, appropriate reference values could only be generated for the link DLR Weilheim – GFZ Oberpfaffenhofen.

A detailed description of the according data evaluation process as well as the correlation results for the whole comparison period can be found in [Ref. 5]. Fig. 6 shows an excerpt of the obtained correlation data (including the local residual hardware delays) which are relevant for the analysed PRARE passes.

The graph is subdivided into two sections, because the Rb clock is subject to automatic frequency control every three days to maintain coherency to UTC. The standard deviation of each of the two parts is – after linear adjustment – around ±20 ns, a usual result for routine Common View observations with GPS. The obtained accuracies and precisions of the two-way PRARE and the GPS Common View clock comparison results are summarized in table 1.

The table confirms the surpassing potential of the two-way PRARE CV time transfer technique compared to routine GPS CV measurements, when

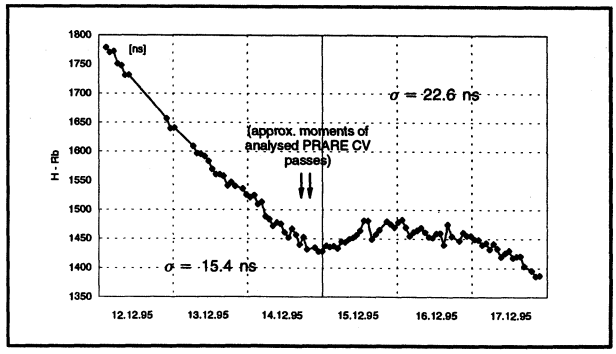


Fig. 6: Time offset and time drift between H-maser clock Weilheim and Rb clock Oberpfaffenhofen from GPS Common View observations between 12 and 17 December 1995. σ-values after drift removal.

precision is discussed. As far as accuracy is concerned, a nearly constant time bias of around +40 ns between the data sets was identified during the experiment. This could be due to the fact that each of the used AOA GPS receivers is specified with an absolute accuracy of around ±30 ns. Additional measurements with the help of a third GPS time receiver, which was operated in Weilheim and Oberpfaffenhofen for a week in each case to inter-calibrate the primary AOA receivers, resulted in similar biases [Ref. 5].

Table 1: Precision and accuracy of quasi-simultaneous two-way PRARE CV and GPS CV clock comparison measurements between H-maser Weilheim and Rb Oberpfaffenhofen (*: estimated).

	PRARE CV			GPS CV		difference	
	offset H – Rb	σ_{30}	σ_{300}	offset H – Rb	σ	PRA – GPS	σ
14.12.95, 19:30 h	1475.8 ns	1.083 ns	0.539 ns	1438.3 ns	15.4 ns	37.5 ns	45 ns*
14.12.95, 21:15 h	1472.0 ns	1.619 ns	0.862 ns	1431.7 ns	15.4 ns	40.3 ns	45 ns*

6 CONCLUSIONS

The time transfer experiment described in this paper proves the applicability and very high precision of differential ground clock comparison, which is achievable with the PRARE two-way Common View technique. The only existing disadvantages are due to the low earth orbit PRARE is currently operating in and could easily be overcome by installing the system onboard a high flying satellite.

When combining it with an ultra-stable frequency standard, the system could not only serve for comparing ground-based clocks, but also provide global time transfer with an unparalleled stability

and accuracy. As the standard PRARE system features, like near-real time ranging, simultaneous range-rate determination by doppler/phase measurement, and high capacity data transfer, could be fully maintained, its consideration in a new European satellite navigation system is strongly recommended [Ref. 1].

ACKNOWLEDGMENTS

The authors would like to thank Mr R. Manetsberger and Mr Ch. Lehner (DLR Weilheim) as well as Dr R. Neubert and Mr H. Fischer (GFZ Potsdam) for their intensive local experiment support, which was highly appreciated.

REFERENCES

[1] HODGE, C. C.; DAVIS, J. A.; GALLOP, J. C.; ALLAN, D. W.; ASHBY, N.; BEDRICH, S.; CUTLER, L. S.; HAHN, J.; NAU, H.; KERN, R. H.; MALEKI, L.; VESSOT, R. F. C. (1996): Towards 10 Millimeter Real-Time Position Determination and 30 Picosecond Time-Transfer Capability with the Next Generation of Global Navigation Satellite Systems (GNSS). *Proc. ION GPS-96*, Kansas City.

[2] REIGBER, CH.; HARTL, PH. (1990): Das Satellitenbeobachtungssystem PRARE. *ZfV* **12/1990**, 512–519.

[3] BEDRICH, S.; FLECHTNER, F.; REIGBER, CH.; TEUBEL, A. (1996): PRARE: A New Tool for Satellite Navigation and Geosciences. *Ann. Geophys.* **14** (Suppl. I: Abstr. XXI Gen. Ass. EGS, The Hague), C259.

[4] FLECHTNER, F.; BEDRICH, S.; MASSMANN, F. H.; REIGBER, CH. (1997): PRARE/ERS-2: System Status and Results. To be publ. in *CSTG Bull.* **13**.

[5] HAHN, J.; BEDRICH, S. (1996): "Common View" Clock Synchronization of Remote Atomic Clocks Using GPS and PRARE onboard ERS-2. *Proc. 10th EFTF*, Brighton, 393–398.

[6] LEWANDOWSKI, W.; THOMAS, C. (1991): GPS Time Transfer. *Proc. IEEE* **79**(7), 991–1000.

[7] GENDT, G. (1997): Analyse der IGS-Daten und Ergebnisse. To be publ. in *Proc. 41. DVW-Seminar*, GFZ, Potsdam.

RECENT WORK IN THE FIELD OF TWO-WAY SATELLITE TIME TRANSFER CARRIED OUT AT THE TUG

D. Kirchner*, H. Ressler** and R. Robnik**

*Technical University Graz, Austria
**Space Research Institute, Graz, Austria

ABSTRACT

After a brief report on two-way satellite time and frequency transfer (TWSTFT) software and hardware developments at the Technical University Graz (TUG) results of measurements of the differential delay of TWSTFT stations are reported and discussed. These results indicate that carrying out TWSTFT measurements together with measurements of the differential station delay frequency comparisons at the 10^{-15} accuracy level for averaging times less than one day should be possible. In addition envisaged activities involving both TWSTFT earth stations available at the TUG are mentioned.

Keywords: Two-way satellite time and frequency transfer (TWSTFT), satellite simulator, signal delay, stability, accuracy.

1. INTRODUCTION

The two-way satellite time and frequency transfer (TWSTFT) technique is well known for decades, but only recently it was used in a large scale experiment, the so-called INTELSAT field trial, carried out in 1994 and 1995 by six European laboratories and two laboratories in the United States of America; one of these laboratories was the Technical University Graz (TUG) (Refs. 1, 2). Based on experience gained during this experiment and previous TWSTFT activities, at the TUG extensive work was done to be ready for the planned follow-up experiments which should lead to operational use of the TWSTFT technique and should furthermore prove its frequency comparison capability at the 10^{-15} accuracy level for averaging times below one day. The activities concerned software and hardware developments and studies of the time transfer stability of TWSTFT systems.

2. SOFTWARE

For the exchange of TWSTFT data two formats are recommended by the Radiocommunication Study Group 7 of the International Telecommunication Union (Ref. 3). One is to be used for reporting individual 1 s measurements and the other one for reporting the results of a quadratic fit. In the first format in one data file only data from one satellite track can be reported, but giving the full data set. The purpose of the second format is to reduce the amount of data to be exchanged and to be able to report in one data file more than one track involving different partner stations and different satellite links. Moreover the exchange of related information concerning measurement uncertainty, delay calibration and delay stability, meteorological and other parameters necessary to process the data is possible. Software has been developed at the TUG to process data files according to the first format in an automated way producing numeric and graphic output and to convert data given in the first format together with the necessary additional data into the second format and to process data given in this format (see Fig. 1).

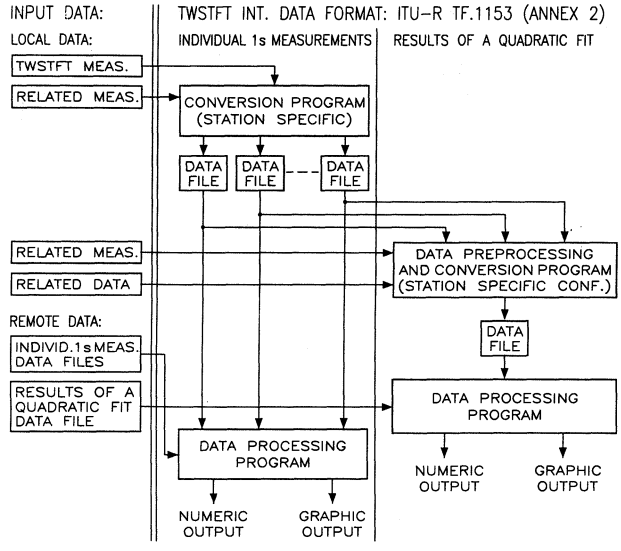


Figure 1 Overview of software developed at the TUG and the relationship to the ITU-R TWSTFT data formats.

3. HARDWARE

Two VSAT-type satellite earth stations (1.8 m antenna diameter) for TWSTFT operations are available at the TUG. Both stations are equipped with a satellite simulator (SATSIM) each in order to monitor signal delay variations and fully automated measurement systems are employed. Furthermore a third satellite simulator to be operated in some distance from the antennas is available. Details are given in (Refs. 4, 5). The stations and measurement systems can be arranged in different ways for example allowing to switch the measurement system of station 1 between both stations (see Fig. 2). Station 2, together with its measurement system, is designed also to be used as a mobile station allowing the calibration of other stations. For this purpose the complete system can be transported by a small trailer.

4. SIGNAL DELAY STABILITY

The crucial parameters for the performance (stability, accuracy) of a TWSTFT system are the stability of the difference between the transmit and receive delays and the knowledge of this differential delay or the difference of the differential delays of two systems assuming that the stability is not limited by other parts of the system such as time and frequency distribution, electronic counter and modem (Ref. 6). The frequency comparison accuracy of a system is limited by its time transfer stability - given by the stability of the differential delay - and the time comparison accuracy is given by the uncertainty of the knowledge of the differential delay, but cannot be better than the stability of the differential delay.

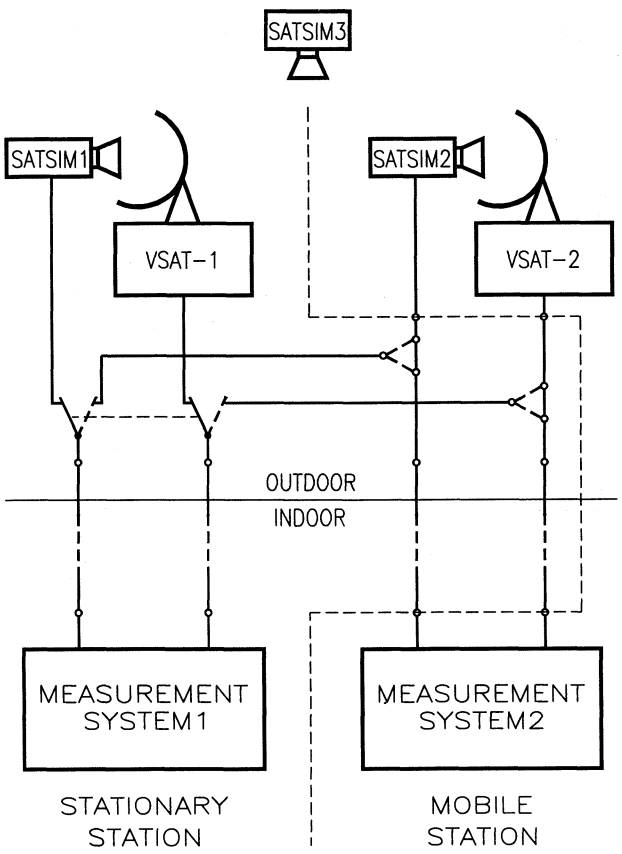


Figure 2 Scheme of the possible arrangements of both satellite earth stations and measurement systems.

4.1 Measurements

At the TUG measurements of the stability of the differential delay of station 1 were carried out during the INTELSAT field trial for a period of about 15 months together with the time transfer measurements which were performed three times per week (Ref. 6). A first evaluation of the stability of the differential delay of station 2 was carried out during February 1997 over a period of 27 days performing successive measurements each lasting about 12 minutes. A measurement of the differential delay consists of different delay measurements each lasting 100 seconds. Using this measurement scheme and performing the measurements under the same conditions as the time transfer measurements the measurement error

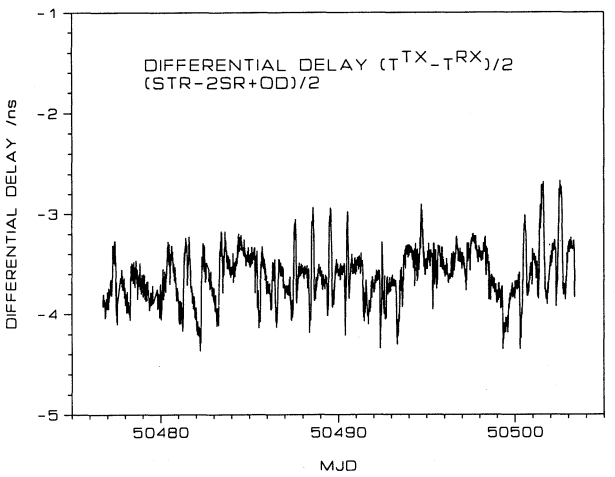


Figure 3 Differential station delay.

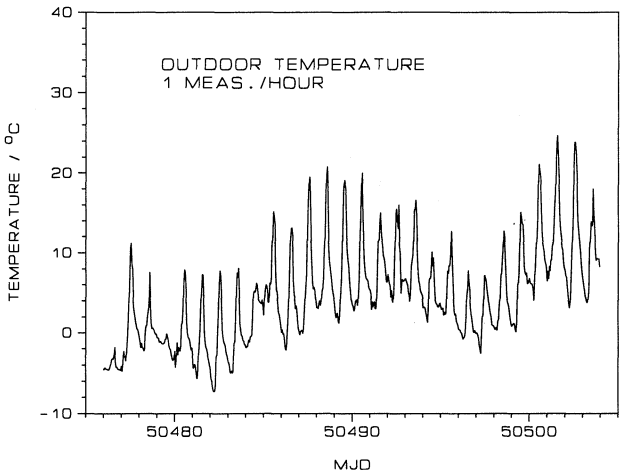


Figure 4 Outside temperature.

estimated by an error budget is smaller than 50 ps. The measured differential delay divided by 2 - as it enters the two-way equation (Ref. 7) - and the outside temperature are shown in Figs. 3 and 4, respectively. There is an obvious correlation between outside temperature and delay variation. During the reported period a peak-to-peak variation of the temperature of about 32°C and of about 1.7 ns of the differential delay can be observed. The stability calculated from the data given in Fig. 3 is depicted in Fig. 5, which is a composite time and frequency stability plot, thus allowing from one plot to estimate the time and frequency transfer capability (Ref. 5). Also given in Fig. 5 is the delay stability of station 1 as measured during the INTELSAT field trial.

4.2 Discussion

Performing one measurement per second TWSTFT usually shows white-noise PM behaviour (the actual noise level depending on the modem and the carrier-to-noise power density) up to averaging times of minutes and is then dominated by clock noise or using highly stable clocks by noise from the measurement system and the satellite earth station (Refs. 5, 6, 8). Fig. 5 shows that for the system used at the TUG the dominating noise source for averaging times longer than a few minutes is the satellite earth station. Measuring the differential station delay by means of a SATSIM and using this measurements to correct the TWSTFT measurements should allow to reduce this noise and therefore to improve the already excellent TWSTFT stability as evaluated by common-clock experiments (Ref. 6, 8). The possible improvement can be estimated from the stability calculated for a period of about 40 hours around MJD 50479 with only a smooth temperature variation and such allowing to estimate from the uncertainty of the SATSIM measurements a resulting frequency comparison capability of the 10^{-15} accuracy level for averaging times below one day (see Fig. 5). Such a figure has already been reported for a special TWSTFT set-up performing a common-clock experiment using VSATs of the same type - maybe causing cancellation of delay variations caused by temperature and humidity effects - and a satellite simulator on top of a mountain (Ref. 9). For comparison purposes the stability obtained for this experiment and also stability data reported for the GPS carrier-phase technique (Ref. 10) are given in Fig. 5.

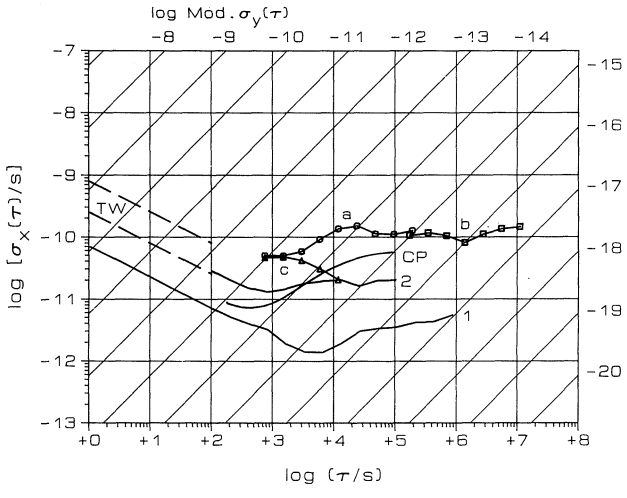


Figure 5 Stability of TWSTFT measurements for averaging times up to 100 s (TW) and of the differential delays of station 1 (b) and station 2 (a) and for a selected period for station 2 (c). Also given is the stability of a common-clock two-way time transfer between MITREX modems (ideal conditions)(1) and between stations of the same VSAT type employing a satellite simulator (2) and of the GPS carrier-phase technique (CP).

5. CONCLUSION AND ENVISAGED ACTIVITIES

Using the existing measurement set-up - the same modem is used for TWSTFT measurements and SATSIM measurements - both measurements cannot be performed simultaneously. The shortest repetition period for a combined TWSTFT and SATSIM measurement is about 6 minutes. This may slightly improve the stability for a certain averaging time compared with the 12 minute period used for the results presented which already indicate a frequency comparison capability of the 10^{-15} accuracy level for averaging times less than one day.

Using both earth stations equipped with SATSIMs available at the TUG, common-clock experiments will be carried out to further evaluate the frequency transfer capability of the TWSTFT technique. This will be done in addition to the regular INTELSAT measurements resumed in February 1997 and performed 3 times per week, for half an hour each within Europe and between Europe and the USA and hopefully for extended measurement periods using a repetition rate of about 6 minutes to verify the frequency transfer capability of the 10^{-15} level for averaging times less one day.

Furthermore the mobile station shall be used for repeated calibration trips to also evaluate the time transfer capability of the TWSTFT technique.

ACKNOWLEDGMENTS

The loan of a MITREX modem by the Institute of Navigation of the University of Stuttgart and of another one by the U.S. Naval Observatory is deeply appreciated. The work was supported by the Austrian Academy of Sciences and the Jubilee Fund of the Austrian National Bank.

REFERENCES

[1] Davis, J.A., P.R. Pearce, D. Kirchner, H. Ressler, P. Hetzel, A.Soering, G. de Jong, P. Grudler, F. Baumont and L. Veenstra: European Two-Way Satellite Time Transfer Experiments Using the INTELSAT (VA-F13) Satellite at 307°E, IEEE Trans. on Instr. and Meas., IM-44 (2), pp. 90-93, 1995.

[2] Davis, J.A., W. Lewandowski, J.A. DeYoung, D. Kirchner, P. Hetzel, T. Parker, W. Klepzcynski, G. De Jong, A. Söring, F. Baumont, K. A. Bartle, H. Ressler, R. Robnik and L. Veenstra: Comparison of Two-way Satellite Time and Frequency Transfer and GPS Common-view Time Transfer During the INTELSAT Field Trial, Proc. 10th European Frequency and Time Forum, pp. 382-387, 1996.

[3] ITU, International Telecommunication Union - Radiocommunication Study Group 7: Recommendation ITU-R TF.1153, The Operational Use of Two-way Satellite Time and Frequency Transfer Employing PN-Codes, Annex 2, Geneva, Oct. 1996.

[4] Reßler, H., D. Kirchner and R. Robnik: Satellite Earth Stations for Two-way Time Transfer at the Technical University Graz, 11th European Frequency and Time Forum, this Proceedings, 1997.

[5] Kirchner, D., H. Reßler and R. Robnik: An Automated Signal Delay Monitoring System for a Two-Way Satellite Time Transfer Terminal, Proc. 9th European Frequency and Time Forum, pp. 75-79, 1995.

[6] Kirchner, D., H. Reßler and R. Robnik: Signal Delay Stability of a Ku-band Two-way Satellite Time Transfer Terminal, Proc. 27th PTTI, pp. 303-311, 1995.

[7] Kirchner, D.: Two-Way Time Transfer via Communication Satellites (Invited Paper), Proc. IEEE (Special Issue on Time and Frequency), 79 (7), pp.983-990, 1991.

[8] Hackman, Ch., S.R. Jefferts and Th.E. Parker: Common-Clock Two-way Satellite Time Transfer Experiments, Proc. 1995 IEEE Frequency Control Symposium, pp. 275-281, 1995.

[9] ITU, International Telecommunication Union - Radiocommunication Study Groups: Doc. 7-2/Temp/5-E, Frequency Transfer at the 10^{-15} Accuracy Level Using Two-Way Satellite Time-and-Frequency-Transfer Systems, Geneva, Nov. 1994.

[10] Giffard, R.P.G., L.S. Cutler, J.A. Kusters, M. Miranian and D.W. Allan: Toward 10^{-15} and Sub-Nanosecond International Frequency and Time Metrology, 10th European Frequency and Time Forum, pp. 528-533, 1996.

RUBIDIUM STANDARDS AND FLYWHEELS

Chairman: A. De Marchi

STUDY OF THE FREQUENCY STABILITY OF LASER-PUMPED RB GAS-CELL FREQUENCY STANDARDS[1]

J. Q. Deng [+], G. Mileti [°], J. M. López-Romero [*], D. A. Jennings, F. L. Walls, R. E. Drullinger

National Institute of Standards and Technology, Time & Frequency Division, 325 Broadway, Boulder, CO 80303

[+] *Wuhan Institute of Physics, The Chinese Academy of Sciences, Wuhan, 430071, P. R. China*

[°] *Supported by a Swiss fellowship of the "Fonds National Suisse de la Recherche Scientifique"*

[*] *División de Tiempo y Frecuencia, Centro Nacional de Metrología, C.P. 76900, México*

ABSTRACT

We report the improvement of short term stability ($5 \cdot 10^{-13}\ \tau^{-1/2}$) that has been obtained with a laser-pumped passive gas-cell frequency standard.

Three types of lasers, including extended cavity (EC), solitary and DBR lasers, have been used with different stabilization schemes. Their amplitude modulation (AM) noise and phase modulation (PM) noise have been measured and compared.

The requirements on the *accuracy* and on the *stability* of their frequency stabilization have been established by light shift measurements which were performed with simultaneous heterodyne detection of the laser frequency. The use of the D_1 line instead of the D_2 line of Rb^{87} is also discussed. The possibility of simultaneously making the two light shift coefficients ($\Delta\nu_{clock}/\Delta I_{laser}$ and $\Delta\nu_{clock}/\Delta\nu_{laser}$) zero is described. Some limitations due to the microwave synthesizer are presented. Finally, clock stability measurements are presented and discussed.

Keywords : *Rubidium passive frequency standard, laser optical pumping, light shift, noise.*

1. INTRODUCTION

The paper contains four sections. The first section defines our goals and recalls the frequency stability limits of passive laser-pumped rubidium frequency standards. The second section displays our analysis of the two major potential sources of instabilities: the laser and the microwave interrogation. The third section presents the short term stability obtained with different types of laser diodes and discusses the long term behavior of the clock. The conclusions are given in the fourth section.

Our goal is to build a "super local oscillator" with a short-term stability near $1 \cdot 10^{-14}\ \tau^{-1/2}$ for use in clocks based on laser cooled atoms and ions.
Presently, we are using an atomic resonator containing a buffer-gas cell. This delivers a signal compatible with an ultimate short term stability of $1 \cdot 10^{-13}\ \tau^{-1/2}$. This stability can be called the "shot noise limit" since it is the theoretical limit obtained from the shot noise of the detected light. This noise is typically 1 pA/√Hz, when the DC light intensity has been optimized.

Different processes involving both the laser and the microwave radiation introduce frequency instabilities which are above the shot noise limit. The following section details the phenomena related to the laser and briefly describes some effects due to the microwave radiation. This last aspect will be presented with more details in a separate communication (Ref. 2).

2. CLOCK INSTABILITY SOURCES

2.1 Limitations from the laser

The laser radiation alters the clock stability by adding noise on the photodetector and by modifying its frequency via the light shift effect. Each of these processes results from different mechanisms which are discussed separately below.

2.1.1 Noise on the photocurrent

We have measured the photocurrent noise on the signal detector in typical operating conditions with three types of laser sources. The results obtained at a Fourier frequency of 300 Hz (typical clock modulation frequency) are summarized in Table 1.

These results show that the intrinsic intensity noise of all the laser sources is not significantly higher than shot noise (column 1). However, the values measured after the laser beam has passed through the vapor (column 2) indicate that additional noise is present on the detection photocell. This noise, which can be one order of magnitude higher than shot noise (with the EC laser and the solitary laser), is due to laser phase noise combined with the atomic absorption (Ref. 3).
Table 1 *Photocurrent noise on the photodetector with different types of laser diodes and locking schemes (shot noise = 1 pA/√Hz). The 1st and 2nd columns display the noise before and after the rubidium cell*

[1] Work of the US Government. Not subject to copyright.

respectively. The last column is obtained with the noise cancellation technique described in the text.

Laser and *locking type*	Noise before Rb [pA/√Hz]	Noise after Rb [pA/√Hz]	Noise with cancellation [pA/√Hz]
Solitary *current mod. 70* kHz	2	30	3
EC *piezo mod. 7* kHz	2	30	3
DBR *sideband 12* MHz	1.5	2	1.7

According to theory (Ref. 3), this additional noise should be proportional to the laser linewidth. Our results are in apparent disagreement with this prediction since the noise obtained with the EC laser is as high as the noise obtained with the solitary laser. We attribute this discrepancy to low frequency phase fluctuations of the laser, probably due to mechanical instabilities in the extended cavity system.

The total intensity noise perturbs the clock stability because it adds noise to the signal. For this reason, it is possible to improve the clock stability using a *passive* noise cancellation (Ref. 4). As described in (Ref. 1), this could be successfully implemented in a laser-pumped Rb clock.

As can be seen in the last column of Table 1, the shot noise can be approached with the three laser systems.

2.1.2 Light shift

Many studies on light shift in laser-pumped rubidium clocks have been reported (Ref. 5). In particular, the light shift is a limiting factor on the clock performance (Ref. 6). In this research, we have focused on the problem of *accuracy* and *stability* of the laser frequency, and its relationship with the clock short term and long term behavior.

With Rb^{87}, the laser frequency can be tuned either to the D_1 transition $S_{1/2} \rightarrow P_{1/2}$ (795 nm) or to the D_2 transition $S_{1/2} \rightarrow P_{3/2}$ (780 nm). D_2 has the advantage of more efficient optical pumping (and thus a higher double resonance signal). D_1, as will be shown, could be more interesting as far as the medium and long term performance is concerned.

Since there are two ground state hyperfine sublevels (F=1 and F=2), two groups of transitions are available for both D_1 and D_2. F=2 is usually chosen to be depopulated since the double resonance signal is higher by a factor of 5/3 for an almost unchanged light shift. Thus, the problem of selecting the optimal transition for clock operation is reduced to the choice between two possibilities : D_1 or D_2.

Figure 1 displays the light shift measurement corresponding to the first option. The laser frequency (current) has been swept through the atomic transition at different light intensities. The closest line appears to be approximately 140 MHz higher than the unique zero light shift frequency ($\nu_{LS=0}$). In order to fine tune the laser frequency, we shifted the laser frequency with an acousto-optic modulator (AOM) in a double pass setup. The light shift was then measured at different detuning frequencies, which were precisely known by heterodyne detection of the beat node against a second laser. The result obtained are shown on Figure 2.

Figure 2 is in good agreement with Figure 1 and provides a very precise measurement of the light shift coefficient. From this coefficient, we have deduced the requirements on the *accuracy* and the *stability* of the laser frequency tuning for a specified clock stability. These requirements are given in Table 2.

Table 2 *Summary of the requirements on the laser accuracy and stability deduced from the light shift coefficient of the D_1 transition of Rb^{87} ($\delta\nu_{clock}$ = relative clock frequency fluctuation, $\Delta\nu_{laser}$ = laser frequency detuning from the zero light shift frequency, $\delta\nu_{laser}$ = laser frequency fluctuation, δI_{laser} = laser intensity change).*

Light shift coefficient $\frac{\delta\nu_{clock}}{\Delta\nu_{laser} \cdot I_{laser}}$	$\frac{1 \cdot 10^{-11}}{\text{MHz} \cdot \mu\text{A}}$
Short term stability *laser frequency stability* ($\delta\nu_{laser}$ so that $\delta\nu_{clock} < 10^{-13}$?)	$\delta\nu_{laser}$ **< 5 kHz**
Long term stability *laser frequency stability* ($\delta\nu_{laser}$ so that $\delta\nu_{clock} < 10^{-12}$?)	$\delta\nu_{laser}$ **< 50 kHz**
laser frequency (tuning) accuracy ($\Delta\nu_{laser} \rightarrow \delta I_{laser}=1\% \rightarrow \delta\nu_{clock} < 10^{-12}$?)	$\Delta\nu_{laser}$ **< 2.5 MHz**
Closest saturated line	**20 MHz / 140 MHz** D2 / D1

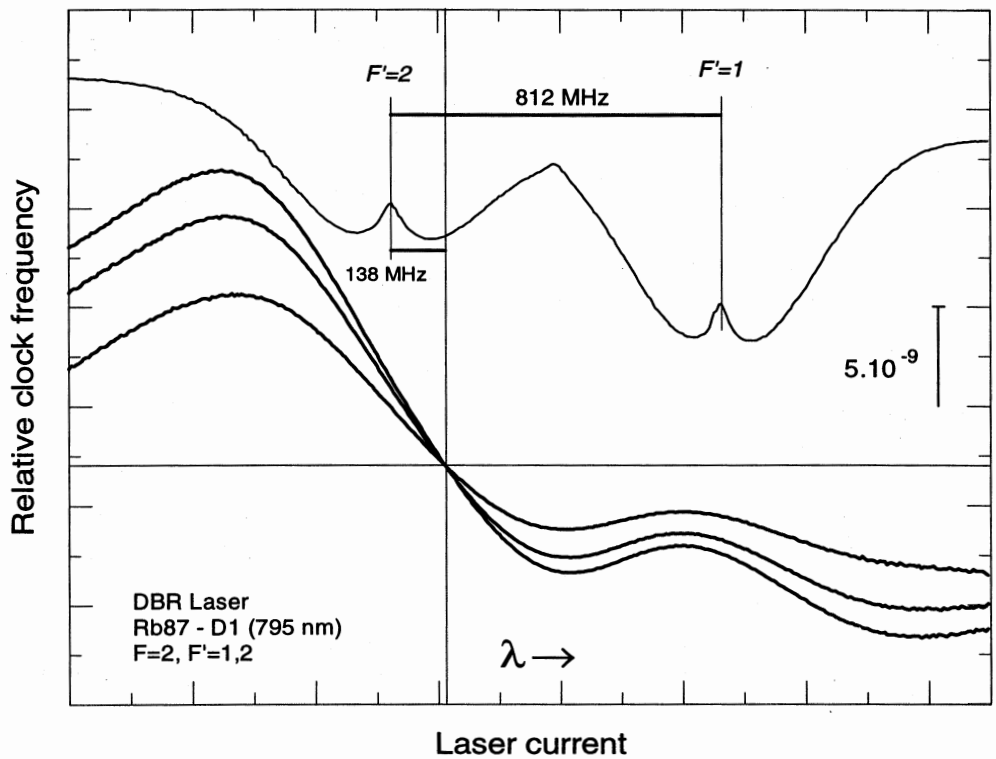


Figure 1 **Light shift coefficient of the D_1 transition of Rb^{87}.** *The lower curves show the clock frequency with three different light intensities, 4.3, 5.7, and 6.4 μA approximately (1 μA ≈ 4 μW/cm²). The upper curve displays the saturated absorption of a separate evacuated Rb^{87} cell.*

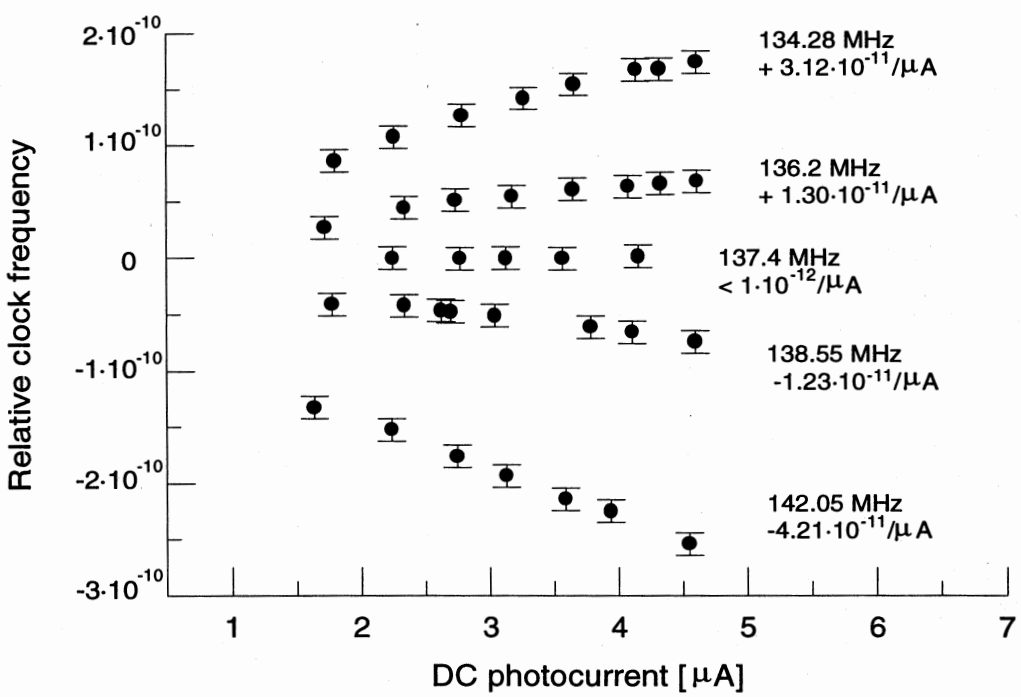


Figure 2 **Fine tuning of the laser frequency to the zero light-shift point.** *The clock transition frequency shift is plotted as a function of the light intensity (1 μA ≈ 4 μW/cm²) for different laser detunings. The beat frequency and the residual light shift is indicated for each curve.*

The requirement on the laser frequency *stability* for a *short term stability* of 10^{-13} (Table 2) is not too severe, and any locking system with a bandwidth higher than 10 kHz fulfills it. Reaching a stability of 10^{-14} will be much more difficult. The requirement on the laser frequency *stability* for good clock *long term stability* is more severe and is not achieved in our present setup.

The requirement on the *accuracy* of the laser frequency tuning to the zero light shift frequency indicates that a fine tuning of the laser frequency is necessary (with an AOM, by some offset locking system, or by fine tuning of the buffer gas, for example) if we want to allow 1% of change in the light intensity. Otherwise, active stabilization of the light intensity is necessary.

Another method for overcoming this requirement on accuracy of laser tuning is provided by the interesting shape of the D_1 transition. In fact, since the detuning between the exited state hyperfine sublevels is large (as compared to the Doppler broadening) two partially resolved Doppler absorption lines and light shift dispersion curves are visible (Fig. 1). In our setup, there is only one zero light shift frequency, but theoretical calculations have shown that at lower buffer gas pressures three zero light shift frequencies are present (Ref. 7). At the transition buffer gas pressure passing from three to one zero, there are two zeros. One of these zeros is a second order zero, which means that at this particular frequency the clock would be insensitive to intensity *and* frequency fluctuations.

The D_2 transition does not have the partially resolved lines and the light shift has a simple dispersion shape. Since there are more saturated transitions, lines closer to the zero light shift frequency can be found. In our particular setup, the closest line (of a separated evacuated cell) was only 20 MHz away from $\nu_{LS=0}$. It is possible that by adjusting the buffer gas pressure this residual detuning might be further reduced.

2.2 Limitations from the microwave interrogation

2.2.1 Noise on the photocurrent

In presence of optical pumping, the microwave field at resonance induces a reduction of the photocurrent. The total absorption of light due to the microwave constitutes the double resonance signal used for stabilizing the quartz oscillator. Unfortunately, it can also introduce additional noise on the photodetector. In our setup, the noise increases to 6 pA/√Hz when the microwave is present. This value is obtained when the microwave frequency is set exactly at resonance "clock" frequency (without modulation). It increases strongly on the side of the resonance, indicating that it is more likely due to PM noise than AM noise. Additional efforts will be needed to reduce this noise.

The theoretical limit of the clock stability is $6 \cdot 10^{-13}\ \tau^{-1/2}$. This estimation is obtained from the experimental resonance signal (discriminator slope ≈ 1 nA/Hz) and the noise level.

2.2.2 Microwave PM noise

Synthesizer PM noise at all the even harmonics of f_m (microwave modulation frequency) limits the stability of passive frequency standards (Ref. 8). We have measured the PM noise of our microwave synthesizer at 100 MHz by using the "three-cornered-hat cross-correlation technique" (Ref. 9) and have obtained the following values :

$$\begin{aligned} S_\Phi(f) &= 3.2 \cdot 10^{-14}\ \text{rad}^2/\text{Hz} \quad @\ 100\ \text{Hz} \\ S_\Phi(f) &= 2 \cdot 10^{-14}\ \text{rad}^2/\text{Hz} \quad @\ 300\ \text{Hz} \\ S_\Phi(f) &= 1.2 \cdot 10^{-14}\ \text{rad}^2/\text{Hz} \quad @\ 600\ \text{Hz} \\ S_\Phi(f) &= 1 \cdot 10^{-14}\ \text{rad}^2/\text{Hz} \quad > 600\ \text{Hz} \end{aligned} \tag{1}$$

According to the quasi-static model described in Ref. 8, with a modulation frequency of 300 Hz, the short term stability limit is $2.4 \cdot 10^{-13}\ \tau^{-1/2}$. This estimation is based on the 100 MHz PM noise measurement. The PM noise at 6.8347 GHz might be higher than expected from the multiplication chain.

3. FREQUENCY STABILITY MEASUREMENTS

We have measured the frequency stability of laser-pumped rubidium frequency standard with the different laser systems and the results are shown in Figure 3.

In each case, the stability is higher than the best reported stability of lamp pumped rubidium frequency standards ($5 \cdot 10^{-13} \cdot \tau^{-1/2} \leq \sigma_y(\tau) \leq 1.5 \cdot 10^{-12} \cdot \tau^{-1/2}$).

We also note that the D_2 transition gives better results than the D_1 transition. In fact, the three corresponding curves are below $1 \cdot 10^{-12}\ \tau^{-1/2}$ and represent the best reported performance of any rubidium passive frequency standards.

We have obtained better results with the EC laser than with the solitary and the DBR lasers. However, all the curves corresponding to the D_2 line are practically within the error-bar range. This remark is in agreement with the fact that the three laser systems all yield results close to the shot noise (2.1.1). Two important conclusions follow:

(1) The laser linewidth does not a limit clock stability for $\sigma_y(\tau) \geq 5 \cdot 10^{-13}\ \tau^{-1/2}$. In fact, there are more than two orders of magnitude difference between the spectral linewidth of extended cavity lasers and solitary lasers.

(2) We have reached the limit of short term stability allowed by the total photocurrent noise (6 pA/√Hz) and there is a potential factor of 5 improvement.

4. CONCLUSIONS

We have built a laser-pumped rubidium clock and demonstrated a short-term stability of $5 \cdot 10^{-13}\ \tau^{-1/2}$. The theoretical shot noise limit for this clock is $1 \cdot 10^{-13}\ \tau^{-1/2}$. One important limitation is the residual AM noise on the detection photocell, we are presently investigating the possibility of reducing it.

Our measurements on the laser frequency *accuracy* and *stability* clearly show that the residual medium and long term clock frequency fluctuations are due to laser frequency changes.

ACKNOWLEDGMENTS

We thank Eva Ferre-Pikal, Maria-Carlotta Delgado, Franklin Ascarrunz, Craig Nelson and Jose-Francisco Garcia for assistance in constructing the microwave synthesizer, Steve Jefferts for helpful discussions, and Pierre Thomann for his comments.

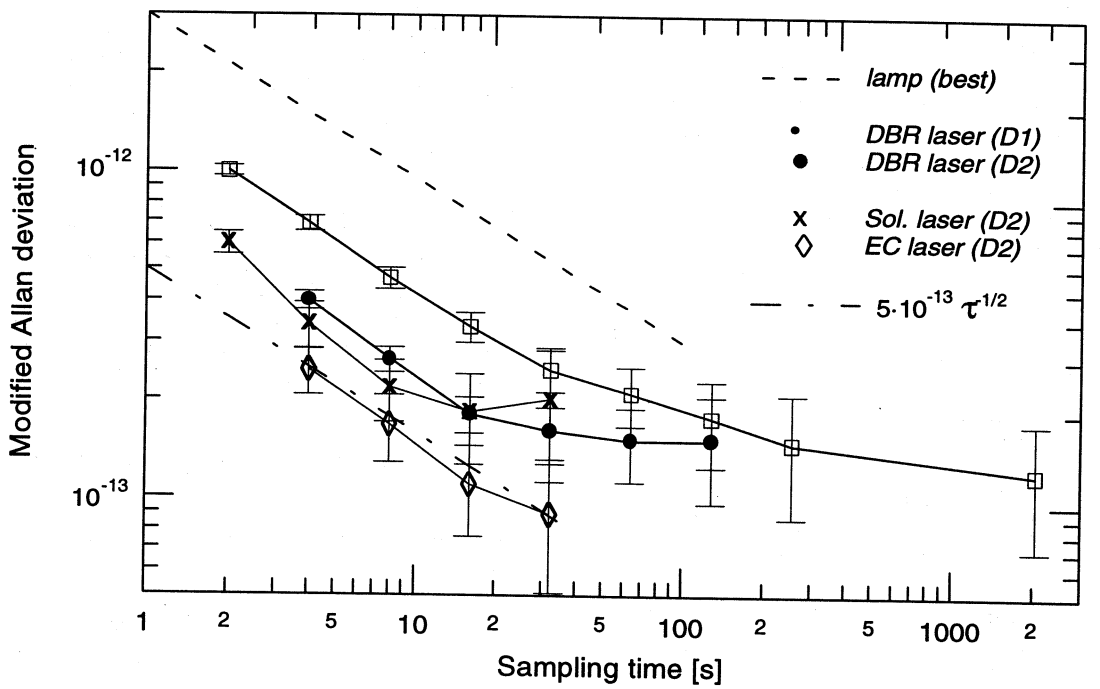


Figure 3 **Short-term stability**. *The experimental points correspond to the solitary (Sol.), extended cavity (EC) and DBR lasers. The corresponding atomic transitions (D_1 or D_2) of Rb^{87} are indicated. The best reported result with a lamp is shown for comparison.*

REFERENCES

[1] G. Mileti, J.Q. Deng, F.L. Walls, J.P. Lowe, R.E. Drullinger, *Proc. 1996 IEEE FCS*, 1066-1072, 1996.

[2] J. Q. Deng, G. Mileti, D. A. Jennings, F. L. Walls, R. E. Drullinger, *Proc. 1997 IEEE FCS*, 1997.

[3] D.H. Mc Intyre, C.E. Fairchild, J. Cooper, R. Walser, *Opt. Lett.*, **18** (21), 1816-1818, 1993.
H.G. Robinson, V.V. Vasiliev, V. L. Velichansky, L. Hollberg, A.S. Zibrov, *Proc. 14th Int. Conf. on Atomic Physics*, 1A7-1A8, 1994.
V. V. Iruvanti, R. Boyd, *Opt. Commun.*, **118**, 46-50, 1995.
M. Bahoura, P. Laurent, G. Santarelli, A. Clairon, to be edited in *Proc. 11th EFTF*, 1997.

[4] P.C.D. Hobbs, *Opt. & Phot. News*, 17-23, 1991.

[5] J.C. Camparo, R.P. Frueholz, C.H. Volk, *Phys. Rev. A*, **27** (4), 1914-1924, 1983.
M. Hashimoto and M Ohtsu, *IEEE Trans. on Instr. and Meas.*, **39** (3), 93-97, 1990.
G. Mileti and P. Thomann, *Proc. 8th EFTF*, 377-384, 1994.

[6] Y. Saburi, Y. Koga, S. Kinugawa, T. Imamura, H. Suga, Y. Ohuchi, *Electr. Lett.*, **30** (8), 633-635, 1994.

[7] B.S. Mathur, H. Tang, W. Happer, *Phys. Rev.*, **171** (1), 11-19, 1968.

[8] C. Audoin, V. Candelier, N. Dimarcq, *IEEE Trans. Instr. & Meas.*, **40** (2), 121-125, 1991.
R. Barillet, V. Giordano, J. Viennet, C. Audoin, *Proc. 6th EFTF*, 453-456, 1992.

[9] F.L. Walls, *IEEE Proc. of Int. FCS*, 289-297, 1993.

On the Use of a Modulated Laser for Hyperfine Frequency Excitation in Passive Atomic Frequency Standards

Filippo Levi*, Aldo Godone*, Carlo Novero* and Jacques Vanier**

***Istituto Elettrotecnico Nazionale, "Galileo Ferraris", Strada delle Cacce, 91 , Torino, Italia.**
****Guest worker, from: Dép. de Physique, Université de Montréal, Montréal, Canada.**

1. ABSTRACT

The use of Coherent Population Trapping, (CPT), for the realisation of a passive frequency standard is examined. The results of a calculation based on a perturbation approach are outlined, making explicit the signal size, its linewidth and the expected light shift. Preliminary experimental results obtained on cells containing cesium and various buffer gases are given. Conclusions are drawn on the possible advantages and expected difficulties to be encountered in the practical realisation of a frequency standard based on this approach.

Keywords: Atomic frequency standard, Dark line, Buffer gas, Light shift, Coherence.

2. INTRODUCTION

The advent of solid state diode lasers has created new avenues in the field of atomic frequency standards. Passive primary cesium standards have been realised in which state magnetic selection and detection are replaced by optical pumping. Many studies and experimental results have also been reported on the use of diode lasers to replace the standard spectral lamp in passive rubidium standards (1). Finally, diode lasers have been used to cool atoms to a very low temperature in schemes allowing the realisation of a clock using the Ramsey cavity approach with an extremely narrow central fringe (2). In these approaches, the high power density of the lasers is used to accomplish exchange of energy between the atoms and the electromagnetic field, altering either the populations of the various energy levels or the dynamic behaviour of the atoms such as their velocities.

Another avenue which has not been exploited as intensely in the field of microwave atomic frequency standards is the use of the laser coherence property itself. The standard level structure in optically pumped atomic frequency standards using alkali atoms consists essentially of two hyperfine ground levels between which microwave transitions are excited, (the clock transition), and an excited state that is used for population inversion through intensity optical pumping. It is possible, however, to excite the microwave transition by means of optical transitions alone making use of the coherence property of two lasers having a frequency difference equal to the hyperfine frequency, ν_{hf}. The resulting physical phenomenon called Coherent Population Trapping (CPT) gives rise to a dark line (3). It has been the object of several experimental studies (4) and elegant analysis of the mechanisms involved have been presented (5,6). In these analysis the creation of the narrow dark line with interesting properties for the field of atomic frequency standards is made explicit. The use of the phenomenon in rubidium for this purpose has been examined and preliminary experimental results have been reported (7). The avenue seems to be promising for the realisation of a passive frequency standard with improved frequency stability, virtual absence of light shift and reduced complexity in view of the absence of a microwave cavity.

In the present paper the CPT phenomenon is examined in connection with its use with the cesium atom which offers some advantages such as high hyperfine frequency, low glass reactivity, natural isotope purity. The results of a perturbation calculation making explicit dark line signal size, line width and light shift are presented. Some preliminary experimental results are also given for cells containing cesium and various buffer gases.

3. THEORETICAL RESULTS

The theory is developed for a three level system as shown in Figure 1.

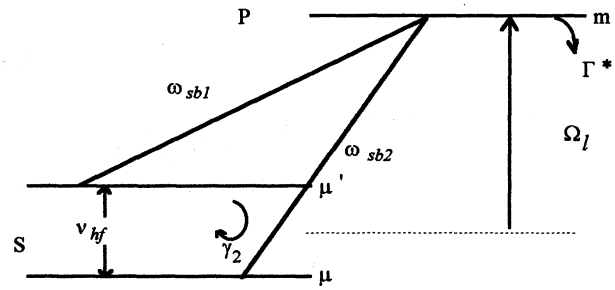


Figure 1. Three level system used in the analysis.

Two radiation fields excite coherently the two transitions $\mu \rightarrow$ m and $\mu' \rightarrow$ m. In practice the need to use two lasers can be avoided through frequency modulation of a single laser at a subharmonic of the desired hyperfine frequency. With the laser frequency Ω_l tuned to the middle of the two ground state hyperfine levels, the first two spectral sidebands ω_{sb1} and ω_{sb2} are made to coincide with the optical transitions.

A detailed analysis has been done of this so-called Λ system excited by two lasers (6). The exact solution obtained, however, does not lead to simple interpretation and analysis of the experimental data. In order to obtain a more transparent solution, we have made a calculation using a perturbation approach based on the assumption that optical

saturation of the spectral lines by the sidebands is low and that the effect of the dark line takes place on a very narrow range of frequencies over a broad optical line. The details of these calculations will be published elsewhere (8). The calculations take into account all coherences created in the ensemble and all relaxation mechanisms in the excited state and the ground state. For simplicity, in the present description of dark-line signal amplitude and linewidth, we have assumed that the optical excitation Rabi frequencies ω_{1R} and ω_{2R} are nearly equal. We set them equal to ω_R, the value of the Rabi frequency when the two sidebands are tuned to their respective transitions. This does not affect substantially the final result and, furthermore, this is close to the situation encountered in our experimental arrangement. The results are:

a) population of excited state, ρ_{mm}:

$$\rho_{mm} = \frac{\omega_R^2}{4}\left\{\frac{1}{\left(\frac{\Gamma^*}{2}\right)^2 + (\Delta\omega_{sb})^2}\right\}\left\{1 - \frac{\frac{\omega_R^2}{\Gamma^*}\left(\gamma_2 + \frac{\omega_R^2}{\Gamma^*}\right)}{\left(\gamma_2 + \frac{\omega_R^2}{\Gamma^*}\right)^2 + (\omega - \omega_{\mu\mu'})^2}\right\} \quad (1)$$

b) linewidth, $\Delta\nu_{1/2}$:

$$\Delta\nu_{1/2} = \frac{1}{\pi}\left\{\gamma_2 + \frac{\omega_R^2}{\Gamma^*}\right\} \quad (2)$$

c) light shift, $\Delta\nu_{ls}$:

$$\Delta\nu_{ls} = -\frac{1}{4}\frac{\Delta\omega_{sb}}{(\Gamma^*/2)^2 + (\Delta\omega_{sb})^2}\left(\omega_{1R}^2 - \omega_{2R}^2\right) \quad (3)$$

In these equations, γ_2 is the relaxation rate of the coherence in the ground state, Γ^* is the decay rate of the excited state taking into account the natural spontaneous emission and the decay caused by collisions with the atoms of the buffer gas, ω is the frequency difference between the two sidebands and $\omega_{\mu\mu'}$ is the hyperfine frequency. In the light shift expression, $\Delta\omega_{sb}$ is the detuning of the sidebands from the selected optical transition frequencies. It is assumed that in our experimental arrangement this detuning is equal for the two sidebands. The analysis is also based on the assumption that $\omega_{iR}/\Gamma^* < 0.1$. Under this condition ρ_{mm} does not take a value larger than 0.01 and it is expected that the perturbation approach is valid within a few percent.

In practice we have used a technique in which the fluorescence of the cesium vapour is monitored. The fluorescence signal is then given by:

$$P_{fluo} = \int_V \hbar\Omega_l\Gamma_f\rho_{mm}n dV \quad (4)$$

where Γ_f is the decay rate of the excited state resulting in fluorescence radiation and n is the Cs density. In the case of some buffer gases, quenching of the fluorescence through collisions may be present resulting in a reduction of the total fluorescence.

Using Equations 1 and 4 it is readily observed that the fluorescence is made of two combined spectral lines at frequencies $\omega_{m\mu}$ and $\omega_{m\mu'}$ over which is superimposed the narrow dark line. In the case of a cell without buffer gas, velocity selective depopulation of the two levels of the ground state takes place. This results in the absence of Doppler effect on the observed fluorescence. In the case where a buffer gas is present, the width of the fluorescence reflects the added homogeneous broadening caused by collisions.

We define the contrast as the ratio of the dark line amplitude over the fluorescence signal. It is observed that in the three level system considered 100% contrast is theoretically possible. In practice, however, one deals with a more complicated level scheme as shown in Figure 2 for cesium.

A detailed calculation using Clebsch-Gordan coefficients of all σ^+ allowed lambda transitions, combined with the effect of the buffer gas, shows that the available contrast is limited to about 3% for the 0-0 transition.

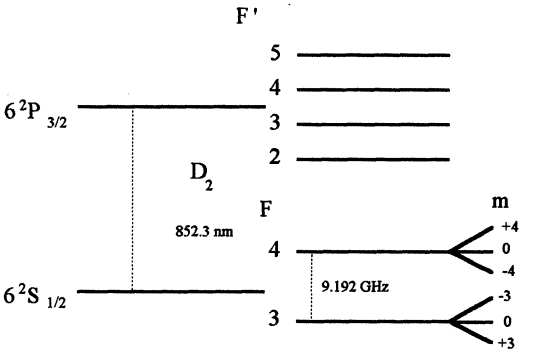


Figure 2. Cesium atom lower energy levels

On the other hand it is observed that the linewidth, as given by equation 2, is function of the optical power (Rabi frequency) with a slope equal to $1/\Gamma^*$.

Finally equation 3 shows that the light shift is proportional to the detuning of the two sidebands and the difference of their power through the respective Rabi frequencies. The light shift is absent when the two Rabi frequencies have the same amplitudes.

4. THE BUFFER GAS

Due to the thermal velocities of the cesium atoms, the dark line has a width that is inversely proportional to the transit time across the laser beam. For a wide beam, Doppler broadening is also present. In order to avoid these effects, a buffer gas may be used as in the classical microwave standard approach (Dicke effect) (9). The linewidth is then limited by diffusion and buffer gas relaxation processes. In practice the system is excited by means of a laser beam of radius 'a' over a length l equal to that of the cell used. A solution of the diffusion equation considering the first diffusion mode and spin exchange relaxation gives a line width equal to:

$$\Delta\nu_{1/2} = \frac{1}{\pi}\left\{\left[\left(\frac{2,4}{a}\right)^2 + \left(\frac{\pi}{l}\right)^2\right] D_0 \frac{P_0}{P} + L_0 \bar{v}_r \sigma_{2bg} \frac{P}{P_0} + \gamma_{se}\right\} \quad (5)$$

where D_0 is the diffusion constant of cesium in the buffer gas at pressure P_0, P is the buffer gas pressure, L_0 is Loschmidt constant, v_r is the relative velocity of the cesium and buffer gas atoms, σ_{2bg} is the collision cross section and γ_{se} is the spin exchange relaxation rate. Table 1 gives result of a calculation using equation 5 for a typical buffer gas with $D_0 = 0.1 cm^2/s$, and P = 10 Torr.

a	γ_2	$\Delta\nu_{1/2}$
0.1 cm	4400 s^{-1}	1400 Hz
1 cm	58 s^{-1}	18 Hz

Table 1 Typical dark-line diffusion limited linewidth expected in cells containing a buffer gas.

5. EXPERIMENTAL RESULTS

The experimental set-up used in the present preliminary studies is shown in Figure 3

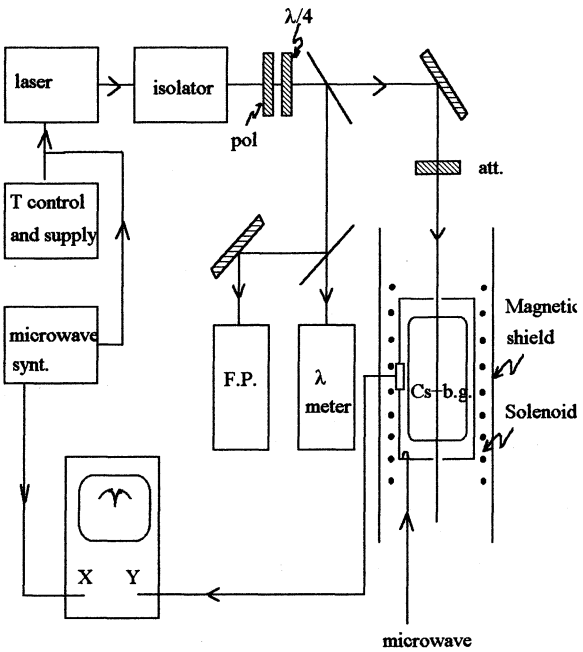


Figure 3 Experimental set-up

This system allows the observation of the hyperfine resonance both through the dark-line approach and the classical approach. In this last case microwave energy at the proper frequency is fed directly in the enclosure containing the cell. Although this enclosure is not a resonant structure at the microwave frequency, sufficient energy can be fed-in, to observe the ground state σ and π hyperfine transitions. This provides a very convenient means of verification of limiting linewidth, power broadening and quick measurement of frequency shifts due to the presence of the buffer gas. For the observation of the dark line the current of the laser diode is modulated at half the hyperfine frequency by means of a frequency synthesiser. In order to avoid $|\Delta m| = 2$ transitions, the laser beam is circularly polarised.

a) Laser frequency modulation. The experiments reported here rely on the possibility of modulation of the laser at microwave frequencies. This can be done directly by modulating the laser driving current at a subharmonic of the desired hyperfine frequency. However diode lasers have a very low input impedance which depends on their construction. The technique works relatively well with open heatsink lasers. A typical result with such a diode is shown in Figure 4, as observed with a Fabry-Perot interferometer having a free spectral range of 1.5 GHz. The modulation frequency in that case was 4.596 GHz, one half the ground state hyperfine frequency of cesium. In our experiments we have used lasers with and without external cavity. We have found that the experimental results on the dark line were not affected by the laser linewidth.

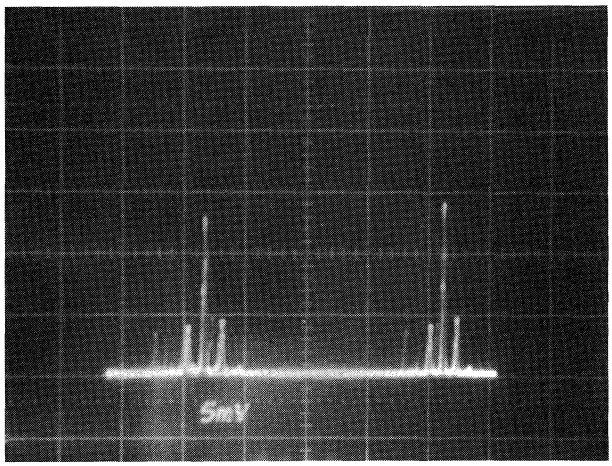


Figure 4 Typical spectrum of a laser diode modulated at a frequency of 4.596 GHz. Due to overlapping of Fabry-Perot modes and free spectral range limitations the sidebands appear close to the carrier.

b) Dark line signals: The dark line signals were observed both with cells without a buffer gas and with a buffer gas.

-No buffer gas. In that case the motion of the atoms is free. Fluorescence is controlled by the natural life time of the excited state and the signal originates from a given group of atoms having a velocity which makes them resonant with the applied optical radiation. The velocity range in interaction is controlled by the width of the excited state, and the fluorescence does not exhibit Doppler broadening, as expected. A typical result is shown in Figure 5. The relative amplitude of the various Δm=0 transitions follows closely the weight dictated by the Clebsch-Gordan coefficients for the particular Λ scheme used in the case of circular polarisation. The linewidths are of the order of 100 kHz and are controlled by the finite transit time of the atoms across the laser beam. The contrast observed for the 0-0 transition is given in Table 2.

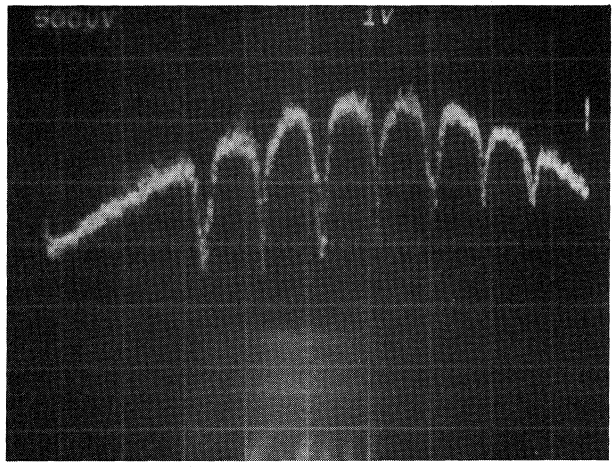


Figure 5. Ground state hyperfine dark lines observed in a cesium cell without buffer gas at room temperature. Horizontal scale: 1MHz/div. First sidebands power: 1 mW/cm^2

-Cells containing a buffer gas. Several cells containing various buffer gases at several pressures were studied. These cells had been fabricated some twenty years earlier and were still functioning as intended. Buffer gases included N_2, Ne, CH_4, C_3H_8, C_4H_{10} and Ar. A typical dark-line result is shown in Figure 6 in the case of a cell containing neon at a pressure of 45.8 Torr. The contrast for the 0-0 transition is given in Table 2.

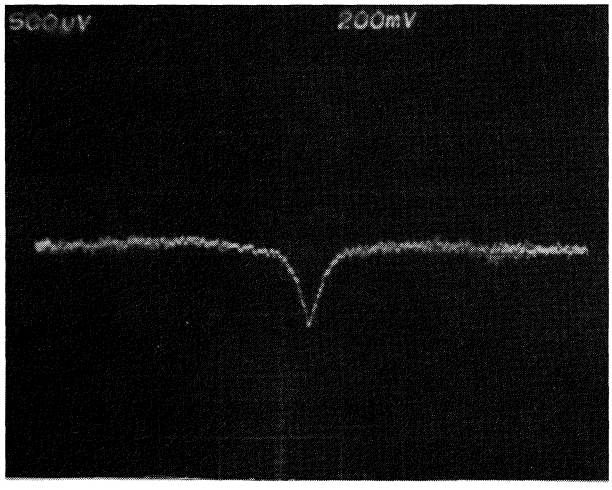


Figure 6 0-0 hyperfine transition Dark line observed in a Cs cell containing Neon at 45.8 Torr at room temperature. Horizontal scale: 1kHz/div. First sidebands power: 0.1 mW/cm^2

It is observed that the lineshape is not well described by equations 1 and 4. The upper part of the line is sharper than a Lorentzian. It has been observed experimentally that the line is closer to a Lorentzian shape when the laser beam is expanded. It appears that the effect is connected to an inhomogeneous saturation of the transitions by the laser beam which varies in intensity over its cross section. A simple calculation over a Gaussian beam shape appears to confirm this effect. Similar results have also been reported in other circumstances and appear to agree with this conclusion (10).

Contrast	Calculation	Experiment
	0-0 transition	
No buffer gas	6.2 %	5%
Buffer gas	3.1%	2 to 3%

Table 2 Calculated and measured contrast of the dark-lines

As expected, the linewidth is a strong function of the laser intensity and of the size of the cross section of the laser beam. This is shown in Figure 7 for the case of a cell containing N_2 at a pressure of 7.4 Torr

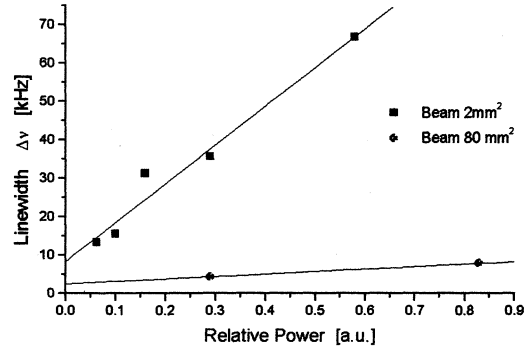


Figure 7 Width of the dark line in a cell containing N_2 at 7.4 Torr as a function of laser power for two settings of the laser beam width

A systematic study of a selected number of cells was made. This consisted in measuring the width of the two spectral lines of the D_2 transition. The limiting line width of the dark line as a function of laser intensity was also measured for a laser beam expanded to about 1 square centimetre. In order to obtain a comparison with the theory presented above the linewidth was also measured as a function of laser power. In theory the slope should be proportional to $1/\Gamma^*$, where Γ^* is the lifetime of the excited state as affected by the buffer gas collisions. These results are summarised in Table 3.

The physical construction of the cells available did not allow an accurate determination of the power density inside the cells. Consequently it was necessary to normalise the data to one particular cell assuming that this density would remain constant from one cell to the other. However it is observed that even with this limitation, the normalised results are in fairly good agreement with the theoretical calculations. Future experiments with cells with optically flat windows should provide better data for this verification.

buffer gas	P (Torr)	$\Delta\nu_{opt}$ (MHz)	$\Gamma^*/2\pi$ (MHz)	γ_2/π (Hz) [3]	$1/\Gamma^*$ (1^{-9} s)	$1/\Gamma^*$ norm [4]	Exp data[5]
none	-	528	5 [1]	-	32	-	-
Ne	45.8	962	690 [2]	56	0.23	0.92	1.1
CH_4	9.3	840	490 [2]	268	0.32	1.28	1.04
CH_4	17.5	980	720 [2]	244	0.22	0.88	0.95
C_3H_8	19.7	962	690 [2]	407	0.23	0.92	0.92
C_4H_{10}	11.5	927	650 [2]	397	0.25	1	1

Table 3 Experimental data on the various cells studied: [1] From natural life time. [2] From a Voigt profile calculation. [3] Limit of observed linewidth (zero power). [4] Normalised to cell containing C_4H_{10} as a buffer gas. [5] As obtained from a measurement of the slopes of linewidth against laser power, in the case of an expanded beam. The optical linewidths reported above as $\Delta\nu_{opt}$ refer to measurements made on the F=3 - F'= 2,3,4 (D_2 line) transitions.

6. APPLICATION TO THE FREQUENCY STANDARDS FIELD

The previous analysis and data has direct application to the frequency standard field. The use of the dark line as a means of detecting the hyperfine transition and locking an external oscillator to it has many basic advantages. First, experimental data confirming theoretical predictions show that signals with relatively high contrast (2 to 3 %) are observed. This should make possible the realisation of a high signal to noise ratio with a resulting good short term frequency stability. Secondly, as highlighted in the analysis, the light shift should be reduced considerably as compared to the classical approach because of the particular combined pulling effect of the sidebands. In fact the simple second order theory developed shows that for equal sidebands amplitudes, large detuning of the laser frequency will not cause a light shift. Experimentally, as a rough check and with a resolution of the order of 10 Hz, we have not been able to detect any dark line shift for laser detuning up to 300 MHz. This is at least one order of magnitude smaller then in the case of the standard optical pumping intensity approach (11).

At this time several studies remain to be done. In particular it appears important to investigate further the behaviour of the dark line signal with temperature. We have found that in our detection technique the dark line signal changes amplitude and shape considerably above 35 ^{0}C. Although the fluorescence level increases with temperature, the amplitude of the dark line decreases. It is believed that this effect is due to multiple scattering which causes a decrease of the coherence in the ensemble. It should be mentioned that although we have used fluorescence as a means of detection of the dark-line, other means, such as a probe beam, could also be used with some advantages. The studies suggested above should also be done in this case.

Finally we wish to conclude by stating that the proposed approach appears to offer several advantages regarding items such as: construction simplicity (no microwave cavity), frequency multiplication only to a subharmonic of the hyperfine transition frequency, strongly reduced light shift compared to the classical intensity pumping approach, and possibly reduced chemical activity of cesium as compared to rubidium.

7 ACKNOWLEDGEMENTS

We would like to thank P. Tavella for her help in many of the calculations reported, F. Bertinetto, M. Inguscio and A. De Marchi for their cooperation and CSELT (Centro Studi E Laboratori Telecomunicazioni) for the instrumentation lent us. We also wish to thank S. Leschiutta for his encouragement. This work was partially supported by ASI (Agenzia Spaziale Italiana).

8 REFERENCES

1) See for examples articles in the proceedings of this Forum and those of previous years. For a general description of the field of atomic frequency standards see also J. Vanier and C. Audoin. "The Quantum Physics of Atomic Frequency Standards, Adam Hilger, 1989.
2) A. Clairon, P Laurent, G. Santarelli, S. Ghezali, S. N. Lea and M. Bahoura, IEEE Trans. Instr. & Meas. Vol. 44,1995, p 128. See also: G. Santarelli et al. 1996 Proc. of the twenty-eighth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting. To Appear.
3) G. Alzetta, A. Gozzini, L. Moi and G. Orriols, Il Nuovo Cimento, 36 B, 5, 1976.
4) See for example Laser Spectroscopy, XII International Conference , Capri, Italy, 1995; World Scientific; editors M. Inguscio, M. Allegrini, A. Sasso.
5) R. G. Brewer and E. L. Hahn, Phys. Rev. 11, p. 1641, 1975.
6) G. Orriols, Il Nuovo Cimento, 53, p. 1, 1979.
7) N. Cyr, M. Têtu and M. Breton, IEEE Trans. Instr.& Meas. Vol. 42, 1993, p. 640.
8) J. Vanier, et al. to be published
9) R.H. Dicke, Phys. Rev. , vol 89, 153, p 472
10) P. Minguzzi, S. Profeti, M. Tonelli and A. Di Lieto J. Opt. Soc. Am. B Vol. 3, 1986, p. 1075.
11) F. Levi, C. Novero, A. Godone, G. Brida, IEEE Trans. Instr. & Meas. Vol. 46, n°2, 1997.

DYNAMICAL BEHAVIOUR OF AN OSCILLATOR SLAVED TO A PERIODICALLY INTERROGATED ATOMIC RESONATOR

C. Audoin, Laboratoire de l'Horloge Atomique
Unité Propre de Recherche CNRS associée à l'Université Paris-Sud
Bâtiment 221 - Université Paris-Sud, 91405 Orsay Cedex, France

G. Santarelli, A. Makdissi and A. Clairon,
BNM/Laboratoire Primaire du Temps et des Fréquences,
Observatoire de Paris, 61 avenue de l'Observatoire, 75014 Paris, France

ABSTRACT

We establish the equation which describe the time domain behaviour of an oscillator controlled by a periodically interrogated atomic resonator, such as those based on an atomic fountain or an ion trap. We give the transient response of the slaved oscillator. We derive its response to systematic frequency changes of the free running oscillator. We point out that a second integration in the feedback loop may not improve significantly the rejection of slow perturbations unless a condition relative to the timing of the atom-field interaction is verified. We consider the effect of the free-running oscillator noise and the detection noise on the long term frequency stability of the standard. Experimental results agree with the theoretical analysis.

Keywords : Frequency Control, Atomic Frequency Standards

1. INTRODUCTION

Advanced high performance atomic resonators, based on an atomic fountain or an ion trap have been developed recently [1,2,3]. In these devices, the atoms are not interrogated permanently like in more traditional atomic resonators, but periodically with dead time between each atom-field interaction. The atomic response being available at the end of each interrogation process only, the control voltage of the oscillator slaved to the atomic resonance is updated periodically, at discrete times.
We summarize here a simple but efficient analysis of the most useful properties of the frequency control loop and of the controlled oscillator. It will be published in more details elsewhere [4]. In this paper, we do not consider the limitation of the frequency stability of the slaved oscillator by down-conversion of the oscillator intrinsic frequency noise [5]. A theoretical description and an experimental evaluation of this effect have been presented at the Forum last year [6].

2. TIME DOMAIN DESCRIPTION OF THE FREQUENCY VARIATION OF THE SLAVED OSCILLATOR

2.1. Sampling of the microwave field by the atoms

Figure 1 shows the succession of events during one of the cycles of operation, of duration T_c. The cycle starts at t_k, immediately after a frequency correction has been applied to the controlled oscillator. The atom-field interaction begins at t'_k such that $t'_k - t_k = T_p$ and it lasts the time interval T_i. In the LPTF cesium fountain, the control voltage of the oscillator is updated at the end of each cycle, e.g. at t_{k+1}.
We denote $\Delta\omega$ any frequency offset of the microwave field with respect to the atom transition frequency, but irrespective of the frequency modulation which is necessary to probe the atomic resonance.

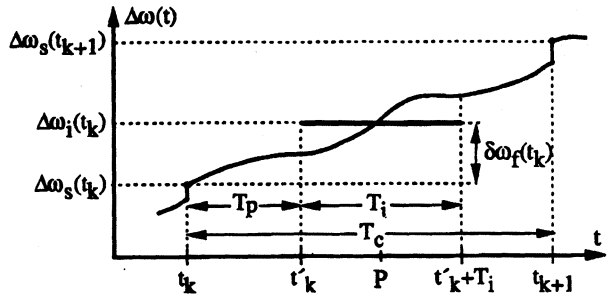


Fig-1. Frequency change of the microwave field during a cycle and its mean value during the interrogation

Between two frequency corrections the frequency of the oscillator evolves freely. We thus have, for $t_k \le t < t_{k+1}$:

$$\Delta\omega(t) = \Delta\omega_s(t_k) + \Delta\omega_f(t) - \Delta\omega_f(t_k), \qquad (1)$$

where the subscripts s and f stand for slaved and free, respectively. Although the frequency varies during the atom-field interaction, we can define an effective frequency offset, $\Delta\omega_i(t_k)$ which is constant and provides the same atomic resonator response as $\Delta\omega(t)$. We have :

$$\Delta\omega_i(t_k) = \Delta\omega_s(t_k) + \delta\omega_f(t_k), \qquad (2)$$

with :

$$\delta\omega_f(t_k) = \frac{1}{g_0 T_c}\int_{t_k}^{t_{k+1}} g(t-t_k)\left[\Delta\omega_f(t) - \Delta\omega_f(t_k)\right]dt. \qquad (3)$$

In this equation, g(t) is the frequency sensitivity function [5,6] and g_0 is its mean value during the cycle considered.

2.2. Error signal

The microwave field is square-wave frequency modulated with period $2T_c$ and modulation depth ω_m. The difference between two successive detection signals at times t_{k-1} and t_k is computed. It gives, at time t_k, an information on the microwave field mistuning during the two half-periods of modulation, from t_{k-2} to t_{k-1} and from t_{k-1} to t_k. Assuming a constant frequency offset, the sign of this difference changes each time k is incremented to k+1. To obtain the error signal $\Delta E(t_k)$ this difference is multiplied by $(-1)^k$. Assuming that $\Delta\omega_s$ and $\delta\omega_f$ are small compared to the atomic line-width, it can be shown that we have :

$$\Delta E(t_k) = \frac{\partial h}{\partial \omega} N_0 \left[\Delta\omega_s(t_{k-2}) + \delta\omega_f(t_{k-2}) + \Delta\omega_s(t_{k-1}) + \delta\omega_f(t_{k-1})\right] + \delta\tilde{N}(t_{k-1}) + \delta\tilde{N}(t_k). \qquad (4)$$

In this equation, N_0 and $h(\omega-\omega_0)$ — an even function of $(\omega-\omega_0)$ — represent the peak to valley height of the resonance pattern and its shape, respectively. The derivative is taken at the frequency which was actually applied between t_{k-1} and t_k. We have set :

$$\delta\tilde{N}(t_k)=(-1)^k\,\delta N(t_k), \qquad (5)$$

where $\delta N(t_k)$ represent the fluctuation of the atomic response, i.e. the detection noise, at time t_k. The time series $\{\delta N(t_k)\}$ is a sequence of independent random variables. We assume that it represents a white noise process, having mean zero and variance $\sigma^2_{\delta N}$. It is easy to show that the $\delta\tilde{N}(t_k)$ have the same autocovariance coefficients as the $\delta N(t_k)$. Thus the $\delta\tilde{N}(t_k)$ have the same variance and power spectral density as the $\delta N(t_k)$.

2.3. Equation of the frequency control loop

Case A. The error signal is added at time t_k to its previously accumulated value $\Delta'E(t_{k-1})$. We thus have :

$$\Delta E'(t_k)=\Delta E'(t_{k-1})+\Delta E(t_k). \qquad (6)$$

This equation is that of a numerical integrator. Its output updates the oscillator frequency at t_k. We thus have, k being a constant :

$$\Delta\omega_s(t_k)=\Delta\omega_f(t_k)+K\Delta E'(t_k). \qquad (7)$$

The equation which describes the dynamical behaviour of the frequency control loop is obtained from Eqs. (4,6,7) :

$$\begin{aligned}&\Delta\omega_s(t_k)-(1-\beta)\Delta\omega_s(t_{k-1})+\beta\,\Delta\omega_s(t_{k-2})\\&=\Delta\omega_f(t_k)-\Delta\omega_f(t_{k-1})-\beta\left[\delta\omega_f(t_{k-1})+\delta\omega_f(t_{k-2})\right]\\&+K\left[\delta\tilde{N}(t_k)+\delta\tilde{N}(t_{k-1})\right].\end{aligned} \qquad (8)$$

The quantity β is the open loop gain, defined by :

$$\beta=-KN_0\,\partial h/\partial\omega. \qquad (9)$$

Case B. One may try to improve the rejection of the error signal low frequency components by adding a recursive filter whose input is $\Delta'E(t_k)$ and whose output is given by :

$$\Delta E''(t_k)=\Delta E''(t_{k-1})+\left(\frac{T_c}{\tau_2}+\frac{\tau_1}{\tau_2}\right)\Delta E'(t_k)-\frac{\tau_1}{\tau_2}\Delta E'(t_{k-1}). \qquad (10)$$

The parameters τ_1 and τ_2 are time constants. For a low enough Fourier frequency, f, its frequency response is $(1+2\pi jf\tau_1)/2\pi jf\tau_2$.
In that case, the equation of the frequency loop becomes :

$$\begin{aligned}&\Delta\omega_s(t_k)-(2-\beta_1-\beta_2)\Delta\omega_s(t_{k-1})\\&+(1+\beta_2)\Delta\omega_s(t_{k-2})-\beta_1\,\Delta\omega_s(t_{k-3})\\&=\Delta\omega_f(t_k)-2\Delta\omega_f(t_{k-1})+\Delta\omega_f(t_{k-2})\\&-(\beta_1+\beta_2)\delta\omega_f(t_{k-1})-\beta_2\,\delta\omega_f(t_{k-2})+\beta_1\,\Delta\omega_s(t_{k-3})\\&+(K_1+K_2)\delta\tilde{N}(t_k)+K_2\delta\tilde{N}(t_{k-1})-K_1\delta\tilde{N}(t_{k-2}),\end{aligned} \qquad (11)$$

with $\beta_1=\beta\tau_1/\tau_2$, $\beta_2=\beta T_c/\tau_2$, $K_1=K\tau_1/\tau_2$ and $K_2=KT_c/\tau_2$.
For a given frequency sensitivity function g(t), Eqs. (8) or (11) contain all what is known about the properties of the slaved oscillator (proportional to that of the microwave field), when sampled at times t_k. The complete description of the behaviour of the output frequency of the standard requires Eq. (1) in addition to Eqs. (8) or (11). The first equation gives the instantaneous frequency change between the sampling times t_k and t_{k+1}.

2. STABILITY CONDITION

Case A. The stability condition is $0<\beta<1$. The transient behaviour is of the damped type for $0<\beta<0.172$ and of the oscillatory damped type for $0.172<\beta<1$. Figure 2 shows examples of the transient response to a step function applied to the oscillator frequency. A time constant and an oscillatory pseudo-frequency of the transient response can be defined [4]. Figure 3 depicts their variation versus the loop gain ß.

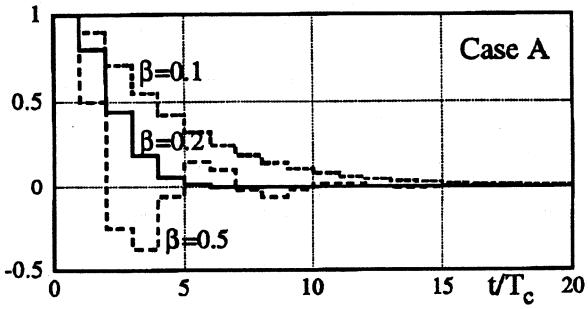


Fig-2. Response to a step function in case A.

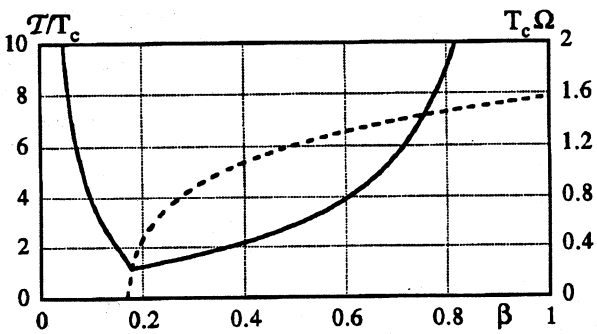


Fig-3. Variation of the time constant and of the pseudo-oscillatory angular frequency versus loop gain.

Case B. The stability condition is the following :

$$0<\beta_2<2\beta_1(1-\beta_1)/(1+\beta_1). \qquad (12)$$

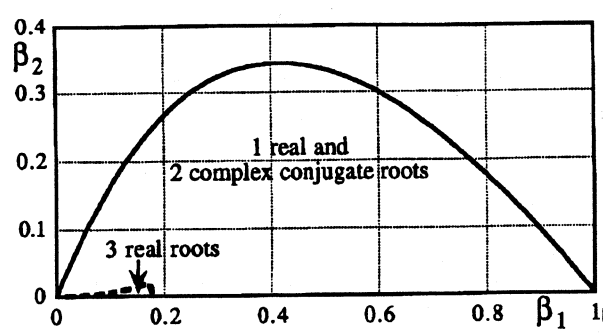


Fig- 4. Shape of the stability domain in case B.

Figure 4 shows the shape of the stability domain in that case. The transient response is oscillatory damped for most of the allowed values of β_1 and β_2.

4. RESPONSE TO SLOW SYSTEMATIC PERTURBATIONS

4.1. Response to a frequency ramp

A frequency ramp of the oscillator, due to ageing for instance, can be represented by :

$$\Delta\omega_f(t) = \omega_0 r t, \qquad (13)$$

where r is the slope of the ramp. Then, $\delta\omega_f(t_k)$ is a constant which does not depend on t_k and we set :

$$\delta\omega_f(t_k) = \omega_0 r T_1, \qquad (14)$$

where T_1 is given by :

$$T_1 = \frac{1}{g_0 T_c}\int_0^{T_c} \theta g(\theta) d\theta, \qquad (15)$$

with $\theta = t - t_k$. Assuming that the sensitivity function is represented by an even function when the time origin is displaced to the middle point between the beginning and the finishing of the atom interrogation process (e.g. at point P in figure 1), it can be shown that we have [4] :

$$T_1 = T_p + T_i/2. \qquad (16)$$

The steady state error is given by :

$$\Delta\omega_s(t_k)/\omega_0 = r\left(\frac{T_c}{2\beta} - T_1\right) \quad \text{in case A,} \qquad (17)$$

$$\Delta\omega_s(t_k)/\omega_0 = -r T_1 \quad \text{in case B.} \qquad (18)$$

4.2. Response to a slow sinusoïdal frequency change

Since the frequency of the oscillator varies linearly between the frequency corrections, applied at discrete times, the frequency change is rT_c over the duration of a cycle. Figure 5 shows the saw-tooth shaped instantaneous frequency of the controlled oscillator.
The oscillator may be perturbed by slow periodic changes of its environmental conditions. We set :

$$\Delta\omega_f(t)/\omega_0 = C_0 \sin(2\pi f t), \qquad (19)$$

and we assume that the period of this sinusoïdal variation is much larger than T_c, i.e. that we have $2\pi f T_c << 1$. Then, between times t_k and t_{k+1}, the frequency offset $\Delta\omega(t)$ given by Eq. 1 can be written as :

$$\Delta\omega(t) = \Delta\omega_s(t_k) + D(t_k)(t - t_k), \qquad (20)$$

where $D(t_k)$ is the derivative of $\Delta\omega_f(t)$ at time t_k.
According to the approximation made, $\Delta\omega(t)$ varies linearly between t_k and t_{k+1}, with a slope which changes very slowly. We thus have, from results of the preceding Section and to first order with respect to $2\pi f T_c$:

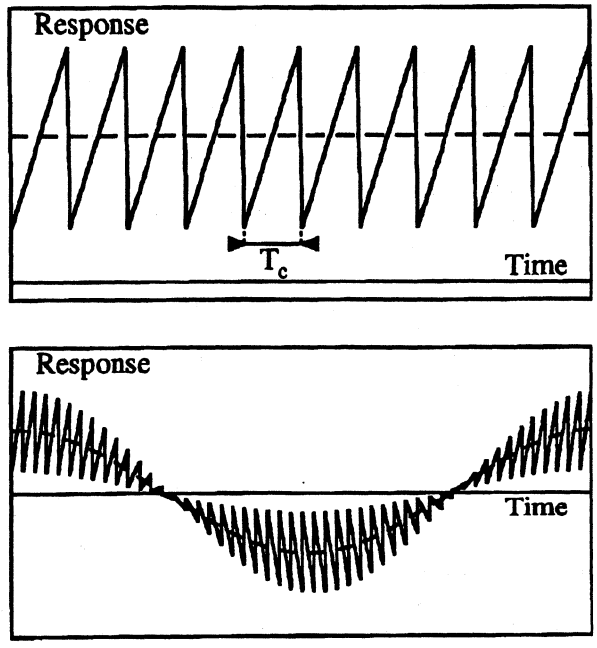


Fig-5. Response to systematic perturbations, Top : to a ramp, bottom : to a sine-wave. The dashed line represents the mean value of the output frequency.

$$\frac{\Delta\omega_s(t_k)}{\omega_0} = 2\pi f T_c\left(\frac{1}{2\beta} - \frac{T_1}{T_c}\right) C_0 \cos(2\pi f t_k) \quad \text{in case A,} \qquad (21)$$

$$\frac{\Delta\omega_s(t_k)}{\omega_0} = -2\pi f T_1 C_0 \cos(2\pi f t_k) \quad \text{in case B.} \qquad (22)$$

Besides the residual sampled sinusoïdal variation, $\Delta\omega_s(t_k)$, of the frequency of the controlled oscillator, a linear frequency change, of amplitude $2\pi f T_c \cos(2\pi f t_k)$, occurs during a cycle. Figure 5 shows the variation of the instantaneous frequency of the controlled oscillator.

4.3. Practical consequences

The characterization of the frequency stability of the atomic frequency standard involves an averaging of the instantaneous frequency of the controlled oscillator over a time interval τ. We will consider the long term stability, such that $\tau >> T_c$. In that case, we have to deal with the mean value of the instantaneous frequency and a term proportional to $T_c/2$ must be added to the right-hand side of Eqs. (17-18) and (21-22). The measurable relative frequency offset y, is thus the following :

a) For a ramp at the input

$$y = r\left(\frac{T_c}{2\beta} - T_1 + \frac{T_c}{2}\right) \quad \text{in case A,} \qquad (23)$$

$$y = r\left(-T_1 + \frac{T_c}{2}\right) \quad \text{in case B.} \qquad (24)$$

b) For a sine-wave at the input

$$y = 2\pi f\left(\frac{T_c}{2\beta} - T_1 + \frac{T_c}{2}\right) C_0 \cos(2\pi f t) \quad \text{in case A,} \qquad (25)$$

$$y = 2\pi f\left(-T_l + \frac{T_c}{2}\right)C_0 \cos(2\pi ft) \qquad \text{in case B.} \quad (26)$$

This frequency offset depends on the loop gain in case A and, in both cases, on the detail and the duration of the atom-field interaction (through T_i) and its timing (through T_p).

It should be noted that, in case B, the steady state error due to a frequency ramp is not equal to zero and the rejection factor of a sine-wave is not of the second order with respect to $2\pi fT_c$. This is a weakness compared to the more traditional situation where the atom interrogation is performed continuously.

However, the drawback disappears if the following condition is satisfied :

$$T_l/T_c = 2. \quad (27)$$

It practically means that the atom-field interaction must be centred on the cycle of duration T_c. Then, one retrieves the familiar properties of continuous time frequency control loops. The steady state error and the rejection factor, respectively, given by Eqs. (23) and (25) are inversely proportional to the open loop gain in case A. In case B, the steady state error is equal to zero and the rejection factor is equal to zero to first order with respect to $2\pi fT_c$.

5. FREQUENCY INSTABILITY RELATED TO OSCILLATOR NOISE

If $S_y^f(f)$ is the power spectral density of the frequency noise of the free-running oscillator, the power spectral density, $S_y(f)$ of the slaved oscillator is given by :

$$S_y(f) = |H_{s,f}|^2 S_y^f(y), \quad (28)$$

assuming $2\pi fT_c << 1$. The transfer function proceeds from the results given in Section 4.3. We have :

$$H_{s,f}(f \to 0) = 2\pi i f T_c\left(\frac{1}{2\beta} + \frac{1}{2} - \frac{T_l}{T_c}\right) \qquad \text{in case A,} \quad (29)$$

$$H_{s,f}(f \to 0) = 2\pi i f T_c\left(\frac{1}{2} - \frac{T_l}{T_2}\right) \qquad \text{in case B.} \quad (30)$$

When condition (27) is fulfilled, the conclusions of Section 4.3 apply. It can be shown that, in case B, the transfer function is of second order with respect to fT_c.

This transfer function can also be derived by considering the frequency response associated with Eqs. (1,8,11) [4].

6. FREQUENCY INSTABILITY RELATED TO DETECTION NOISE

In general, the frequency instability related to the detection noise is larger than that due to the oscillator noise considered in Section 5. Since the time series $\{\delta\tilde{N}(t_k)\}$ represents a white noise process, with variance $\sigma^2_{\delta N}$, the one-sided power spectral density of the detection noise is given by :

$$S_{\delta\tilde{N}}(f) = 2T_c\sigma^2_{\delta N} \quad \text{, for} \quad 0 \le f \le 1/2T_c. \quad (31)$$

It is equal to zero for $f > 1/2T_c$.

The power spectral density of the relative frequency fluctuations of the controlled oscillator is then given by :

$$S_y(f) = \frac{1}{\omega_0^2}|H_{s,n}|^2 S_{\delta\tilde{N}}(f), \quad (32)$$

where $H_{s,n}$ is the frequency response to the input noise $\delta\tilde{N}$. Figure 6 shows the variation of $\beta^2|H_{s,n}|^2/K^2$ versus fT_c, for several values of β.

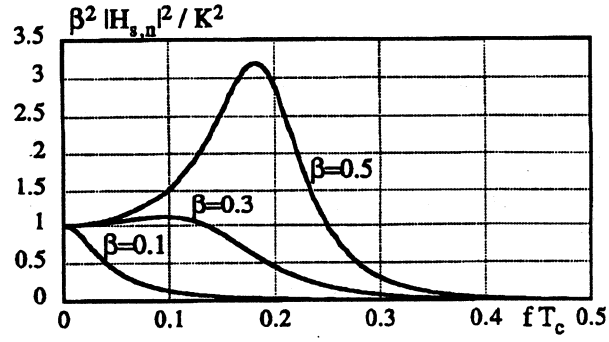


Fig-6. Power transfer function for detection noise

Assuming again $2\pi fT_c << 1$, we have, in cases A and B as well :

$$|H_{s,n}(f \to 0)|^2 = K^2/\beta^2, \quad (33)$$

and therefore, from Eqs. (9), (31), (32), (33) :

$$S_y(f \to 0) = \frac{2T_c}{\omega_0^2(\partial h/\partial\omega)^2}\frac{\sigma^2_{\delta N}}{N_0^2}. \quad (34)$$

In the special case where the atom-field interaction takes place according to the Ramsey method [7], we have :

$$h(\omega - \omega_0) = [1 + \cos(\omega - \omega_0)T]/2, \quad (35)$$

where it is assumed that the transit time T between the two oscillatory fields is much larger than the time spent in each of them. Usually, the modulation depth is equal to the half-width at half-maximum. We thus have $|\partial h/\partial\omega| = T/2$. The atomic quality factor Q_{at} being equal to $\omega_0 T/\pi$, Eq. (34) becomes :

$$S_y(f \to 0) = \frac{8T_c}{\pi^2 Q_{at}^2}\frac{\sigma^2_{\delta N}}{N_0^2}. \quad (36)$$

The related Allan variance is the following, for $\tau >> T_c$:

$$\sigma_y^2(\tau) = \frac{4}{\pi^2 Q_{at}^2}\frac{\sigma^2_{\delta N}}{N_0^2}\frac{T_c}{\tau}. \quad (37)$$

In the servo-loop, the quantity :

$$\nabla N(t_k) = \delta N(t_k) - \delta N(t_{k-1}), \quad (38)$$

which is available after the atomic line has been probed on both sides, is more easily accessible to measurement than $\delta N(t_k)$. It is thus useful to express the result of interest using $\sigma^2_{\nabla N}$ instead of $\sigma^2_{\delta N}$. The Allan variance becomes :

$$\sigma_y^2(\tau) = \frac{2}{\pi^2 Q_{at}^2}\frac{\sigma^2_{\nabla N}}{N_0^2}\frac{T_c}{\tau}. \quad (39)$$

It can be shown that Eq. (34) does not depend on the details of the processing of the error signal.

It should be noted that Eqs. (37) and (39), which have been derived rigorously, differ by a numerical factor from other ones previously published.

7. EXPERIMENTAL RESULTS

Figure 7 shows the experimental set-up [3] which has been implemented to verify the crucial points of the theoretical analysis. The high resolution frequency synthesizer n°1 is frequency locked to the atomic resonance. It receives from the computer : i) the frequency modulation signal at frequency $1/2T_c$ which is necessary to probe the atomic resonance, ii) the frequency control signal according to case A and iii) the frequency perturbing signal. The frequency synthesizer n°2 is identical to n°1. It receives the same signals as n°1, at the exception of the modulation. Its output frequency change is $\Delta\omega_s$ considered in the preceding sections.

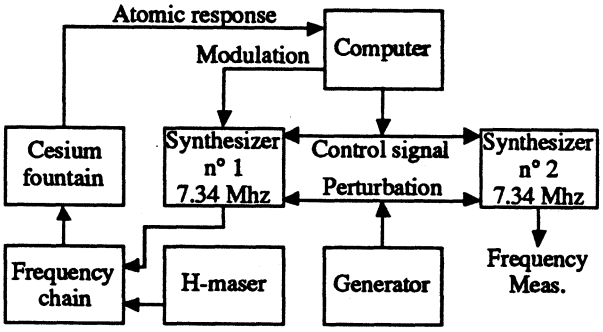


Fig-7. Experimental set-up

The response to a step function is shown in figure 8, where the value of β is known within a few percent. It closely agrees with the predicted one. The response to slow sinusoïdal perturbations has been recorded. Its shape is identical to that shown in figure 5 (bottom). These records have been used to verify the reality of the term T_1/T_c in Eq. (21), which is specifically due to the pulsed operation of the frequency standard. For that purpose, the transfer function :

$$|H'| = 2\pi f T_c \left(\frac{1}{2\beta} - \frac{T_1}{T_c} \right), \qquad (40)$$

relative to the sampled values $\Delta\omega_s(t_k)$ considered in Section 4.2 has been measured. It has been verified that the quantity $|H'|/fT_c$ is a constant for given values of β and T_1/T_c.

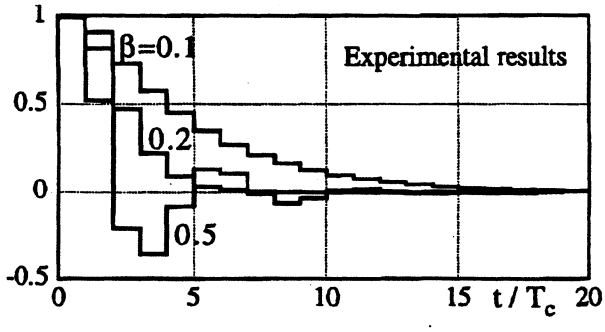


Fig-8. Experimental response to a step function

Figure 9 shows the variation of $|H'|/2\pi fT_c$ versus β, for $f = 25.10^{-3}$ Hz, $T_c = 1.095$ s and $T_1 = 0.503$ s. The 10 % uncertainty of the measurement is due to the frequency noise of the controlled oscillator.

The experimental results agree clearly with Eq. (40) which includes the term T_1/T_c.

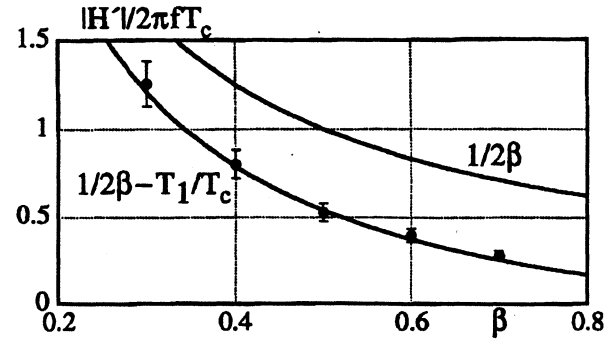


Fig-9. Experimental measurement of the variation of $|H'|/2\pi fT_c$ versus β.

8. CONCLUSIONS

We have established the main properties of an oscillator controlled by an atomic resonator operated sequentially. We have given the stability condition of the feedback loop and the condition to be satisfied to improve significantly the rejection of systematic frequency changes of the oscillator when a second numerical integration is included in the control algorithm. We have derived rigorously the expression for the Allan variance of the relative frequency fluctuations associated with the detection noise.
Experimental results confirm the theoretical analysis.

ACKNOWLEDGEMENTS

We are grateful to Centre National d'Etudes Spatiales for his financial support.

REFERENCES

[1] Prestage J.D., Tjoelker R.L., Dick G.J. and Maleki L. J. Mod. Opt., **29**, pp. 221-232, 1992.

[2] Sellars M.J., Fisk P.T.H., Lawn M.A. and Coles C. Proc. International Frequency Control Symposium , San Francisco, USA, pp. 66-73, 1995.

[3] Clairon A., Laurent P., Santarelli G., Ghezali S., Lea S.N. and Bouhara M. IEEE Trans. on Instrum. Meas., 44, pp. 128-131, 1995

[4] Audoin C., Santarelli G., Makdissi A. and Clairon A. To be submitted to IEEE Trans. on Ultrason., Ferroelec. and Freq. Control.

[5] Dick G.J. Proc. 19th Annual PTTI Applications and Planning Meeting, Redondo Beach, USA, pp. 133-147, 1987.

[6] Santarelli G., Laurent P., Clairon A, Dick G.J., Greenhall C.A. and Audoin C. Proc. 10th European Frequency and Time Forum, Brighton, UK, pp. 66-71, 1996.

[7] Ramsey N.F. Rev. Mod. Phys., 62, pp. 541-551, 1990.

A Novel Approach for the Development of Flywheel Oscillators for Atomic Clocks

Lute Maleki and Steve Yao

Jet Propulsion Laboratory, California Institute of Technology
4800 Oak Grove Dr., Pasadena, CA 91109
(818) 393-9031, Fax: (818) 393-6773
E-mail: Lute.Maleki@jpl.nasa.gov

We present theoretical and experimental results for a novel oscillator that converts continuous light energy into stable and spectrally pure microwave signals. This Opto-Electronic Oscillator (OEO), can generate ultra-stable, spectrally pure microwave reference frequencies up to 75 GHz with a phase noise lower than -140 dBc/Hz at 10 KHz.

The advent of the trapped ion and laser cooled atomic frequency standards with potential for stability beyond $10^{-13}/\sqrt{\tau}$ has created the need for high performance local oscillators. This is because these new atomic standards operate as frequency discriminators which periodically steer the frequency of a flywheel. The period of time required to perform this function can be as short as a second, and as long as many tens of seconds. During the intervals between each frequency comparison steps to steer the flywheel with the atomic line, the flywheel is free-running. Thus the performance of the flywheel for the duration of period between comparison steps determines the ultimate performance that may be obtained with the atomic standard. At the present time, conventional voltage controlled oscillators (VCO) based on quartz crystals significantly degrade the stability which the trapped ion standards are capable to produce.

We report here on a novel oscillator that converts continuous light energy into stable and spectrally pure microwave signals.[1] The OEO consists of a pump laser and a feedback circuit including an intensity modulator, an optical fiber delay line, a photodetector, an amplifier, and a filter. This device holds the promise for the development of a flywheel suitable for the new class of ultra-stable atomic standards.

A MINIATURE ULTRASTABLE QUARTZ OSCILLATOR WITH EXTREMELY GOOD PERFORMANCES

R.J. BESSON, M. MOUREY

Laboratoire de Chronométrie, Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques (ENS2M)
26 Chemin de l'Epitaphe - 25030 BESANÇON CEDEX - FRANCE

Abstract

Until recently, ultrastable quartz oscillators were fabricated in packages whose volumes ranged between 500 and 1000 cubic centimeters. In this paper it is shown that it is possible to considerably reduce volume of oscillators without sacrificing performances.

Basic concepts used for design of quartz oscillators are first rapidly introduced. The particular case of BVA oscillators is also considered because BVA resonators usually exhibit much lower noise level. A method to determine the elements of a quartz oscillator using the negative resistor concept is also recalled. Then several limitations due to various phenomena (including aging, g sensitivity, magnetic sensitivities, second level of drive, frequency jumps, amplitude frequency effects...) are considered. It is also shown that going to a small size by use of surface mounted components is favorable from many points of view.

As a consequence of previous studies a new oscillator extremely small (91 cm^3) has been developed. However main performance parameters are as follows :

aging : better than $2x10^{-11}$/day
Short term stability
at flicker level : $\leq 2x10^{-13}$
Spectral purity : 120 dBc/Hz at 1 Hz from carrier
155 dBc/Hz at 100 kHz from carrier
Stability versus temperature : $2x10^{-10}$ between (-20°C and +50°C)
Accelerometric sensitivity : $<10^{-10}$/g worst axis.

It is also pointed out that 30% of oscillators constructed have aging rates in the 10^{-12} range. Another model of larger volume (less than 450 cm^3) oscillator has been constructed either to obtain better aging rates (then 30% of production exhibit aging rates lower than 1 x 10^{-12}/day) or to obtain very low g sensitivity (better than 5 x 10^{-12}/g in a given plane).

From the upper results it is clearly seen that performances of miniature BVA ultrastable quartz oscillator can compare with usual rubidium standards, including from M.T.B.F. (Mean Time Betwwen Failures) point of view. As a conclusion performances of modern miniature ultrastable quartz oscillators will be reviewed and discussed.

I - INTRODUCTION

Reducing dimensions of a quartz Ultrastable Oscillator (U.S.O.) is a very interesting goal because volume and mass of such an oscillator is an important parameter in many projects. Unfortunately an important reduction in volume is a difficult task mainly because of thermal problems. In this paper we report on a 3 years program that ended at the beginning of 1996 allowing production of a new oscillator extremely small (91 cm^3). However we immediately point out that this effort has been

successfull basically because of the use of a miniature 10 MHz SC cut 3rd overtone bva resonator which had been extensively space evaluated [1]. Our decision to fullfill all quality requirements in resonator production seems to be rewarding in the long run for oscillators performances even if the cost of resonators cannot be drastically reduced under those conditions. In the some order of ideas, the fact that a precise P.I.D. (Process Identification Document) has been used is also very rewarding.

II - BASIC PRINCIPLES

Metrological oscillators are usually of the Pierce, Colpitts or Clapp type more or less modified according to the application. Those three types of oscillators correspond in fact to the circuit of Fig. 1 :

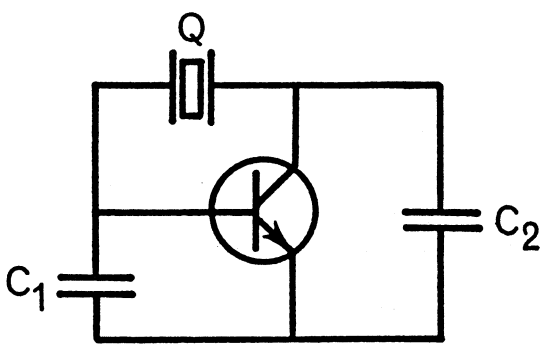


Fig. 1

DC voltages are not represented on Fig. 1 and the circuit is grounded either to emitter (Pierce oscillator) either to collector (Colpitts oscillator) or to base (Clapp oscillator). The oscillator finally choosen in industries or laboratories largely depends on available components and technologies [discrete components, surface mounted components (so said s.m.c.), hybrid circuits or integrated circuits]. They also depend on applications and on personal experience of designers. S.m.c. techniques allow a substantial decrease of volume together with greater reproducibility.

In the case of BVA oscillators practical limitations in $\sigma_y(\tau)$ usually come from the electronic sustaining circuit because BVA resonators usually exhibit much lower noise levels. A very useful concept is to consider that an oscillator is made of two complex dipoles one $R_1 + jX_1$ representing the resonator and the other $R_n + jX_n$ representing the sustaining circuit or negative resistor. A very simple calculation using this concept shows [2] that stability will be somewhat better if the motional inductance L_1 is high enough. (In our case L_1 is in the order of 2 H) but this also means that frequency pulling of resonator will be reduced. In other words the resonator has to be very well frequency adjusted.

III - OVERCOMING SEVERAL LIMITATIONS

We have already called for attention on several phenomena, namely :

- aging phenomena (resonator, oscillator)
- g sensitivities (resonator)
- magnetic sensitivities (resonator, oscillator)
- amplitude frequency effects (resonators)
- noise of components (incl. flicker of resonator)
- temperature fluctuations and variations in oven
- noise in various electrical connections between components.

All these phenomena have been carefully studied so as to obtain state of the art reduction. As a consequence only some components have been selected for use.

IV - REDUCTION OF OSCILLATOR'S VOLUME

This reduction has been first studied from theoretical point of view and all results have been carefully verified on successive experimental models. Specially selected surface mounted components have been used except for the loop transistor. Resonator is surrounded by electronic circuitry which constitutes the oven, the oven itself is surrounded by special foam. This detail is very important for thermal reasons. Finally the oscillator is hermetically sealed in a laser welded package under dry nitrogen. To prevent thermal effects of a 2g tip over a special insulating alumina powder can be used.

Finally it has been possible to reduce oscillator volume down to 91 cm^3 without sacrificing too much on performances.

V - PERFORMANCES OF MINIATURE ULTRASTABLE bva OSCILLATORS

19 units have been built in LCEP before going to larger scale on an industrial basis.
Units were under following specifications :

Power consumption	: 1,5 W
Aging	: better than $2x10^{-11}$/day
Short term stability at flicker level	: $\leq 2x10^{-13}$
Spectral purity	: better than 120 dBc/Hz at 1 Hz from carrier 155 dBc/Hz at 100 kHz from carrier
Stability versus temperature	: $2x10^{-10}$ between (-20°C and +50°C)
Accelerometric sensitivity	: < 10^{-10}/g worst axis
Magnetic sensitivities :	
resonator	: 2 x 10^{-12}/Gauss
oscillator	: a few 10^{-11}/Gauss
Barometric sensitivity	: < 10^{-10}/bar.

• Aging performances after 60 days. Following results have been obtained :

1 unit	< $4x10^{-11}$/day
11 nits	≤ $2x10^{-11}$/day
7 units in the low 10^{-12}/day range	

this includes two units whose aging was in the 10^{-13}/day range.

• Accelerometric sensitivity

All units exhibited sensitivities less than 10^{-10}/g worst axis.
As pointed out recently, it is possible to build improved oscillators [3] so that one of the external walls of the oscillator package is a plane, say P, of "zero" g sensitivity. This has been done using oscillator packages of less than 450 cm^3. Under those conditions g sensitivity in this plane P is less than $5x10^{-12}$/g for any acceleration contained in plane P.

• Larger packages

This larger package (450 cm^3) has also been used to house same ultrastable oscillator. Under those conditions, said oscillator exhibits very low aging rates since 30% of units show aging rates lower than $1x10^{-12}$/day.

VI - CONCLUSION

From the upper results it is clearly seen that performances of miniature bva ultrastable quartz oscillators can compare with those of usual rubidium sources. Choice between the two types of frequency standards largely depends on application. BVA quartz miniature U.S.O. is particularly suitable for applications where low consumption, excellent MTBF together with extremely good spectral purity are needed. It is to be pointed out that from the point of view of aging, g sensitivity and magnetic sensitivities quartz U.S.O. now come close to the regular rubidium units whereas their construction is simpler.

REFERENCES

[1] J.J. BOY, R.J. BESSON, N. FRANQUET, D. THIEBAUD, R. PETIT, Ph. BAVIERE, M. BRUNET, "Evaluation of mini-BVA resonators for space applications : tests program and results", 10th European Frequency and Time Forum, Brighton, 5-7 March 1996, Proc. pp. 31-36.

[2] R.J. BESSON, J.J. BOY, M.M. MOUREY, "BVA resonators and oscillators : a review. Relation with space requirements and quartz material characterization", 1995 IEEE international Frequency Control Symposium, Proc. pp. 590-599.

[3] R.J. BESSON, J.J. BOY, F. DEYZAC, "Acceleration sensitivity of BVA resonators", 1996 IEEE International Frequency Control Symposium, Proc. pp. 457-463.

MATERIALS

Chairman: P. W. Krempl

OH IMPURITIES IN GALLIUM ORTHOPHOSPHATE

F. Krispel°,*, P. W. Krempl°, P. Knoll*, W. Wallnöfer°

°AVL List GmbH, Kleiststraße 48, A-8020 Graz, Austria
*Karl Franzens Universität, Institut für Experimentalphysik, Universitätsplatz 5, A-8010 Graz, Austria

ABSTRACT

The synthetic crystal gallium orthophosphate ($GaPO_4$) possesses a very good stability in temperature up to 900°C and thus it is very suitable for high temperature applications. Since OH impurities influence the material properties their spectroscopy is investigated in more details. Infrared absorption in the range of 2600 to 3700 cm^{-1} is used to characterize the quality of $GaPO_4$ crystals. IR spectra of $GaPO_4$ show one strongly polarized OH band which is attributed to a gallium vacancy and a broad weakly polarized one. Most IR bands in the $GaPO_4$ spectra are assigned to intrinsic lattice vibrations of the crystal. It is possible to reduce the IR absorption of $GaPO_4$ by exposing the crystals to high temperatures.

Keywords: gallium orthophosphate, $GaPO_4$, IR, OH, high temperature

1. INTRODUCTION

Gallium orthophosphate belongs to the same point group as quartz (32), where silicon is substituted by gallium and phosphorus alternately. Thus the unit cell is about twice that of quartz in z-direction. In opposition to quartz $GaPO_4$ does not occur in nature (i.e. no natural seed plates). Compared with quartz there are some advantages as higher piezoelectricity, higher coupling constant and the much better temperature stability of all physical constants up to 900°C. The growth mechanism is hydrothermal because of the phase transition at 933°C, where the α-quartz structure changes to a cristobalite-like one. In quartz impurities are characterized mostly with infrared radiation by investigating the region near 3000 cm^{-1}, where only the light hydrogen atom vibrations absorb. The Q_{IR} value derived from the IR spectrum limits the maximal mechanical Q of a resonator, if it is designed well. Because of the similarities with quartz the IR absorption is also a good help to estimate the quality of $GaPO_4$ crystals.

2. INFRARED SPECTRA OF $GAPO_4$

2.1 Comparison of some quartz homeotypes

The infrared transmission spectra of quartz and two homeotypic materials are shown in Fig.1. The purest and best documented (Ref. 1) material is SiO_2, which shows narrow and strongly polarized bands with low absorption. Most bands in quartz are interpreted as OH defects near an aluminium atom substituting a silicon one. These OH impurities are responsible for the degradation of the intrinsic material quality, i. e. the quality factor Q (Ref. 2). $AlPO_4$ (berlinite) shows a high, broad, weakly polarized infrared absorption band without remarkable peaks. Thus most impurities are believed to be inclusions (growth solution or water). $GaPO_4$ possesses also a broad band which has medium polarization and absorption and in opposition to berlinite there are two strong absorption peaks too.

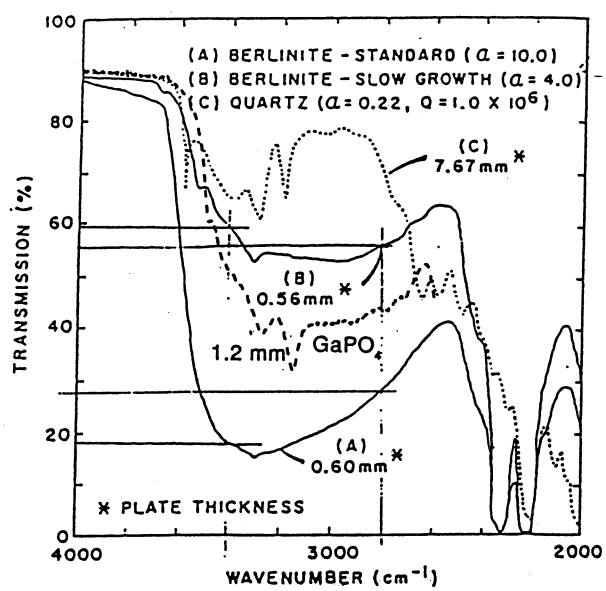


Fig. 1: Typical infrared transmission of three quartz homeotypic materials (Ref. 3): SiO_2 ....., $AlPO_4$ —— and $GaPO_4$ - - - -

2.2 Low temperature IR spectrum of $GaPO_4$

In most cases it is simpler to use the IR absorption instead of transmission:

$$\text{absorption}\left[\text{cm}^{-1}\right] = \frac{1}{\text{d[cm]}} \cdot \log \frac{T_{Ref}}{T}$$

d is the thickness of the plate in cm and T means the IR transmission in %. T_{Ref} is the transmission in % at a wavenumber, where no absorption should occur (3800 cm^{-1}).

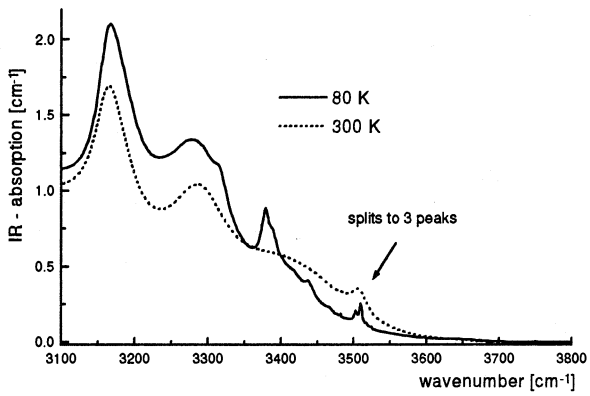


Fig. 2: Infrared absorption of $GaPO_4$ at 80K and 300K

In a typical infrared spectrum of gallium orthophosphate the broad band is not resolved down to 15 K (Fig. 2). But the band at 3508 cm^{-1} is splitted to three parts. Most bands don't change but some new ones appear which can be attributed to OH vibrations.

2.3 Polarized spectra:

Polarized infrared spectra of $GaPO_4$ at room temperature are shown in Fig. 3.

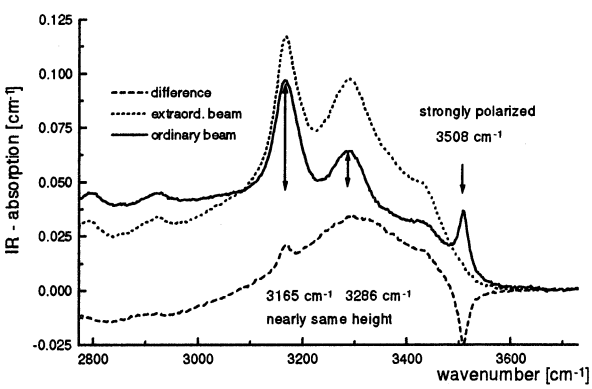


Fig. 3: Polarized IR absorption of $GaPO_4$ at 25°C

The extraordinary beam (electric field is polarized parallel to z) is absorbed more strongly in the region of the broad band than the ordinary one (electric field is polarized normal to z), but the 3508 cm^{-1} band vanishes. The difference between the extraordinary and the ordinary beam is also shown in the figure. The resulting broad band looks like the absorption of pure water (Ref. 4). The two significant bands at 3165 and 3286 cm^{-1} are weakly polarized only.

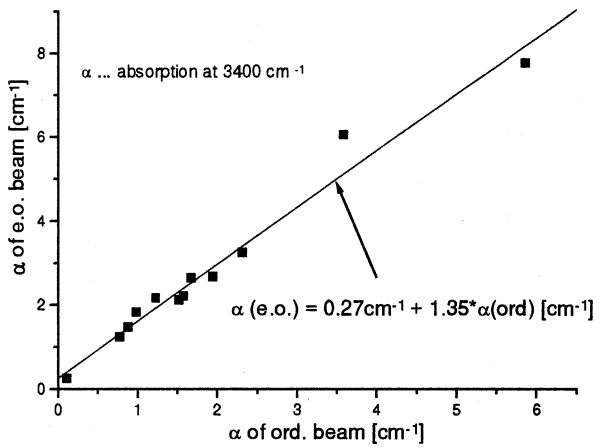


Fig. 4: Plot of absorption of extraordinary against ordinary beam at 3400 cm^{-1} for plates with different OH values

The absorption α at 3400 cm^{-1} of the extraordinary beam is plotted against the ordinary one (Fig. 4). The correlation of these two absorptions is linear. Its slope which gives the ratio of extraordinary to ordinary absorption at 3400 cm^{-1} equals 1.35, independent of absolute

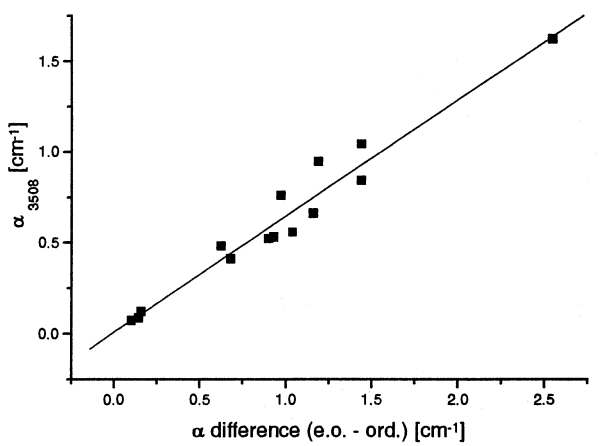


Fig. 5: Correlation between the absorption at 3508 cm^{-1} and the polarized part of the broad band

OH contents. Therefore attention should be given to the measurement of samples with different cuts because the absorptions of x- and z-cut samples are not the same. The peak at 3508 cm^{-1} seems also to be correlated with the difference of the absorption at 3400 cm^{-1} of extraordinary and ordinary beam (Fig. 5).

2.4 Intrinsic gallium orthophosphate bands

Many different samples have been compared by us to check if there are intrinsic vibrations around 3000 cm^{-1}. Some bands, especially the strong ones at 3165 and 3286 cm^{-1}, have always the same height of absorption peaks although the absorption of the broad band varies strongly (Fig. 6).

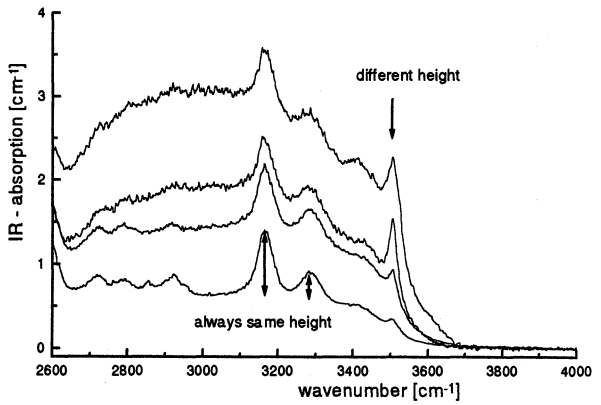


Fig. 6: Different $GaPO_4$ samples with different OH contents

Both bands are assigned to third overtone vibrations because at two thirds of their wavenumbers, bands with roughly the same absorption ratio can be detected if the samples are very thin (< 0.1 mm) as shown in Fig. 7.

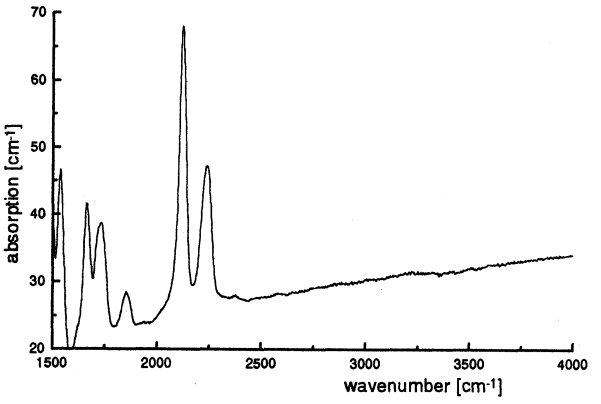


Fig. 7: Infrared absorption of a $GaPO_4$ sample with a thickness of 0.073 mm

It should be noted that some authors (Ref. 3, 5) attribute these two peaks to POH and GaOH vibrations.

2.5 Typical gallium orthophosphate spectrum

All measured spectra can be used to explain the infrared spectrum of $GaPO_4$ (Fig. 8): Most bands are intrinsic bands of $GaPO_4$, which is well seen in infrared spectra of samples with low OH contents as Fig. 11. Two peaks at 2850 and 2920 cm^{-1} can be avoided if the samples are cleaned carefully (shown later). The one real OH peak at 3508 cm^{-1} is strongly polarized and correlated with the broad background from 3600 cm^{-1} to 2600 cm^{-1} and lower.

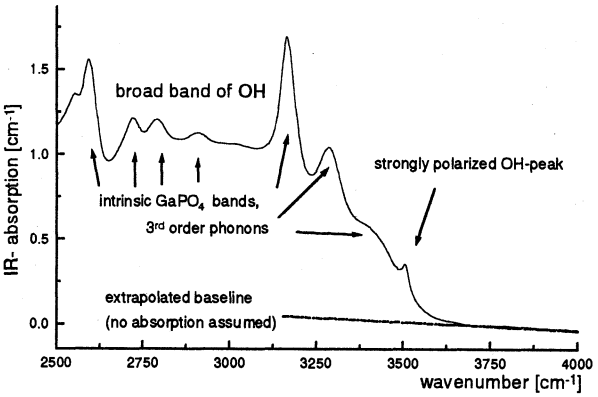


Fig. 8: Infrared absorption of $GaPO_4$

The absorption of the significant intrinsic bands are given below (mean value of 57 samples):

3165 cm^{-1} : 0.697 cm^{-1}
3286 cm^{-1} : 0.285 cm^{-1}
3400 cm^{-1} : 0.078 cm^{-1}

3. ESTIMATION OF OH CONTENTS

Different methods are in use to estimate the OH contents of $GaPO_4$ (Ref. 3, 5). We suggest:

$$\alpha\left[cm^{-1}\right] = \frac{1}{d[cm]} \cdot \log \frac{T_{3800}}{T_{3400}} - \alpha_0\left[cm^{-1}\right]$$

$$\alpha_0 = 0.078\ cm^{-1}$$

d means the thickness of the sample and T the transmission at 3400 (OH absorption) respectively 3800 cm^{-1} (no OH absorption assumed = baseline). α_0 is the absorption of the intrinsic lattice vibrations of $GaPO_4$ at 3400 cm^{-1} and can be neglected in most cases.

To get an idea about the quantity of OH (in ppm/wt) one can multiply the α - value with 65. This factor is estimated by the absorption of water in quartz glass (Ref. 6) but we have made no experiments to confirm this value, so we cannot recommend it. The wavenumber of absorption 3400 has historical reasons and originates from time when it was believed, that water is the dominant impurity in $GaPO_4$ crystals. Its infrared maximum value lays near 3400 cm^{-1} (Ref. 4).
Because of the partly polarization of the IR spectrum different cuts lead to different α values. For our estimation method we suggest to use a factor of 1.26 to get α_x or α_y out of the absorption of a z-cut, α_z.
If the surface is not polished, the baseline of the infrared absorption raises linearly with higher wavenumber. In Fig. 9 one spectrum represents a $GaPO_4$ sample with a rough surface and the second one shows the same sample coated with immersion oil. The difference of both is represented too. The latter curve shows that the baseline of a rough surface is really a straight line. The peaks at 2850 and 2920 cm^{-1} vanish by the use of immersion oil as well as by cleaning and are attributed to CH - vibrations by surface contamination as finger prints. An example of the determination of α at 3400 cm^{-1} is also shown. Both spectra result in the same α value.

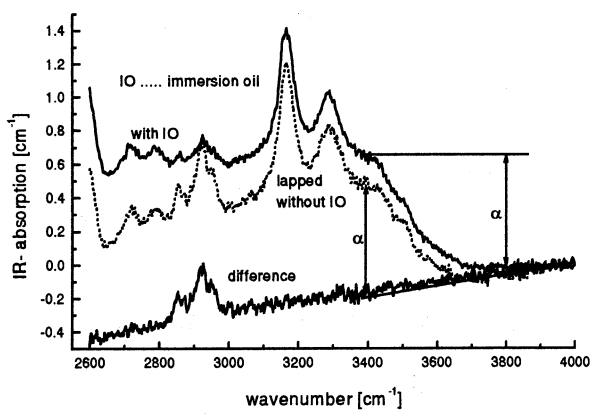


Fig. 9: Influence of a rough surface at the baseline of IR absorption

4. INFLUENCE OF OH

OH impurities influence many physical properties as the dielectric constants shown by Foulon et al. (Ref. 7). Also influenced are the quality factor of resonators Q, well established for quartz, and the electrical insulation. Frequency-temperature characteristics of the same cut show different behaviour, shown in berlinite (Ref. 3), if OH is very high.

$GaPO_4$ samples with higher OH contents become opaque if they are exposed to high temperatures. Two possible models can explain this fact. The first suggest the formation of water bubbles by single OH molecules which cause scattering. In the simpler second model little inclusions extend and destroy the lattice which causes the scattering.
Ways to reduce the OH content are known for quartz. Electrolysis under vacuum at high temperatures force the hydrogen atoms to leave the crystal. The influence of growth conditions has been investigated for the last years. High temperature growth and slow growth rates reduce the OH content. In the growth process of gallium orthophosphate OH is lowered by improving the transport of solution.

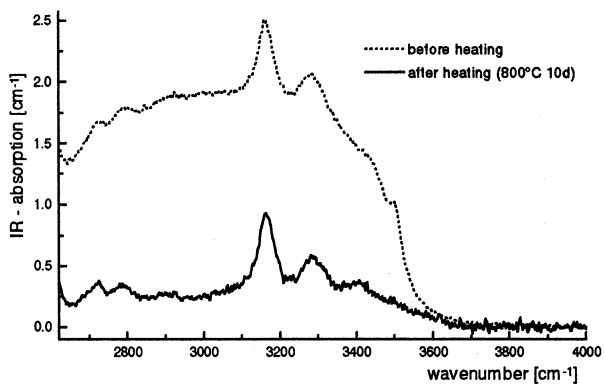


Fig. 10: Reduction of infrared absorption after heating to higher temperatures

If $GaPO_4$ crystals are tempered at very high temperatures, for example 830°C for 4 days, the OH absorption band in the infrared spectrum is reduced in a strong manner (Fig. 10). Apparently hydrogen diffuses out of the crystal. The disadvantage of this method is that all disturbtions stay which are caused during the growth process and where OH indicates the imperfections of the crystal.

5. DISCUSSION

For the absorption peak at 3508 cm^{-1} one possible model would be the occupation of a gallium vacancy by 3 hydrogen atoms. Studies on the influence of OH on physical properties are in work. The reduction method by heating the crystals is not satisfying so we research on lowering OH by growth processes. One infrared absorption spectrum of a gallium orthophosphate sample with an OH content near zero is shown in Fig. 11.

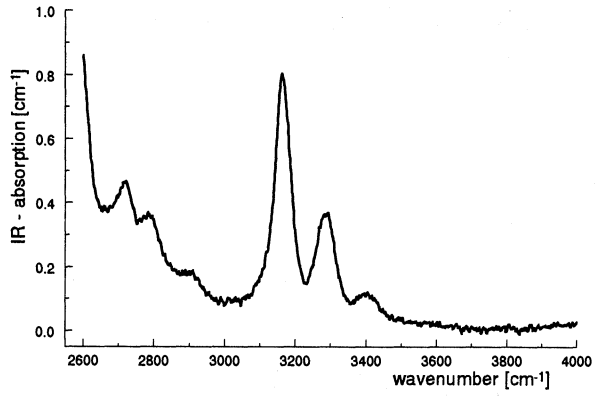


Fig. 11: Infrared absorption spectrum of a $GaPO_4$ sample with very low OH content

6. REFERENCES

1. A. Kats, Philips Res. Rep. 17 (1962), 133 - 195 and 201 - 279
2. D. M. Dodd, D. B. Fraser, J. Phys. Chem. Solids 26 (1965), 673 - 686
3. R. F. Steinberg, M. K. Roy, A. K. Estes, B. H. T. Chai, R. C. Morris, Proc. IEEE Ultrasonics Symp. 1 (Dallas, 1984), 279 - 284
4. R. D. Aines, G. R. Rossmann, J. Geophys. Res. 89 (1984), 4059 - 4071
5. D. Palmier, A. Goiffon, B. Capelle, J. Detaint, E. Philippot, J Cryst. Growth 166 (1996), 347 - 353
6. M. S. Paterson, Bull. Mineral. 105 (1982), 20 - 29
7. J. - D. Foulon, J. - C. Giutini and E. Philippot, Eur. J. Solid State Inorg. Chem. 31 (1994), 245 - 256

ACOUSTIC WAVES PROPAGATION IN LANGASITE UNDER UNIAXIAL STRESS

G.D. MANSFELD
Institute of Radioengineering and Electronics
11, Mokhovaya ave, 103907, MOSCOW - RUSSIA

J.J. BOY
Laboratoire de Chronométrie, Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques (ENS2M)
26 Chemin de l'Epitaphe - 25030 BESANÇON CEDEX - FRANCE

Experimental investigations of acoustic wave progagation in crystallographic directions (X, Y, Z) of Langasite crystals submitted to uniaxial pressure have been performed. In our case, the samples are parallepipedic and the pressure is applied in a direction perpendicular to this of the acoustic wave propagation which is generated by a ZnO electroacoustic transducer deposited on the largest surface. All surfaces of the samples are optically polished. The system LGS-ZnO creates a multi-frequency high overtones bulk acoustic waves composite resonator with resonances in the frequency range [0.1 - 1 GHz].

The relative frequency shifts due to the uniaxial stress belonging to the range [0 - 800 MPa] is linear and proportional to the "natural" sound velocity ; hence nonlinear constants can be measured.
Furthermore, we have analyzed the slopes of experimental curves which can be positive or negative depending on the directions of acoustic wave and strain.

Some experimental evidence of the onset of twinning when pressure is applied at 45° angle to the Z-axis of X-cut samples was found and also discussed.

HF LANGASITE MONOLITHIC FILTERS FOR GSM STANDARD

S.A.Sakharov, A.V.Medvedev, O.A.Buzanov
FOMOS Co.,

Bujeninova St., 16, 105023 Moscow, Russia

1. ABSTRACT

The paper describes technological peculiarities of the langasite crystal growth, HF monolithic filter and SAW filter fabrication. The results of works aimed at smaller defect number in crystals of greater diameters up to 3 inch are presented. Langasite resonator characteristics operating at the fundamental frequency up to 200 MHz are also given.

Keywords: langasite , crystal growth, resonator, filter, wafer, orientation

2. SINGLE CRYSTAL GROWTH

The procedure of initial charge preparation and langasite crystal growth is described in detail in our paper [1]. However, principal aspects have been revealed, which required a more detailed investigation.

The initial charge synthesis has been made by the method of selfdistributing high temperature synthesis (SHS). Lanthanum oxide (99.99%), silicon oxide (99.99%), gallium oxide (99.99%) and metal gallium (99.999%) have been used as initial components. Substitution of gallium oxide for a part of metal gallium during the SHS process compared with that described in [1], made it possible to significantly increase the part of melted phase (up to 85%) and, consequently to decrease losses of easily volatile components caused by evaporation at the stage of melt homogenization. The total impurity contents in the charge did not exceed their level in the initial components (not over $5*10^{-3}$ mass. %)

Langasite crystals have been grown by using Czochralski method in "Kristall-3m" pulling machine along the C-axis. The usual "as grown" crystals had the inscribed circumference diameter of the cylindrical part equal to 82 mm and the total length over 150 mm. The crystal mass exceeded 4 kg (Fig.1). The growth process of langasite crystals having such dimensions is connected with many difficulties. It was been found, in particular, that the material of the crucible and its geometrical dimensions are very critical for obtaining high quality langasite crystals. Besides of this, thermal gradients in the crystallization chamber, the geometrical shape of the conical part of the crystal and cinematic process parameters exert especially strong influence on the quality of the grown crystal. Crystal growing process proceeded with the availability of the flat crystallization front, which made it possible to avoid defects described in [2]. The problem consisted in the fact, that the growth of the whole cylindrical part occurred with the flat interphase boundary.

The "as grown" single crystals were subjected to high temperature annealing after growth on the air.

Fig.1 - Photo of the langasite crystals with the diameter of 86 mm

2. TECHNOLOGICAL ASPECTS OF CRYSTAL MECHANICAL TREATMENT

LGS crystal mechanical treatment has a number of peculiarities differing this technological process from the process of quartz crystal treatment. Langasite is considerably softer then quartz, its Moh's hardness is equal to 5.5. Technological aspects of LGS mechanical treatment have been described in detail in papers [3,4]. In the course of this work the main attention was paid to langasite wafer treatment having large (3 inch) diameter.

The main problems for wafer manufacturers with one-sided polishing are the following ones:

- obtaining defectless polished surface (absence of scratches, cracks, microinclusions at the polished surface;
- nonuniform thickness at wafers with different treatment of their sides.

For mechanical wafer lapping the diamond laps with special bonded abrasive have been used, the abrasive grain size was 80/63. The use of bonded abrasive enabled us to obtain a satisfactory planeparallelism and thickness nonuniformity.

The main peculiarities of wafer polishing are connected with the use of polishing slurries based on silica sol (colloidal solution) with surface active additives analogous to Nalco and Syton.

The use of cerium-, iron- and aluminum oxides does not allow to obtain a wafer surface free of scratches. It is necessary to note also the influence of polishing pad material on the shape of wafers obtained.

The use of the soft materials (polyvel) ensures the best abrasive material holding at the surface, but with this wafer edge rounding downwards occurs due to wafer pressing into the polishing pad material. When rigid polishing pads are used (polyurethane foam) it is necessary to take into account the fact that these materials should not contain metal oxide inclusions and should hold well polishing soles at their surface.

Chemical etching of langasite crystal wafers is also one of the most important problems in obtaining high quality products operating both at bulk acoustic waves (BAW) and surface acoustic waves (SAW).

Langasite crystals can practically be etched in all acids, but the etching rate differs significantly with different acids, etching temperatures and concentrations. The main problem during the technological etching process development consists in finding suitable polishing composition.

We have investigated different etchants influence on the surface quality obtained. The best results have been achieved when etching langasite crystals in a mixture of hydrofluoric and hydrochloric acids. This composition has been used for etching LGS-wafers for resonators operating at the fundamental frequencies of 180- 210 MHz. Langasite plates have been polished up to the thickness of 70 μm , then reversed mesastructure has been etched up to the thickness of 6 μm, i.e. approximately 32.0 μm has been etched off from each side. The wafer surface in the region of mesastructure has no traces of selective etching at which defect revealing occurs at the surface and within the crystal volume causing through etching of a crystal surface as a result.

The surface quality obtained with plate thickness of 6 μm enabled us to manufacture resonators with the parameters listed in table 1.

Table 1

Fs	dF_{ar}	Q	R_1	L_1	C_1	C_0
MHz	%		Ω	H	pF	pF
186.0	0.218	2200	56.8		0.013	4.35

The technological process of chemical reversed mesastructure forming at polished LGS crystal plates has been used by the authors for monolithic filter manufacture operating over the frequency range from 30 to 100 MHz.

3. HIGH-FREQUENCY LANGASITE MONOLITHIC FILTERS

In our earlier works we have many times described the application of langasite crystals for the development and manufacture of monolithic filters. These filters operating over the frequency range from 2 to 25 MHz are manufactured by using conventional technology without the use of mesastructure. The typical characteristic of the 8th order MF at 10.7 MHz with the pass band ± 25 kHz is presented at Fig. 2.

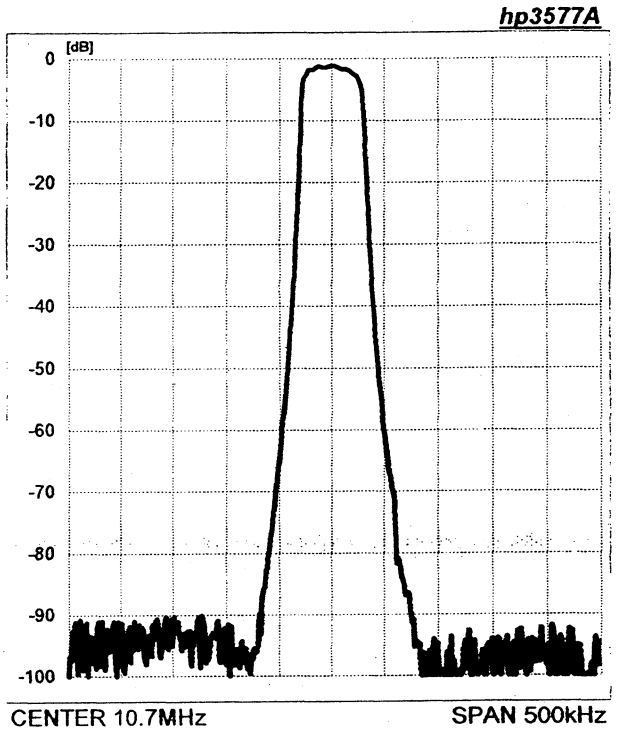


Fig. 2 - Amplitude-frequency characteristic of LGS MF at 10.7 MHz, n= 8

Table 2 lists the main characteristics of langasite monolithic filters operating over the frequency range from 2 to 25 MHz. The principal attention in this work has been paid to the development of high-frequency LGS monolithic filters. In the high-frequency region from 50 MHz and above the electrode dimensions have become smaller as well as the dimensions between them. For MF manufacture the mask technology of electrode evaporation is used. At present it is possible to fabricate masks with the minimum interelectrode distance of 90 μm. This causes the limitation of the pass band of high frequency filters. For quartz monolithic filters operating over the frequency range from 50 to 100 MHz the optimum

Table 1

TYPE	NOMINAL FREQUENCY	ORDER	PASSBAND		ATTEN. BANDWIDTH		RIPPLE	INSERTION ATTEN.	OPERATING TEMP.
	MHz		dB	kHz	dB	kHz	dB	dB	°C
FP3P6-1.9-18	1.992	2	3	±8.8	15	±22	1.0	2.0	-40...+70
FP3P6-10.7-50	10.7	2	3	±25	20	±80	0.5	1.5	-40...+70
FP3P4-575	10.7	8	6	±25	80	±75	2.0	4.0	-40...+70
FP3P6-10.7-75	21.4	2	3	±37.5	20	±125	0.5	1.5	-40...+70
FP3P6-25-100	25.0	2	3	±50	20	±200	1.0	2.0	-40...+70
FP3P6-23-80	23.0	2	3	±40	20	±135	0.5	1.5	-40...+70

of langasite crystals in an analogous frequency range allows to manufacture MF with the relative pass band from 0.2 to 0.6%.

The authors have developed and manufactured MF at the frequency of 71 MHz with the pass band +90 kHz.

The most interesting from these filters are filters at 71 MHz, which is the intermediate frequency for cellular radio telephones of GSM standard. At present analogous quartz crystal filters are manufactured both with discrete elements and in a monolithic form. The main drawbacks of such filters are their large size and the value of spurious responses in the stop band. In Fig 3 the amplitude-frequency characteristic of langasite MF at 71 MHz is given.

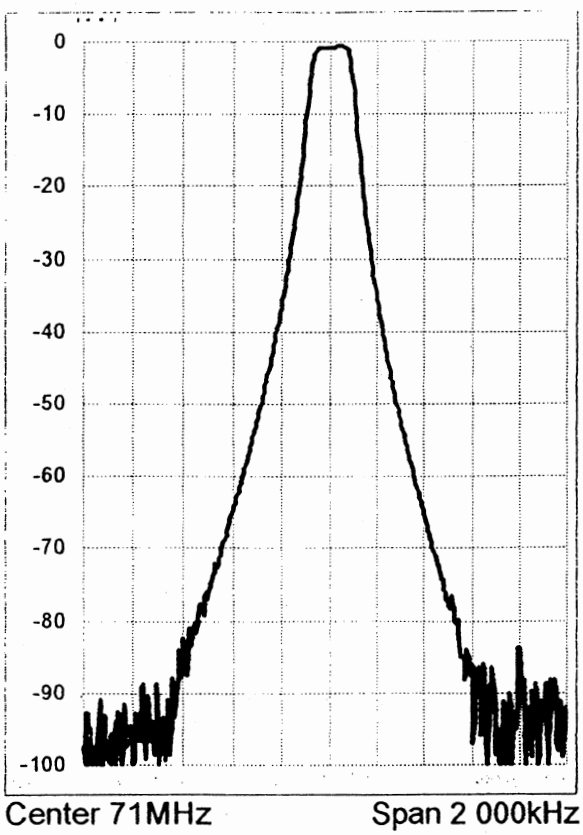


Fig.3 - Amplitude-frequency characteristic of LGS MF at 71 MHz, n = 4;

Fig. 4 shows group delay distortion of this filters. Fig 5 shows package outlines of the 4th order MF at 71 MHz.

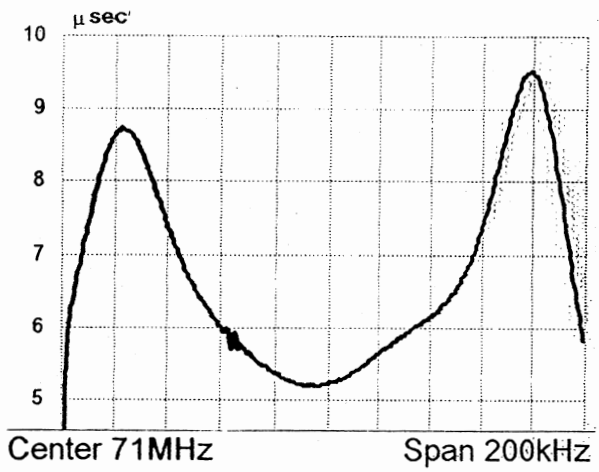


Fig.4 - Group delay of LGS MF at 71 MHz,

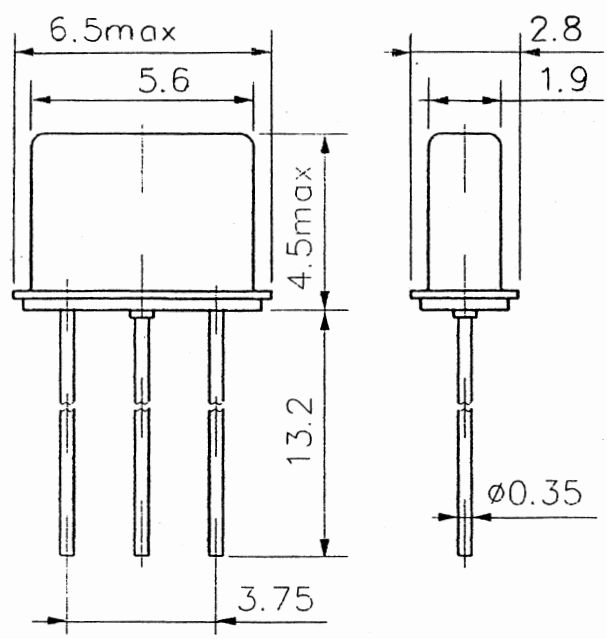


Fig.5- Package outlines of LGS MFat 71 MHz

Table 3 lists the main parameters of this LGS monolithic filter.

Table 3

1	Center Frequency, f_0	71 MHz
2	Min. Pass Bandwidth, $\Delta f_{3\,db}$	± 80 kHz
3	Max. Pass Bandwidth, $\Delta f_{3\,db}$	± 95 kHz
4	Min. Stop Band Attenuation:	
	at $f_0 \pm 130$ kHz	13 dB
	at $f_0 \pm 150$ kHz	17 db
	at $f_0 \pm 200$ kHz	28 dB
	at $f_0 \pm 400$ kHz	50 dB
	at $f_0 \pm 600$ kHz	65 dB
	at $f_0 \pm 800$ kHz	70 dB
	at $f_0 - 12$ MHz	85 dB
5	Min. Spurious:	
	f_0 to 500 kHz	30 dB
	+500 kHz to +910 kHz	42 dB
6	Min Spurious Guardband:	
	f = 71600 ± 2 kHz	60 dB
	f = 71800 ± 2 kHz	70 dB
	f = 72000 ± 2 kHz	70 dB
7	Max. Insertion Loss (at f_0)	5 dB
8	Max. Pass Band Ripple (within $f_0 \pm 45$ kHz)	1 dB
9	Max. Group Delay Distortion (within $f_0 \pm 50$ kHz)	1.5 μsec
10	Input Level	- 2 dBm
11	Operating Temperature Range	0 - 70 °C

4. LANGASITE CRYSTAL APPLICATION FOR SAW FILTERS

Results of large diameter LGS crystals growth enabled the authors to make a number of investigations aimed at application of this crystal for surface acoustic wave (SAW) filters. As it has been described in clause 1, these crystals were grown along the Z-axis and , as a consequence of this, SAW characteristics were investigated around Z-orientations. Three inch diameter wafers with one-sided polishing have been fabricated for this purpose. Interdigital electrodes have been evaporated by photolithography method. The main task for the authors was the search for orientations with the minimum frequency vs temperature dependence. The orientations around the Z-axis have been investigated with acoustic wave distribution along the Y-axis. The temperature coefficient of frequency has the values 10 ppm.

The SAW velocity is 2460 m/c.

The authors consider that the orientations exist in langasite crystals with higher temperature stability , allowing to apply these crystals for SAW filters. For exact determination of these orientations more thorough investigations are necessary both around Z-orientations and in other directions. At present the works aimed at growing langasite crystals along the Y-axis are initiated. Fig.6 shows the characteristic of a SAW filter manufactured at langasite substrate.

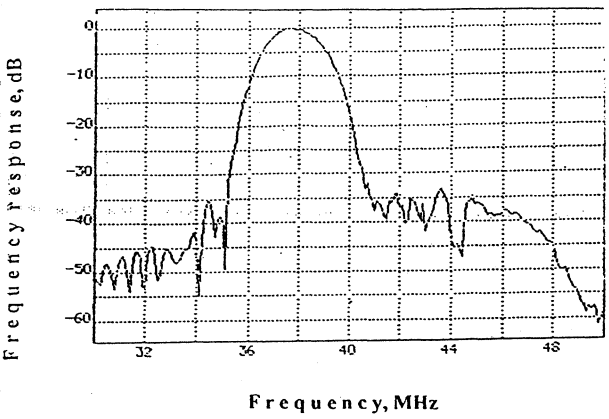


Fig.6 - Amplitude-frequency characteristic of LGS SAW filter

5. CONCLUSION

Results obtained by the authors in langasite HF resonator and filter manufacture, in growing langasite crystals of large diameter with a satisfactory quality show the possibility of this material application in piezoelectric engineering. The most promising area of this material use seems to be SAW filters.

The authors hope that the results presented will find interest of specialists working in the field of piezoelectric engineering and acoustoelectronics.

6. REFERENCES

1. O.A.Buzanov, A.V.Naumov, V.V.Nechaev, S.N.Knyazev. New approach to the growth of langasite crystals.- Proc. 1996 IEEE Int. Frequency Control Symp., pp. 131-136.

2. K. Shimamura, H. Takeda, T. Kohno, T. Fukuda. Growth and characterization of lanthanum gallium silicate $La_3Ga_5SiO_{14}$ single crystals for piezoelectric applications.- Journal of Crystal Growth, v. 163 (1996), pp. 388-392.

3. S. Laffey, M.Hendrickson and J.R.Vig. Polishing and etching langasite and quartz crystals. - Proc. 1994 IEEE Int. Frequency Control Symp.

4. S.Sakharov, Y.Pisarevsky, A.Medvedev. Surface and volume defect in langasite crystals.- Proc. 1994 IEEE Int. Frequency Control Symp. pp. 642-646

SOME HIGHER ORDER EFFECTS IN THE ELECTROMECHANICAL SYSTEMS BASED ON VOLUME AND SURFACE ACOUSTIC WAVES

Jaroslav NOSEK
Technical University of Liberec
Hálkova 6, CZ 461 17 LIBEREC , Czech Republic

ABSTRACT

The paper deals with properties of piezoelectric resonators and resonant mode sensors and focused on non-linear effects, especially in elastic domain. The effects of these nonlinearities in electrical circuits are considered together with the possibilities of using them in the sensors. The effect of excitation and intermodulation in the BAW quartz resonators is described and measured.Futher, an attention is paid to the change of velocity of SAW propagation on the quartz delay line, caused by electric field. The arrangement of new electric field sensor is presented.The optimal orientation and dimensions of X-length strip quartz resonator as a temperature sensor are found.

Keywords: non-linearities of the BAW and SAW quartz resonators, some results

1. NON-LINEAR ELECTRO-ELASTIC EQUATIONS

The linear theory of piezoelectricity is a suitable tool for describing the propagation of acoustic waves with small amplitudes. Also, the vibrating solids must be subjected to any elastic prestressing or strong electric field. The non-linear theory takes into consideration the final deformation η_{ij}. Elastic stiffness of higher order can be derived from the internal energy U relating to a unit volume of the given substance by

$$c_{ijkl\ldots pq} = \left(\frac{\partial^n U}{\partial \eta_{ij} \partial \eta_{kl} \ldots} \right)_{D,\sigma = konst.}, n \geq 2 \qquad (1)$$

where n is the order of stiffness. As shown in (Ref. 1), the elastic coefficients of n-order, as well as higher order coefficients characterising the dielectric, piezoelectric and pyroelectric properties of crystals, can be derived from various thermodynamic potentials. The electro-optical, electrostriction and other coefficients can be specified by the same method. The thermodynamic stress t_{LM}, the thermodynamic electric displacement $\mathcal{D}_N$ and the thermo-dynamic intensity of electric field $\mathcal{E}_N$ can be taken as the basis for the study of non-linear electroelastic properties of piezoelectric substances, the material co-ordinates a_N are denoted by capital letters for the study of material properties of piezoelectric substances. The relationships between the thermodynamic parameters $\mathcal{D}_N$, $\mathcal{E}_N$ and D_i , E_i (electric displacement and intensity of the electric field), in a space -coordinate x_i are the following:

$$\mathcal{D}_N = J \frac{\partial a_N}{\partial x_i} D_i, \quad \mathcal{E}_N = \frac{\partial x_i}{\partial a_n} E_i \qquad (2)$$

where J is the functional determinant (Jacobean). For an adiabatic process, the components $t_{LM}^{\varepsilon,\sigma}$ of thermodynamic stress tensor can be expressed

$$t_{LM}^{\varepsilon,\sigma} = c_{LMCD}^{\varepsilon,\sigma} \cdot u_{C,D} + \frac{1}{2} c_{LMCD}^{\varepsilon,\sigma} \cdot u_{K,C} \cdot u_{K,D} + \frac{1}{2} c_{LMCDEF}^{\varepsilon,\sigma} \cdot u_{C,D} \cdot u_{E,F} + \frac{1}{2} c_{LMCDEF}^{\varepsilon,\sigma} \cdot u_{K,C} \cdot u_{K,D} \cdot u_{E,F} + e_{NLMCD}^{\sigma} \cdot \phi_{,N} \cdot u_{C,D} + e_{NLM}^{\sigma} \cdot \phi_{,N} - \frac{1}{2} H_{NALM}^{\sigma} \cdot \phi_{,A} \cdot \phi_{,N} \qquad (3)$$

$$\mathcal{D}_N^{\eta,\sigma} = e_{NAB}^{\sigma} \cdot u_{A,B} + \frac{1}{2} e_{NAB}^{\sigma} \cdot u_{K,A} \cdot u_{K,B} + \frac{1}{2} e_{NABCD}^{\sigma} \cdot u_{A,B} \cdot u_{C,D} + \frac{1}{2} e_{NABCD}^{\sigma} \cdot u_{K,C} \cdot u_{K,D} \cdot u_{A,B} - H_{NMAB}^{\sigma} \cdot \phi_{,M} \cdot u_{A,B} - \frac{1}{2} H_{NMAB}^{\sigma} \cdot \phi_{,M} \cdot u_{K,A} \cdot u_{K,B} + \varepsilon_{NM}^{\sigma} \phi_{,M} + \varepsilon_{NMP}^{\sigma} \cdot \phi_{,D} \cdot \phi_{,M} \qquad (4)$$

where symbols $c_{LMCD}^{\varepsilon,\sigma}$ and $C_{LMCDEF}^{\varepsilon,\sigma}$ denote elastic stiffness of the second and third order, symbols e_{NAB}^{σ} and $e_{NABCD}^{\varepsilon,\sigma}$ denote linear and quadratic piezoelectric stress-tensor components, symbol H_{NALM}^{σ} denotes the electrostriction coefficients, and the symbols $\varepsilon_{NM}^{\eta,\sigma}$ and $\varepsilon_{NMP}^{\eta,\sigma}$ denote the components of the tensor of linear and quadratic permittivities. The non-linear piezoelectric stresses e_{NABCD}^{σ} read

$$e^{\sigma}_{NABCD} = -\frac{\partial^2 t^{\varepsilon,\sigma}_{LM}}{\partial \eta_{CD} \partial \varepsilon_N} \qquad (5)$$

and non-linear piezoelectric strains

$$d^{\sigma}_{NABCD} = \frac{\partial^2 \eta^{\varepsilon,\sigma}_{LM}}{\partial t_{CD} \cdot \partial \varepsilon_N} \qquad (6)$$

In the simplified pure elastic case, the description of propagation of a finite amplitude wave in a non-linear medium leads to the wave propagation equation with effective non-linear elastic constants of the 3rd and 4th orders. In the presented problem these non-linearities are at the origin of harmonic generation with its consequences

- amplitude - frequency effect, and
- intermodulation.

2.AMPLITUDE-FREQUENCY EFFECT

To some extent, the resonant frequency of piezoelectric resonators depends on the magnitude of excitation current passing through the resonator. Considering the results of (Ref. 2), the following approximate relationship for the relative change in the resonant frequency of AT resonators (in shape of a the plate) is used:

$$\frac{\Delta f}{f} = \frac{9}{16} \cdot \frac{\gamma}{c^D_{66}} \cdot \frac{h^2 \pi^2}{8 \cdot e^2_{26}} \cdot \left[\frac{\varepsilon_{22}}{(2a)^2} \left(1 + \hat{k}^2_{26} \right) V_K^2 + \frac{I_K^2}{\omega_0 (2b \cdot 2l)^2} \right], \qquad (7)$$

where

$$\gamma = \frac{1}{2} c^E_{22} + c^E_{266} + \frac{1}{6} c^E_{6666}, \qquad \hat{k}^2_{26} = \frac{k^2_{26}}{1 - k^2_{26}} \qquad (8)$$

and $2a$ is the resonator thickness, c^E_{22}, c^E_{266}, c^E_{6666} are elastic stiffness of the second, third and fourth orders, respectively, measured under a constant electric field, e_{26} is the piezo-electric stress, k_{26} is the electromechanical coupling coefficient, ε_{22} the permitivity along the thickness axis of the plate, 2b, 2l are the width and length, respectively, of a fully plated resonator, ω_0 is the angular resonant frequency, V_K is the voltage and I_K is the excitation current applied to the resonator.

In Eqn. (7) , the first term in the square brackets is usually negligible compared to the second term, and Eqn. (7) acquires the form

$$\frac{\Delta f}{f} = B I_k^2 \qquad (9)$$

where

$$B = \frac{9}{16} \frac{\gamma}{c^D_{66}} \frac{h^2 \pi^2}{8 e^2_{26}} \frac{1}{\omega_0 (2b \cdot 2l)^2} \qquad (10)$$

The influence of the excitation current on the frequency and its mathematical description makes necessary to introduce a non-linear impedance characteristic of the piezoelectric resonator (Nosek in Ref. 1). This influence was modelled by the non-linear equivalent circuit, as shown in Fig. 1.

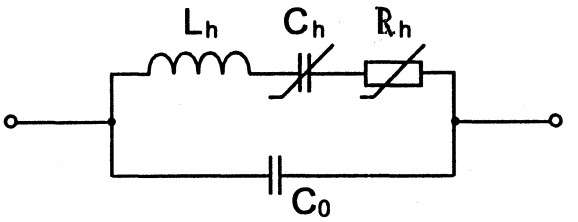


Fig. 1: Non-linear equivalent electrical circuit for piezoelectric resonator

The equivalent series resistance $R_h(I)$ and dynamic capacitance $C_h(I)$ are taken to be functions of the amplitude of the excitation current:

$$R_h(I) = R_h(1 + \beta I^2) \qquad (11)$$

$$C_h(I) = C_h \frac{1}{1 + 2BI^2} \qquad (12)$$

R_h and C_h are the equivalent series resistance and dynamic capacitance relating to infinitely small amplitudes of excitation current, the magnitudes of the constants β and B depend on vibration mode, elastic and piezoelectric properties of the resonator and the dimensions of the electrodes. The constant B relating to AT resonator is defined by Eqn.(10). We analysed the above equivalent circuit by the method of equivalent linearisation. This method consists of determining the average values R_{hs}, C_{hs} of the resistance $R_h(I)$ and the capacitance $C_h(I)$, which fulfil the condition of equal voltage drops across the non-linear elements for first harmonics of the excitation current. It follows from this condition that

$$R_{hs} = R_h \left(1 + \frac{3}{4} \beta I_1^2\right) \qquad (13)$$

$$C_{hs} = C_h \frac{1}{1 + BI_1^2} \qquad (14)$$

where I_1 is the amplitude of the first harmonic of the current flowing through the dynamic branch of the equivalent circuits, i.e. through R_{hs}, C_{hs} and L_{hs} in series. The relationships between the amplitudes of the voltage V_1 and the first current harmonics I_1 read

$$I_1 = \frac{V_1}{\left[R_h^2(1+\frac{3}{4}\beta I_1)^2 + (\omega L_h - \frac{1+BI_1^2}{\omega C_h})^2\right]^{\frac{1}{2}}} \quad (15)$$

and is shown in Fig. 2.

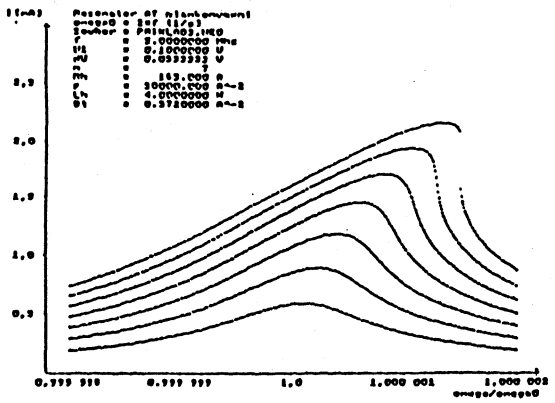


Fig. 2:Computed dependence $I(\omega/\omega_0)$ with parameter V_1 (AT resonator 5 MHz).

Calculated dependence of the phases - frequency with parameter V_1 is given in Fig.3.

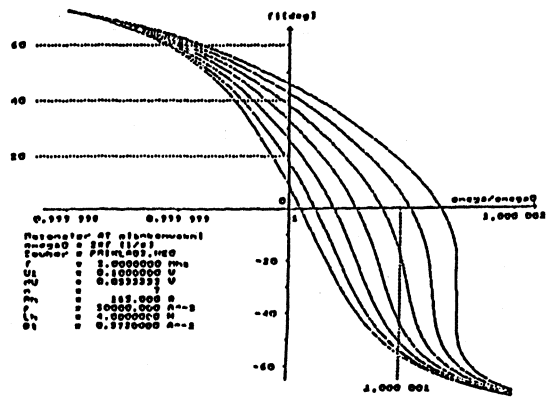


Fig. 3:Phases- frequency dependence of the excited AT resonator 5 Mhz.

3.INTERMODULATION OF PIEZOELECTRIC RESONATORS

If two (or more) harmonic signal levels V_1 and V_2 and frequencies ω_1 and ω_2 positioned uniformly with respect to resonant frequency ω_N act simultaneously, the intermodulation signal of the voltage $\hat{V}_\Omega$ and angular frequency $\Omega = 2\omega_1 - \omega_2$ is created due to the cubic non-linearities. This frequency Ω will be located also within the bandwidth, and, therefore will not be filtered.

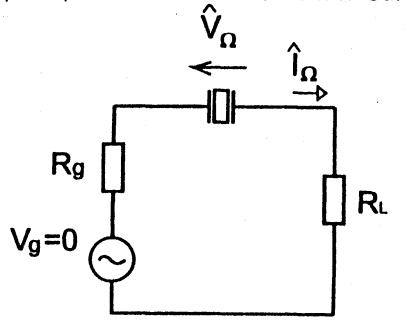


Fig. 4: Circuit with intermodulation current

The complex value of the intermodulation voltage for the circuit in Fig. 4 can be derived from (Ref. 5). After a simple arrangement we get an interesting relation for $\hat{I}_\Omega$:

$$\hat{I}_\Omega = -(\hat{Y}_\Omega \cdot \hat{V}_\Omega + jA_{1\Omega} \cdot A_4 \frac{V_1^2 V_2}{\Delta\Omega}), \quad (16)$$

where $\hat{Y}_\Omega$ is the resonator admittance at angular frequency Ω. The significance of symbols for admittance $\hat{Y}_\Omega$, $A_{1\Omega}$, A_4 and $\Delta\Omega$ are shown in (Ref. 5).

It follows from Eqn. (16), that the intermodulation current consist of a component dependent on the admittance $\hat{Y}_\Omega$ and a component affected by control signals of levels V_1, V_2, and effective elastic stiffness γ, which include the linear and non-linear elastic stiffness of 2nd, 3rd and 4th order.

The formulae given above makes the selection of a suitable method for the measurement of the intermodulation signal magnitude possible.

A measuring set schematically showing Fig. 5 contains generators in the part IM. The used generators are crystal controlled oscillators of 10 MHz with mistuning of $\pm 1 \cdot 10^{-3}$ and stability of the order $(10^{-6} \div 10^{-7})$ / day.

Mixing circuit SO is a key part of the control signal block. It makes possible to increase the insertion loss of channel A with respect to generator approximately to 40 dB and the second harmonic rejection increases to -- 85 dB. The forward direction insertion loss is approximately 1 dB. The

mixing network consists of the crystal bandpass filters of central frequencies f_A = 9.997,7 kHz and f_B = 10.002,3 kHz symmetrical with respect to the nominal frequency 10 MHz. Insertion loss characteristic of realised mixing circuit are in Fig. 6.

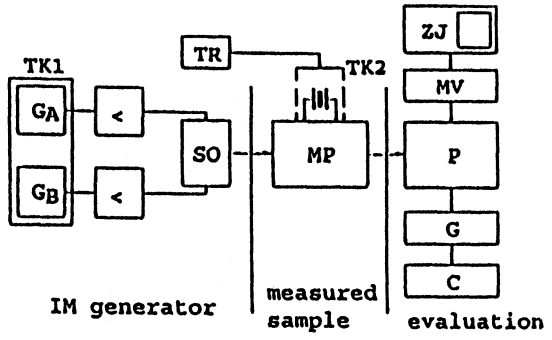


Fig. 5: Measurements of IM products

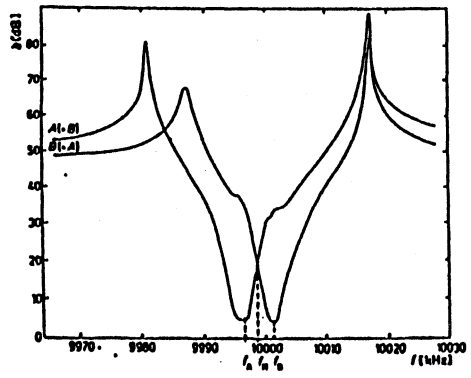


Fig. 6: Measured insertion loss characteristic of mixing circuit.

Evaluation part consist of commercial set with measuring transmitter, generator and measuring receiver.

The measured values V_L = - 98 dB and c^D_{6666} = 76 x 10^{-11} Nm^{-2} (V_g = 4,5 dB, Δf = +- 1 kHz, f_N = 9.995,35 kHz, R_L = 135 Ω) are valid for the AT-cut quartz resonator of 10 MHz.

4.PIEZOELECTRIC SUBSTRATE SUBJECTED TO ELECTROSTATIC FIELD

The effect of a strong electroelastic field acting simultaneously with a small varying field was studied for the first time for GT - cut quartz resonators at the TU Liberec in 1961. The resulting effect was termed by Hruska as a polarisation effect (Ref. 7 and other papers). In a later work (Ref. 4) the effect of a strong electrostatic or slowly varying electric field acting on a piezoelectric cut vibrating near its resonance was described with the aid of a change of elastic modules. The elastic modules c^E_{LMCD} change on the value of c^*_{LCMD} and piezoelectric modules e_{LMN} change on e^*_{LMN}, and, are expressed by these relations

$$c^*_{LMCD} = c^{\varepsilon}{}_{LMCD}\left(1 + \frac{e_{NLMCD}}{c^{\xi}{}_{LMCD}}\phi_{,N}\right) \qquad (17)$$

and

$$e^*_{NLM} = e_{NLM}\left(1 - \frac{1}{2}\frac{H_{NALM}}{e_{NLM}}\phi_{,A}\right) \quad , (18)$$

The symbol c^{ε}_{LMCD} was used for the components of the second order linear elastic modules measured at constant thermodynamic electric field strength. The symbol ϕ is the potential related to the electric field components by $\varepsilon_N = -\phi_{,N}$ and H_{NALM} are components of the electrostriction coefficient. Due to the fact that elastic modules c^*_{LCMD} and piezoelectric modulus e^*_{NLM} are linear functions of the electric field ε_N, the above mentioned module will change by the activity of the electric field, and the amplitude of the change will be a linear function of the field.
Let us assume a thin piezoelectric plate with orthogonal co-ordinate system as in Fig. 7.

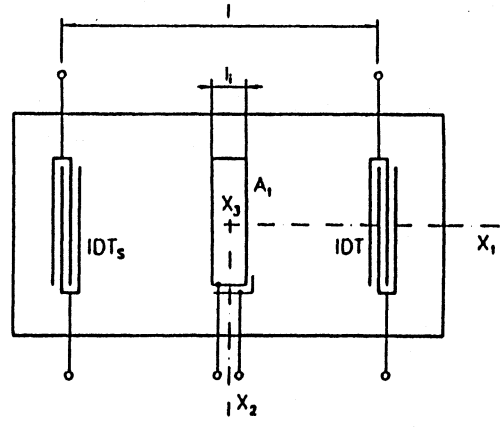


Fig.7: Control electrodes on piezoelectric plate

Two interdigital transducers IDTs and TDT are on its surface, in the x_1 x_2 - plane. An important part of the arrangement is that there is a pair of electrodes A_i, A_i' deposited both on top and the bottom plate surfaces between the interdigital transducers. These electrodes create an electric field in the plate volume. The field results in a change of the elastic module in the plate between electrodes A_i, A_i' from the value c_{LCMD} to c^*_{LCMD}. Also due to the piezoelectric effect, the length l_i is changed by Δl_i. The relative length change can be given by the relation

$$\frac{\Delta l_i}{l_i} = d^*_{311} E_3 \qquad (19)$$

where d^*_{CAB} are the components of tensor of piezoelectric coefficients. As a consequence of the SAW velocity change from the value v to v' and the relative length change Δl_i, the time τ for SAW transmission between transmitting and receiving transducers is changed by

$$\tau' = \tau + \Delta\tau_i , \qquad (20)$$

where $\tau = 1/v$. The ratio of velocities v and v' is a complicated function of linear and non-linear elastic module c^{ξ}_{LMCD}, c^{*}_{LMCD} and piezoelectric modules e_{NLM} and e^{*}_{NLM}. Let us assume the simplified equation

$$\frac{v}{v'} = \sqrt{\frac{c^{\xi}_{LMCD}}{c^{*}_{LMCD}}} , \qquad (21)$$

After substitution and arrangement, it is possible to obtain the dependence of the transmission time on electric field

$$\frac{\Delta\tau_i}{\tau_i} = d_{CAB}\left(1 - \frac{1}{2}\frac{R_{CDAB}}{d_{CAB}}\phi_{,C} - \frac{1}{2}\frac{e_{NLMCD}}{c^{\xi}_{LMCD}}\right) . \qquad (22)$$

Using the DC or slowly varying voltage u = ϕ, it makes the possibility to change the SAW time transmission continuously. In principle, an A/D converter of linear dependence of time delay on control voltage with large stability can be realised.

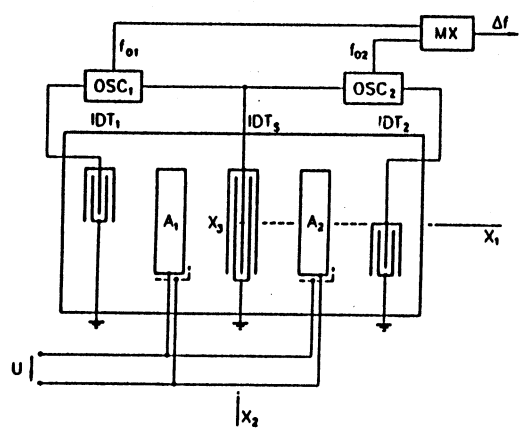


Fig. 8 An SAW sensor with two delay lines and two mistuned oscillators

5.TEMPERATURE DEPENDENCE OF RESONANT FREQUENCY

The behaviour of resonators submitted to temperature variations described by means of thermal expansion coefficients and temperature coefficients of the elastic constants. The non-linear elastic constants combined with the fundamental temperature derivatives coefficients are necessary for describing nonuniformly heated resonators. In this case, the temperature gradients induce thermal stresses and strains which, in addition to thermal expansion and temperature dependence of the fundamental elastic constants, contribute to the frequency shifts by non-linear elastic effects(Ref.3):

$$\frac{f - f_0}{f_0} = Tf^{(1)}(T - T_0) + Tf^{(2)}(T - T_0)^2 + Tf^{(3)}(T - T_0)^3 + \tilde{a}\frac{dT}{dt} \qquad (23)$$

$Tf^{(1)}, Tf^{(2)}, Tf^{(3)}$ represent the 1st, 2nd and 3rd order static temperature coefficients of frequency, and $\tilde{a}$ the dynamic temperature coefficient. It follows from a static case from the Eqn. (23) that the linear dependence, which is useful for a temperature sensor, can be achieved if the 2nd and 3rd order temperature coefficients and $\tilde{a}$ have been zero values. Such condition is valid for double rotated cut YXbl 8°26′/13° noted as LC cut, of the temperature coefficient $Tf^{(1)}$=4,79 x 10^{-5} K^{-1} and typically frequency change fo 1000 Hz K^{-1}.

The progress in the analysis and in the design of the strip resonators will enable a large reduction of the resonator dimensions in the field of the frequency - temperature sensitive quartz resonators. The reduction of the strip resonator in dimensions is accompanied by the elastic coupling of the thickness-shear vibration with other modes. Considering these effects, we investigate the influence of the thickness - length flexure, width-shear and width-length flexure modes of vibration. The rotated Y-cut and LC-cut quartz strip resonators are selected as a suitable temperature - sensitive sensors. The aim of the theoretical consideration is to optimise the dimensions and the orientations of the strip to determine the parameters leading to the minimum impact of the coupled modes, high temperature coefficient and the smooth curve expressing the frequency-temperature characteristic in the given temperature range.

The linear elastic equations for small vibrations superimposed on the thermal induced deformation by steady state and uniform temperature changes were used by the derivation of the resonant frequency -temperature characteristic of strip resonators (Ref.6). The thermally biased homogenous strain was included in the constitutive equations by means of the

thermal expansions coefficients and thermally dependent elastic stiffness. The incremental one dimensional equations of motion are solved for the modes of vibration in the strip width. The description of the vibrations of the strip was done by including the fundamental thickness-shear mode vibration and expanding the mechanical displacement and electric potential in series of trigonometric functions of the width co-ordinate. The incremental equations of motion were solved by neglecting the piezoelectric effect for the traction-free ends of the strip. Considering the equations for the displacement, many procedures provide the linear homogenous equation in the amplitude A_{mr}

$$(g_{mn})(\gamma_{mr} \cdot A_{4r}) = 0\,, \quad \text{m,n} = 1,2,3,4 \qquad (24)$$

where g_{mn} is γ_{mr}, is an amplitude ratio, r = 1,2,3,4. The nontrivial solution of Eqn. (24) is obtained, when the determinant of the coefficient g_{mn} vanishes, i.e. when

$$|g_{mn}| = 0 \qquad (25)$$

The roots of the transcendental Eqn.(25) enable to determine the normalised frequency Ω vs dimensions ratio of the strip resonator read

$$\Omega = \omega\left(\frac{\pi}{2b}\sqrt{\frac{c_{1212}}{\rho}}\right)^{-1} \qquad (26)$$

The frequency - temperature coefficients $Tf^{(1)}, Tf^{(2)}, Tf^{(3)}$ can be found from the obtained frequency - temperature dependence.

The experimental investigation showed, that for frequency f_0 = 4.197 MHz the coefficients are

$$Tf^{(1)} = 58{,}914 \cdot 10^{-6} K^{-1},$$
$$Tf^{(2)} = 65{,}032 \cdot 10^{-9} K^{-2},$$
$$Tf^{(3)} = 87{,}67 \cdot 10^{-12} K^{-3}.$$

The optimal size of the frequency - temperature sensor in the form of a small quartz strip resonator was determined for one orientation of the strip and the width/thickness ratio c/b = 3.1 . For the orientation $YXl_{5.5^\circ}$ the optimal l/b ratioes (length/thickness) are 15,9, 19,5 and 20.9. $YXl_{5.3^\circ}$ -cut strip have the largest first order temperature coefficient $Tf^{(1)}$ and k_{26}. The coefficients of the 2nd and 3rd order are $Tf^{(2)} = 82{,}18 \cdot 10^{-6} K^{-2}$ and $Tf^{(3)} = -126{,}00 \cdot 10^{-12} K^{-3}$.

6.CONCLUSION

The phenomena of the non-linear behaviour of the quartz resonators is topical because the influences of the non-linearities have a large significance for the applications of resonators both in the oscillators and in the sensors.
This behaviour is described rather by effective constants with 2nd, 3rd and 4th order fundamental constants . The theoretical and experimental evaluation of some fundamental constants of 3rd and 4th order was realised (after more years) mainly for quartz.

ACKNOWLEDGEMENT

A part of this work was supported by a Grant (Code VS96006) of Ministry of Education of Czech Republic.

REFERENCES

1. J.Zelenka: Piezoelectric Resonators and their Applications. Academia Prague, Elsevier Amsterdam 1986
2. H.F.Tiersten: Non-linear electro-elastic equations cubic in the small field variables. J.Acous.Soc.Am.57,3, (1975),pp.667-681.
3. J.J.Gagnepain: Non-linear constants and their significance. Proc. 41st AFCS,1987,pp. 266-276.
4. J.Nosek: Highly precise and stable sensors with surface acoustic wave. Proc. 15th Int.Congress on Acoustics ICA 95, Trondheim,Norway,26-30 June 1995,Vol.I,pp.397-400.
5. J.Nosek: Intermodulation products of quartz resonators and their measurement. Proc. 9th European Frequency and Time Forum EFTF 95, Besancon, France, 8-10 March 1995,pp. 447-449.
6. J.Nosek and J.Zelenka: Quartz strip resonator as a temperature sensor for mechatronics. Proc. 3rd Int. Conference MECHATRONICS 96 and M2VIP 96, Guimaraes, Portugal,18-20 September 1996, pp.233-237.
7. K.Hruska: On the linear polarizing effect with α quartz AT plates. IEEE Trans.Son.Ultrason. SU-28,p.108,1981.

INFLUENCE OF METAL THICKNESS ON THE TEMPERATURE OF SURFACE ACOUSTIC WAVE USING A PERTURBATION METHOD

S. BALLANDRAS, E. HENRY, E. BIGLER

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32 avenue de l'Observatoire - 25044 Besançon Cedex - France

ABSTRACT

An accurate adjustment of the central frequency of Surface Acoustic Wave (SAW) devices requires a good knowledge of their sensitivity to technological parameters such as metal thickness of interdigital transducers and reflectors. This parameter is known to strongly influence the turnover temperature of (ST,X) cut of quartz SAW devices widely used in resonator and filter applications. A model was implemented to calculate this phenomenon by Minowa. However, this approach did not provide an analytical expression of the first order Temperature Coefficient of Frequency (TCF) used to properly predict the turnover temperature change versus metal thickness. Based on a perturbation method developed by Tiersten and Sinha, a new analytical approach is proposed to predict the influence of metal thickness on the first order temperature coefficient. First, boundary conditions relative to the metal layer are analytically derived and injected in the variational perturbation equations. This allows to compute a correction factor to the TCF as an additional term proportional to the metal thickness. The turnover temperature is then calculated by combining these first order terms with the second order TCF obtained by the classical thermal variation of the elastic constants of the substrate. The results are finally compared with theoretical and experimental data of Minowa and a very good agreement between the different approaches is found. Moreover, the proposed analysis gives access to an analytical expression of the correction factor of the TCF well suited for systematic investigations of new quartz cuts exhibiting a frequency temperature dependence with a low sensitivity to metal thickness.

1. INTRODUCTION

Surface Acoustic Wave (SAW) devices are generally built on crystal substrates allowing the compensation of temperature effects on the elastic propagation. In the case of Rayleigh wave devices, the (ST,X) quartz cut [1] is considered as an industry standard because of its low sensitivity to quasi static thermal effects with a confortable tolerance on the propagation direction. A strong effort was performed to describe accurately the frequency-temperature characteristics of such devices using for example perturbation methods [2] or a correction of the classical constant variation approach [3]. However, most of these theoretical models address the academic problem of a Rayleigh wave propagating on a load-free surface. Actually, the presence of metallic transducers and reflectors on the propagation surface induces light modifications of the thermoelastic properties of the wave. These changes may be neglected for uncompensated substrates as $LiNbO_3$ or $LiTaO_3$, but not for devices exhibiting a first order temperature coefficient of delay (TCD) close to zero like Rayleigh wave on (ST,X) quartz plates. Only little work has been devoted to take into account the influence of the metallization on the thermal stability of SAW. Minowa has proposed a calculation method based on the thermal variations of effective elastic constants [4]. The SAW is considered to propagate under a uniform metal film and both the metal and substrate constants are varied. However, it should emphasized that approaches using non rotationally invariant coefficients were found inadequate for the accurate prediction of temperature compensated crystal orientations [5]. As a consequence, Minowa's model is assumed to suffer from the same limitations.

To avoid this problem, it is recommended to use a perturbation method according to Sinha and Tiersten work [2]. This approach is based on the combination of variational equations describing small fields superposed on a bias [6], and requires the so-called first-order temperature derivatives of fundamental elastic constants [7]. Since these derivatives are calculated using the 3rd order elastic constants and a set of representative frequency-temperature measurement for a given crystal, they were only available for quartz [7] for which these different data were measured accurately. However, a simple procedure has been proposed to compute these derivatives for isotropic or cubic material using their 3rd elastic constants and their effective thermoelastic coefficients [8]. Consequently, Sinha-Tiersten's perturbation method can be adapted to model SAW device with arbitrary metallization on the propagation surface. The first section of the paper is devoted to the perturbation equation description in the case of a semi-infinite substrate covered by a thin metal layer. The following section to establish boundary conditions due to this layer at the propagation surface which are necessary to compute the perturbation equations. Finally, results provided by the proposed analysis are compared to Minowa's theoretical and experimental data [4], yielding discussion concerning improvement of the models.

2. PERTURBATION ANALYSIS

2.1 Combination of variational equations

Figure 1 shows the axis-definition adopted for the present work. The SAW is assumed to propagate along a_1 axis and to decrease exponentially along a_2 axis (material coordinates). According to [2, 6], variational equations are derived to describe the wave propagation in both unperturbed and perturbed states as follows :

$$-\rho_o \omega_o^2 \iiint_v u_i \, u_i^{o*} \, dv =$$

$$\iint_s u_i \, C_{ijk\ell} \frac{\partial u_k^{o*}}{\partial a_\ell} \nu_j \, dS - \iiint_v \frac{\partial u_i}{\partial a_j} C_{ijk\ell} \frac{\partial u_k^{o*}}{\partial a_\ell} dV \qquad (1)$$

$$-\rho_o \omega^2 \iiint_v u_i^o \, u_i^* \, dv =$$

$$\iint_s u_i^o \, \overline{A_{ijk\ell}} \frac{\partial u_k^*}{\partial a_\ell} \nu_j \, dS - \iiint_v \frac{\partial u_i^o}{\partial a_j} \overline{A_{ijk\ell}} \frac{\partial u_k}{\partial a_\ell} dV$$

where u_i represents the mechanical displacements, $C_{ijk\ell}$ are the fundamental elastic constants, $\overline{A_{ijk\ell}}$ the perturbed elastic constants expressed in the natural state, ρ_0 the unperturbed mass density and ω the angular frequency. The upperscript ^{o} denotes an unperturbed solution and * a complex conjugation. ν_j is the normal vector pointing

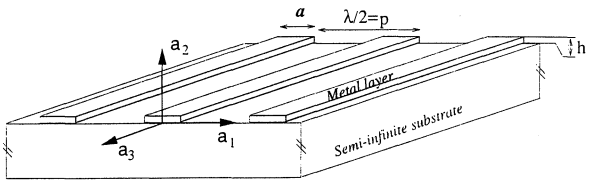


Fig. 1 : *Axis definition, case of a semi-infinite crystal substrate covered by a thin isotropic or cubic layer*

outward the substrate (see Fig. 1) and consequently written as ($\nu_1 = 0, \nu_2 = 1, \nu_3 = 0$). By considering proper assumptions [2], Eq. (1) can be combined to give access to the angular frequency variations as follows

$$\frac{\omega-\omega_o}{\omega_o} = \frac{1}{N}\iiint_v \frac{\partial u_i^{o*}}{\partial a_j}\,\overline{H_{ijk\ell}}\,\frac{\partial u_k^o}{\partial a_\ell}\,dv - \frac{1}{N}\iint_S u_i^{o*}\,\overline{H_{ijk\ell}}\,\frac{\partial u_k^o}{\partial a_\ell}\,\nu_j\,dS \tag{2}$$

with

$$N = 2\rho_o\,\omega_o^2 \iiint_v u_m^o\,u_m^{o*}\,dv$$

where $H_{ijk\ell}$ appears as a perturbation tensor which is written in the case of quasi static thermal effects applied on a plate free to expand in all directions as follows

$$\overline{H_{ijk\ell}} = C_{ijk\ell uv}\,\alpha_{uv} + C_{pjk\ell}\,\alpha_{pi} + C_{ijq\ell}\,\alpha_{qk} + \frac{dC_{ijk\ell}}{d\theta}\,(\theta-\theta_o) \tag{3}$$

where α_{uv} is the thermal expansion tensor, θ the temperature and $\theta_0 = 25°C$. Note that Eq. (3) is obtained assuming a fixed center of the plate and no rigid rotations around this point.

The surface integral of Eq. (2) results from the combination of surface integrals of Eq. (1). In the case of a surface free of mechanical loads, these integrals vanishes because of the conservation of the boundary conditions written in the natural state as follows

$$C_{ijk\ell}\,\frac{\partial u_k^o}{\partial a_\ell}\cdot\nu_j\,\Big|_{a_2=0} = 0\,;\quad \overline{A_{ijk\ell}}\,\frac{\partial u_k}{\partial a_\ell}\cdot\nu_j\,\Big|_{a_2=0} = 0 \tag{4}$$

However, in the presence of a metal layer loading the propagation surface, Eq. (4) are no longer valid and are written as follows

$$C_{ijk\ell}\,\frac{\partial u_k^o}{\partial a_\ell}\cdot\nu_j\,\Big|_{a_2=0} = T_{ij}^{'o}\,\nu_j\,\Big|_{a_2=0}$$
$$\overline{A_{ijk\ell}}\,\frac{\partial u_k}{\partial a_\ell}\cdot\nu_j\,\Big|_{a_2=0} = \overline{T'_{ij}}\,\nu_j\,\Big|_{a_2=0} \tag{5}$$

where $T^{'o}_{ij}$ and $\overline{T'_{ij}}$ represent respectively the stresses in the metal layer in the unperturbed and perturbed states.

As a consequence, the surface integral of Eq. (2) can be written as follows

$$\frac{1}{N}\iint_S u_i^{o*}\,\overline{H_{ijk\ell}}\,\frac{\partial u_k^o}{\partial a_\ell}\,\nu_j\,dS = \frac{1}{N}\iint_S u_i^{o*}\left(\overline{T'_{ij}} - T_{ij}^{'o}\right)\cdot\nu_j\,dS \tag{6}$$

It is shown in the next section how the expression of the stresses $T^{'o}_{ij}$ and $\overline{T'_{ij}}$ are derived from thermoelastic constants of the metal using proper assumptions.

2.2 Surface boundary conditions

Figure 1 shows the propagation surface, mechanically loaded by a metal layer. The profile of this layer is assumed to be described by a derivable function $f(a_1)$ as follows

$$a_2 = h\,.\,f(a_1) \tag{7}$$

where h is the maximum height of the layer. A stress-free condition similar to Eq. (4) is applied at the top surface of the metal layer. Assuming the normal vector ν_j pointing outwards this surface given by

$$\nu'_1 = -h\,\frac{\partial f(a_1)}{\partial a_1}\quad;\quad \nu'_2 = 1\quad;\quad \nu'_3 = 0 \tag{8}$$

and assuming h much smaller than the acoustic wavelength λ_{ac} allows to develop the above mentioned stress conditions as a Taylor-McLaurin series limited to the first order in h

$$T_{i2}^{'o}(a_1,0) + h f(a_1)\,\frac{\partial T_{i2}^{'o}}{\partial a_2}(a_1,0) - h\,\frac{\partial f(a_1)}{\partial a_1}\,T_{i1}^{'o}(a_1,0) = 0 \tag{9}$$

The continuity of the mechanical displacements is also assumed at the interface ($a_2 = 0$). This condition allows to determine the dependence of u'_i along a_1. Since the dependence of the mechanical displacements in the substrate are described by the function $e^{-j(\omega_0/V_R)a_1}$ (V_R is the Rayleigh wave phase velocity), the following relations are considered

$$\frac{\partial u_i^{'o}}{\partial a_1} = -j\,\frac{\omega_o}{V_R}\,u_i^{'o}\quad,\quad \frac{\partial T_i^{'o}}{\partial a_1} = -j\,\frac{\omega_o}{V_R}\,T_{ij}^{'o} \tag{10}$$

Finally, the continuity of normal stresses at the interface ($a_2 = 0$) are written to complete Eq. (9)

$$T_{i2}^{o}(a_1,0) = T_{i2}^{'o}(a_1,0) \tag{11}$$

It is now necessary to express the derivatives of $u_i^{'o}$ along a_2 to solve boundary conditions (9) and (11). It must be underlined that all the derivatives along a_3 vanish because the mechanical displacement field does not depend on this variable. According to Eq. (10), the following stress components are obtained

$$T_{11}^{'o} = -j\,\frac{\omega_o}{V_R}\,C'_{11}\,u_1^{'o} + C'_{12}\,\frac{\partial u_2^{'o}}{\partial a_2}\;;\quad T_{32}^{'o} = C'_{66}\,\frac{\partial u_3^{'o}}{\partial a_2}$$
$$T_{22}^{'o} = -j\,\frac{\omega_o}{V_R}\,C'_{12}\,u_1^{'o} + C'_{11}\,\frac{\partial u_2^{'o}}{\partial a_2}\;;\quad T_{13}^{'o} = -j\,\frac{\omega_o}{V_R}\,C'_{66}\,u_3^{'o} \tag{12}$$
$$T_{12}^{'o} = C'_{66}\left(-j\,\frac{\omega_o}{V_R}\,u_2^{'o} + \frac{\partial u_1^{'o}}{\partial a_2}\right)$$

Note that Eqs. (12) are valid for both cases of isotropic and cubic materials. These equations are now used to develop the propagation equations in the metal layer. These equations are then used to determine the derivatives of the stresses along a_2 which appear in Eq. (9). This yields the following expressions

$$\frac{\partial T'^{o}_{12}}{\partial a_2} = C'_{11}\left[\left(\frac{\omega_o}{V_R}\right)^2 - \left(\frac{\omega_o}{V'_L}\right)^2\right] u'^{o}_1 + j\frac{\omega_o}{V_R} C'_{12}\frac{\partial u'^{o}_2}{\partial a_2}$$

$$\frac{\partial T'^{o}_{22}}{\partial a_2} = C'_{66}\left[\left(\frac{\omega_o}{V_R}\right)^2 - \left(\frac{\omega_o}{V'_T}\right)^2\right] u'^{o}_2 + j\frac{\omega_o}{V_R} C'_{66}\frac{\partial u'^{o}_1}{\partial a_2} \qquad (13)$$

$$\frac{\partial T'^{o}_{32}}{\partial a_2} = C'_{66}\left[\left(\frac{\omega_o}{V_R}\right)^2 - \left(\frac{\omega_o}{V'_T}\right)^2\right] u'^{o}_3$$

with

$$V'_L = \sqrt{\frac{C'_{11}}{\rho'_o}} \quad \text{and} \quad V'_T = \sqrt{\frac{C'_{66}}{\rho'_o}}$$

Equations (12) and (13) are now injected into the boundary conditions (9). This provides the following result

$$C'_{66}\frac{\partial u'^{o}_1}{\partial a_2} + h\left(j\frac{\omega_o}{V_R} f(a_1) - \frac{\partial f(a_1)}{\partial a_1}\right) C'_{12}\frac{\partial u'^{o}_2}{\partial a_2}$$

$$= j\frac{\omega_o}{V_R} C'_{66}\, u'^{o}_2 - h \quad f(a_1)$$

$$\left[\left(\frac{\omega_o}{V_R}\right)^2 - \left(\frac{\omega_o}{V'_L}\right)^2 + j\frac{\omega_o}{V_R}\frac{\partial f(a_1)}{\partial a_1}\right] C'_{11}\, u'^{o}_1$$

$$C'_{11}\frac{\partial u'^{o}_2}{\partial a_2} + h\left(j\frac{\omega_o}{V_R} f(a_1) - \frac{\partial f(a_1)}{\partial a_1}\right) C'_{66}\frac{\partial u'^{o}_1}{\partial a_2} \qquad (14)$$

$$= j\frac{\omega_o}{V_R} C'_{12}\, u'^{o}_1 - h \quad f(a_1)$$

$$\left[\left(\frac{\omega_o}{V_R}\right)^2 - \left(\frac{\omega_o}{V'_T}\right)^2 + j\frac{\omega_o}{V_R}\frac{\partial f(a_1)}{\partial a_1}\right] C'_{66}\, u'^{o}_2$$

$$T'^{o}_{32} = -h\left\{\left[\left(\frac{\omega_o}{V_R}\right)^2 - \left(\frac{\omega_o}{V'_T}\right)^2\right] f(a_1) + j\frac{\omega_o}{V_R}\frac{\partial f(a_1)}{\partial a_1}\right\} C'_{66}\, u'^{o}_3$$

where all the stresses, displacements and displacement gradients are expressed in $a_2 = 0$. It can be remarked that since T'^{o}_{32} only depends on u'^{o}_3, it can be directly used to establish Eq. (11) for $i = 3$. Equation (14) also shows that the displacement gradients $\partial u_1'^{o}/\partial a_2$ and $\partial u_2'^{o}/\partial a_2$ are coupled together and can be expressed as combinations of displacements $u_1'^{o}$ and $u_2'^{o}$. This relation is written under the following simplified form

$$\begin{Bmatrix} \partial u'^{o}_1 / \partial a_2 \\ \partial u'^{o}_2 / \partial a_2 \end{Bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{Bmatrix} u'^{o}_1 \\ u'^{o}_2 \end{Bmatrix} \text{ in } a_2 = 0 \qquad (15)$$

with

$$A = \frac{-h}{C'_{11}C'_{66}}\left(C'^{2}_{11}\left[\left(\frac{\omega_o}{V_R}\right)^2 - \left(\frac{\omega_o}{V'_T}\right)^2\right] + C'^{2}_{12}\left(\frac{\omega_o}{V_R}\right)^2\right)$$

$$\times f(a_1) + j\frac{\omega_o}{V_R}\frac{\partial f(a_1)}{\partial a_1}\left(C'^{2}_{11} + C'^{2}_{12}\right)$$

$$B = j\frac{\omega_o}{V_R} \; ; \; C = j\frac{\omega_o}{V_R}\frac{C'_{12}}{C'_{11}} \; ; \; D = h\left(\frac{\omega_o}{V'_T}\right)^2 \frac{C'_{66}}{C'_{11}}$$

Using the matricial relation (15) together with stress distribution (12) provides the expressions of stresses T'^{o}_{12} and T'^{o}_{22}.

Finally, expressing the continuity of the displacements in $a_2 = 0$ and considering Eq. (11) provide the definitive form of the boundary conditions

$$T'^{o}_{12}(a_1, 0) = h\left[\left(\rho'_o\,\omega_o^2 - \left(\frac{C'^{2}_{11} + C'^{2}_{12}}{C'_{11}}\right)\left(\frac{\omega_o}{V_R}\right)^2\right) f(a_1)\right.$$

$$\left. - j\frac{\omega_o}{V_R}\frac{\partial f(a_1)}{\partial a_1}\left(\frac{C'^{2}_{11} + C'^{2}_{12}}{C'_{11}}\right)\right] u^o_1(a_1, 0)$$

$$T'^{o}_{22}(a_1, 0) = h\;\rho'_o\,\omega_o^2\, u^o_2(a_1, 0) \qquad (16)$$

$$T'^{o}_{32}(a_1, 0) = h\left[\left(\rho'_o\,\omega_o^2 - C'_{66}\left(\frac{\omega_o}{V_R}\right)^2\right) f(a_1)\right.$$

$$\left. - j\frac{\omega_o}{V_R}\, C'_{66}\frac{\partial f(a_1)}{\partial a_1}\right] u^o_3(a_1, 0)$$

It is important to remark that if the function $f(a_1)$ describes a rectangular profile (obtained by combination of Heaviside functions), then Eq. (16) correspond to boundary conditions established by Datta and Hunsinger [9]. Also, if the profile $hf(a_1)$ is assumed rectangular (see Fig. 1), considering the condition $h \ll \lambda_R$ respected, Eq. (16) provides the so-called Tiersten boundary conditions [10]. These equations can be used to compute the frequency shift of the Rayleigh wave only due to mass loading using the following perturbation formula

$$\frac{\Delta\omega}{\omega_o} = \frac{\iint_S T^o_{i2}\, u^o_i\, u^{o*}_i\, dS}{2\rho_o\,\omega_o^2 \iiint_V u^o_\ell\, u^{o*}_\ell\, dv} \qquad (17)$$

It is now necessary to establish the expressions of $T'^{o}_{i2}(a_1,0)$ to calculate the surface integral of Eq. (6). In that case, a particular attention should be payed to avoid mixing statical and dynamical terms. The latter are always considered unperturbed, and only the material fundamental constants are modified.

$$\overline{T'_{12}}(a_1,0) = h\left[\frac{\overline{A'_{66}}}{C'_{66}}\left(\rho'_o\,\omega_o^2 - C'_o\left(\frac{\omega_o}{V_R}\right)^2\right) f(a_1)\right.$$

$$\left. -j\frac{\omega_o}{V_R}\frac{\partial f(a_1)}{\partial a_1}\left(\frac{\overline{A'_{11}}\,C'_{11}+\overline{A'_{12}}\,C'_{12}}{C'_{11}}\right)\right] u_1^o(a_1,0)$$

$$\overline{T'_{22}}(a_1,0) = h\left[\frac{\overline{A'_{11}}}{C'_{11}}\rho'_o\,\omega_o^2\right. \tag{18}$$

$$\left. -\left(\frac{\overline{A'_{12}}\,C'_{11}+\overline{A'_{11}}\,C'_{12}}{C'_{11}}\right)\left(\frac{\omega_o}{V_R}\right)^2\right] u_2^o(a_1,0)$$

$$\overline{T'_{32}}(a_1,0) = h\left[\frac{\overline{A'_{66}}}{C'_{66}}\left(\rho'_o\,\omega_o^2 - C'_{66}\left(\frac{\omega_o}{V_R}\right)^2\right) f(a_1)\right.$$

$$\left. -j\frac{\omega_o}{V_R}\,\overline{A'_{66}}\,\frac{\partial f(a_1)}{\partial a_1}\right] u_3^o(a_1,0)$$

In the case of a metal layer with rectangular profile, the right-hand side of Eq. (6) can be simply written as follows :

$$\frac{1}{N}\int\!\!\int_S u_i^{o*}\,(\overline{T'_{i2}} - T'^{o}_{12})\,dS) = -h\frac{a}{p}\,{}^{\theta}\beta_i\;V_i \tag{19}$$

where

$${}^{\theta}\beta_1 = \frac{\overline{H'_{66}}}{C'_{66}}\left[\left(\frac{\omega_o}{V_R}\right)^2 C'_o - \rho'_o\,\omega_o^2\right]$$

$${}^{\theta}\beta_2 = \left(\frac{\omega_o}{V_R}\right)^2\left(\frac{\overline{H'_{12}}\,C'_{11}-\overline{H'_{11}}\,C'_{12}}{C'_{11}}\right) - \rho'_o\,\omega_o^2\frac{\overline{H'_{11}}}{C'_{11}}$$

$${}^{\theta}\beta_3 = \frac{\overline{H'_{66}}}{C'_{66}}\left[\left(\frac{\omega_o}{V_R}\right)^2 C'_{66} - \rho'_o\,\omega_o^2\right]$$

According to the definitions of [11], V_i combines unperturbed wave parameters as follows :

$$V_i = \frac{\sum_{r,s=1}^{4} A_r^*\,A_s\;{}^o u_i^{o(r)*}\;{}^o u_j^{o(s)}}{2\rho_o\,\omega_o\,V_R\sum_{r,s=1}^{4}\dfrac{A_r^*\,A_s\;{}^o u_m^{o(r)*}\;{}^o u_m^{o(s)}}{j(n_2^{(r)*} - n_2^{(s)})}} \tag{20}$$

The final form of the perturbation equation can be established as follows

$$\frac{\Delta\omega}{\omega_o} = ({}^{\theta}\alpha + h\frac{a}{p}\,{}^{\theta}\beta)(\theta-\theta_o) \tag{21}$$

In Eq. (21), the coefficient ${}^{\theta}\alpha$ is the same as the one provided by the usual load-free calculation [11]. The influence of the metal layer on the frequency stability of the SAW is given by a linear relation along h due to the truncature of the Taylor-McLaurin series in Eq. (9). Note that a higher order truncature would provide a more accurate description of the metal layer contribution to the thermoelastic properties of the considered devices. However, most of them are built using very thin metal layer which largely respect the condition $h \ll \lambda_{ac}$.

A stronger assumption was made in establishing Eq. (21) by neglecting the stresses generated at the metal-substrate interface by the differential thermal expansion of both materials. These stresses should be calculated using finite element analysis because of the substrate anisotropy which prevents any simple analytical model. Nevertheless, calculations have been performed to check the importance of this assumption in regard with Minowa's experimental results. One can remark that Minowa's calculations also neglect these stresses and provide an overestimated influence of the metal layer on the turnover temperature variations of SAW on ST quartz [4]. However, the proposed perturbation method is well adapted to take these additive terms into account by using an adequate expression of the perturbation tensor $\overline{H_{ijk\ell}}$ and a discretized form of Eq. (2) as proposed in [12].

3. COMPARISON WITH MINOWA'S THEORETICAL AND EXPERIMENTAL DATA

As mentioned in introduction, Minowa's theoretical calculation is based on a model of dispersive Rayleigh-like wave propagating under a uniform metal layer deposited on a piezoelectric crystal. This model is used to compute the SAW frequency shift due to temperature perturbation by the variations of effective thermoelastic constants of both crystal and metal.

These approaches provide first the SAW sensitivity to mass loading, but also the first and second order TCF. These latter coefficients are used to evaluate the turnover temperature, and finally its variations along the metal layer thickness.

Experiments of Ref. [4] have been performed on three-transducer-type SAW devices [13] with a $\lambda_{ac} = 20.56$ µm and a finger width of 5.14 µm (finger to gap ratio equals 1), built on quartz cuts close to (ST,X) orientation (θ respectively equal to 40° and 42.75°) with a propagation direction varying from 0° to 6°. Since Minowa's model does not allow to consider periodic gratings as built practically, an equivalent thickness h equal to one half of the experimental metal srip height was considered for theoretical calculations of Ref. [4] since the actual profile can be considered in the present approach, perturbation computations have been performed for both cases of a half thickness homogeneous layer and of a rectangular metal strip (calculations performed over one mechanical period $\lambda_{ac}/2$). Figures 2 and 3 show a comparison between theoretical and experimental data of Ref. [4] concerning the SAW on ST quartz sensitivity to mass loading and the theoretical calculation corresponding to Eq. (19), using Slobodnik's constants for quartz [14] and Landolt-Bornstein data for aluminum [15]. The proposed approach provides results closer to the experimental data than Minowa's calculations ones. This may be due to the fundamental constants used in Ref. [4] which do not comply accurately with those found in [14] and [15]. Moreover, the description of the analytical model of propagation in [4] does not allow to check Minowa's calculation procedure. It should be precised that the perturbation calculation corresponding to Eq. (19) is low sensitive to the variations of the elastic constants of the metal, which can considered either isotropic or cubic (change of C'_{66}). The precision of the prediction is better than $1.5.10^{-4}$, instead of $5.1.10^{-4}$ for Minowa's theory.

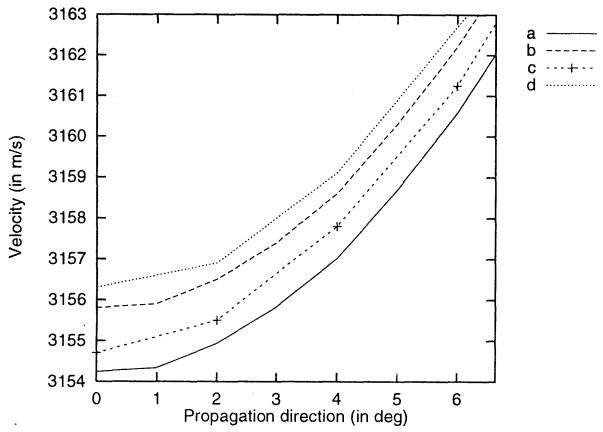


Fig. 2 : *Comparison of Rayleigh wave velocity versus propagation direction between Minowa's theoretical and experimental data [4] and the present results case of an ST quartz cut loaded by a 1 000 Å Aluminum layer (C'_{66} = 25 GPa)*
a) loaded substrate (this work)
b) unloaded substrate (this work)
c) experiments [4]
d) theoretical results of [4]

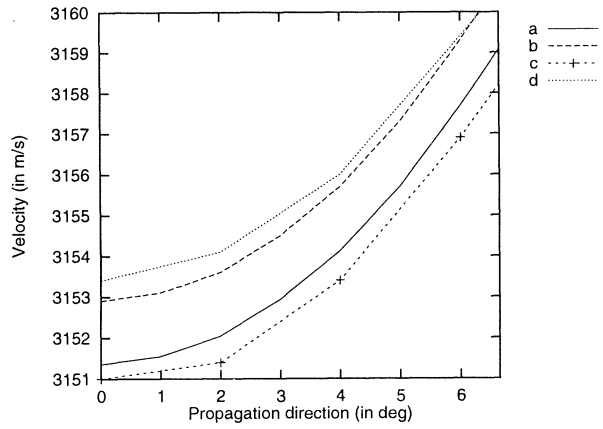


Fig. 3 : *Comparison of Rayleigh wave velocity versus propagation direction between Minowa's theoretical and experimental data [4] and the present results case of an (Φ = 0°, θ = 40°) quartz cut loaded by a 1 000 Å Aluminum layer (C'_{66} = 25 GPa)*
a) loaded substrate (this work)
b) unloaded substrate (this work)
c) experiments [4]
d) theoretical results of [4]

In order to provide comparative results for the frequency-temperature analysis, the turnover temperature has been calculated according to Minowa's procedure (first and second order TCF calculated using the effective constant variation approach) but taking into account the correction term of Eq. (21) proportional to the metal thickness h. This yields the following expression of the turnover temperature

$$\theta_{turn} = \theta_o - \left(\frac{\alpha_1^{(o)} + h^{\theta}\beta}{2\,\alpha_2^{(o)}} \right) \qquad (22)$$

where $\alpha_1^{(o)}$ and $\alpha_2^{(o)}$ are respectively the first and second order TCF according to Minowa's calculation [4].

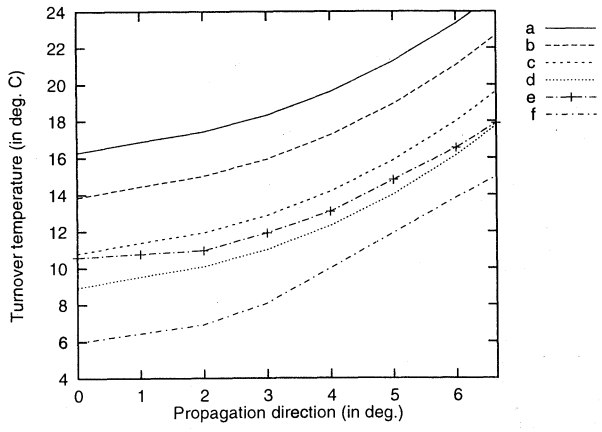


Fig. 4 : *Comparison of turnover temperature versus propagation direction between Minowa's theoretical and experimental data [4] and the present results, case of an ST quartz cut loaded by a 1000 Å Aluminum layer*
a) unloaded substrate (this work)
b) loaded substrate, C'66 = 28.3 GPa (this work)
c) loaded substrate, C'66 = 25 GPa
d) loaded substrate, C'66 = 23 GPa
e) experiments [4]
f) theoretical results of [4]

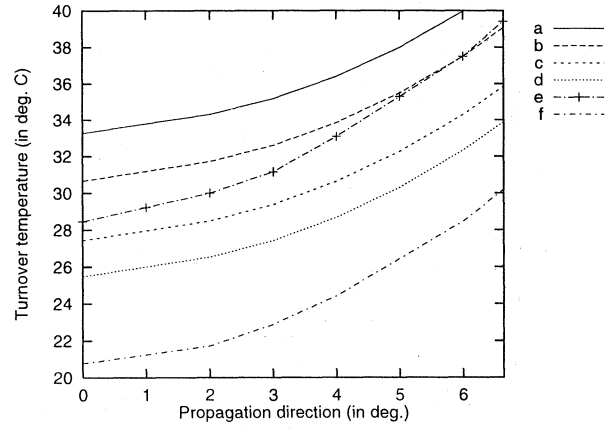


Fig. 5 : Comparison of turnover temperature versus propagation direction between Minowa's theoretical and experimental data [4] and the present results, case of an (Φ = 0°, θ = 40°) quartz cut loaded by a 1000 Å Aluminum layer
a) unloaded substrate (this work)
b) loaded substrate, C'66 = 28.3 GPa (this work)
c) loaded substrate, C'66 = 25 GPa
d) loaded substrate, C'66 = 23 GPa
e) experiments [4]
f) theoretical results of [4]

Figures 4 and 5 compare the variation of the turnover temperature versus propagation direction for Aluminum (1000 Å thick) on ST quartz resulting from the proposed theoretical analysis and Minowa's theoretical and experimental data [4]. The present calculations have been performed using Aluminum nonlinear elastic constants from Ref. [15] and first derivatives of the linear elastic constants of Aluminum calculated according to [8] (Refs. [15] and [16]). It is important to emphasize the fact that the proposed calculation is very sensitive to these data. Particularly, Figs. 5 and 6 underline the influence of the shear elastic constant C'_{66} of the metal layer. Actually, a variation of 10 % of C'_{66} induces a 3 to 5°C shift of the turnover temperature. However, using the proposed set of coefficients allows a more accurate prediction of the turnover temperature than Minowa's analysis.

4. CONCLUSION

A rigorous analysis of the influence of thin isotropic or cubic layers (metal or oxyde) on the velocity and on the thermoelastic behavior of Rayleigh waves on quartz has been developed and implemented. Comparison between theoretical and experimental data already published and the present results have been exposed, emphasizing the good accuracy of the proposed model in regard with experiments. However, the gain of precision is trongly influenced by the good knowledge of fundamental thermoelastic constants of both layer and substrate.

Although these data are well-known in the case of quartz and for most of the isotropic and cubic materials, the properties of thin layer materials may be subjected to some variation in regard with the bulk material ones. As a consequence, the proposed model provides accurate predictions only if these properties are precisely characterized. Even if a satisfying set of constants was found for Aluminum, greatly improving the theoretical analysis, it should be noted that more experimental data are necessary to confirm the present results and to extend the model to other materials.

References

[1] M.B. Schultz, B.J. Matzinger, M.G. Holland, "Temperature dependence of surface acoustic wave velocity on α quartz", J. Appl. Phys., Vol. 41, n°7, pp. 2755-2765 (1970).

[2] H.F. Tiersten and B.K. Sinha, "A perturbation analysis of the attenuation and dispersion of surface waves", J. Appl. Phys., vol. 49, n° 1, pp. 87-95 (jan. 1978).

[3] R.L. Rosenberg, "Thermal shearing effects on the temperature stability of SAW devices", IEEE Trans. Sonics Ultrasonics, vol. SU-27, n° 3, pp. 130-133 (may 1980).

[4] J. Minowa, "A method for accurately adjusting the center frequency of surface acoustic wave filters", Review of the Electrical Communication Lab.", vol. 26, n° 5-6, pp. 797-807, 1978.

[5] B.K. Sinha and H.F. Tiersten, "On the temperature dependence of the velocity of surface waves in quartz", J. Appl. Phys., vol. 51, n° 1, pp. 4659-4665 (jan. 1980).

[6] J.C. Baumhauer and H.F. Tiersten , "Nonlinear electroelastic equations for small fields superposed on a bias", J. Acoust. Soc. Am., vol. 54, n° 4, pp. 1017-1074 (oct. 1973).

[7] B. K. Sinha, H. F. Tiersten, "First temperature derivatives of the fundamental elastic constants of quartz", J. Appl. Phys., Vol. 50, n° 4, pp. 2732-2739 (1979).

[8] S. Ballandras, E. Gavignet, E. Bigler, "Temperature derivatives of the fundamental elastic constants of isotropic materials", submitted to Appl. Phys. Lett.

[9] S. Datta, B. Hunsinger,"First order reflexion coefficient of surface acoustic waves from thin strip overlays", J. Appl. Phys., vol. 50, n° 9, pp. 5661-5665 (1979).

[10] B.A. Auld, *Acoustic fields and waves in solids*, edited by R.E. Krieger Publishing Company (Malabar, Florida, 1990), vol. II, pp. 159-170.

[11] S. Ballandras, E. Bigler, "Surface acoustic wave devices with low sensitivity to mechanical and thermoelastic stresses", J. Appl. Phys., vol. 72, n° 8, pp. 3272-3281 (october 1992).

[12] S. Ballandras, E. Bigler, "Calculation of SAW sensitivity to mechanical effects using a perturbation method coupled with finite element analysis", 9th Piezoelectric Conf. (Piezo'96), Waplewo (Poland), oct. 1996.

[13] Y. Koyamada *et al.*, "Elastic surface wave filter using three transducers", Rep. of 1973 Spring Meeting Acoustical Society of Japan, 3-4-2, 1973.

[14] A.J. Slobodnik, E.D. Conway, R.T. Delmonico, "Microwave Acoustics Handbook, vol. 1A : Surface Wave Velocities", Air Force Cambridge Research Labs, 1973.

[15] Landolt-Börnstein, *Numerical Data and Functional Relationships in Science and Technology, group III : Crystal and Solid State Physics*, vol. 11, Springer-Verlag Berlin - Heidelberg - New York, 1979.

[16] Bonczar L.J., G.R. Barsch, J. Appl. Phys., vol. 46, p. 4339 (1975).

OSCILLATORS

Chairman: B. Schlueter

TRANSIENT SIMULATION IN QUARTZ CRYSTAL OSCILLATORS

R. Brendel*, N. Ratier*, L. Couteleau*, G. Marianneau*, P. Guillemot**

*Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32 avenue de l'Observatoire - 25044 Besançon Cedex - France

** Centre National d'Etudes Spatiales
18 avenue E. Belin - 31055 Toulouse Cedex - France

ABSTRACT

By using an approach based on the full nonlinear Barkhausen criterion, it is possible to describe oscillator behaviour under the form of a nonlinear characteristic polynomial whose coefficients are functions of the circuit components and of the oscillation amplitude. Solving the polynomial in the frequency domain leads to the steady state oscillations amplitude and frequency. In the time domain, the characteristic polynomial represents a nonlinear differential equation whose solution gives the oscillator signal transient. It is shown how symbolic manipulation capabilities of commercially available softwares can be used to automatically generate the coding of the oscillator characteristic polynomial from the SPICE description netlist. The numerical processing of such an equation in the time domain leads to inadmissible computer time because of the high quality factor of the oscillator circuits involved. Nevertheless, by using the slowly varying amplitude and phase method, it is possible to transform the initial nonlinear differential equation into a nonlinear first order differential equation system in the amplitude and phase variables. The solution of this system directly gives the designer the most relevant features of the oscillation ; that is, the amplitude, phase or frequency transients which can be accurately obtained within a short computer time by using classical numerical algorithms.

1. INTRODUCTION

Increasing performance of quartz crystal oscillators as well as predictability requirements when developing these devices need more and more accurate and powerful design tools. For several years, much software has been made available to help the designer. Nevertheless, these programs are either quite general, so that they are not well suited to quartz oscillator design or they are highly specialized in the design of a limited number of oscillator structures. In most cases, this software gives steady state features of the oscillation ; namely, amplitude, frequency and noise. However, because of the high quality factor involved in the quartz oscillator circuits, these programs usually fail to accurately describe the transient behaviour.

For several years, a simulation program dedicated to quartz crystal oscillators has been developed, mainly grounded on the characteristic polynomial which represents the nonlinear Barkhausen criterion. Previous papers [1, 3] have described the oscillator circuit reduction method leading to the characteristic polynomial from the circuit description.

By using commercially available symbolic calculation programs it is possible to produce the calculation code giving the characteristic polynomial coefficients. However, automatic circuit reduction should be made with care because the brute force method using symbolic analysis often results in an explosion of terms in equations. On the other hand, a too high encapsulation level may hide possible simplifications leading to an unnecessarily high polynomial degree which may be difficult to solve. It will be shown that, by properly analyzing and reducing the circuit, it is possible to safely, concisely and quickly generate the polynomial coefficients.

Once obtained, the characteristic polynomial can be solved in the frequency domain so as to obtain amplitude, frequency and other features of the steady state oscillation [2].

In the time domain, the characteristic polynomial represents a nonlinear differential equation. Even by using a numerical method, this equation is quite difficult to solve in the general case because of the high quality factor of the circuit resulting in an inadmissible computer time. Fortunately this particularity which is a drawback for the general method turns into an advantage when using asymptotic methods.

The so-called slowly varying amplitude and phase method transforms the initial nonlinear differential equation into a nonlinear first order differential equation sytem in the amplitude and phase variables. The solution of this system is no longer the oscillator signal itself but the amplitude, phase or frequency computation which can be accurately obtained within a short computer time by using classical numerical algorithms.

2. OSCILLATION CONDITION

It has been long known that a quartz crystal oscillator can be represented as shown in Fig. 1.

Fig. 1 : *Simplified block diagram of a quartz crystal oscillator*

The resonator exerts a positive reaction from the output to the input of an amplifier "A" leading to growing oscillation whose frequency is mainly determined by the resonator while the amplitude is limited by the amplifier nonlinearity. Although the simple form shown in Fig. 1 is appealing because it is easy to understand, it gives very few insights into the influence of the various circuit components on the oscillation features.

So as to go proceed in the oscillator analysis, the nonlinear oscillator behaviour has to be described under a suitable form. In the present work, the amplifier will be described under the form of a large signal admittance two port circuit [4] as shown in Fig. 2.

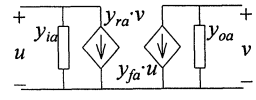


Fig. 2 : *Large signal admittance two port representation of an amplifier*

In this representation, the y-parameters are no longer constant as in small signal representation, rather they are functions of the amplitude of the input or output voltage. More precisely y_i and y_f depend on the input voltage amplitude $\bar{u}$ while y_r and y_o depend on the output voltage amplitude $\bar{v}$.

So as to simplify the automation of the circuit reduction process, it is convenient to consider that in all cases the resonator takes up one of the three basic positions with respect to the amplifier as shown in Fig. 3.

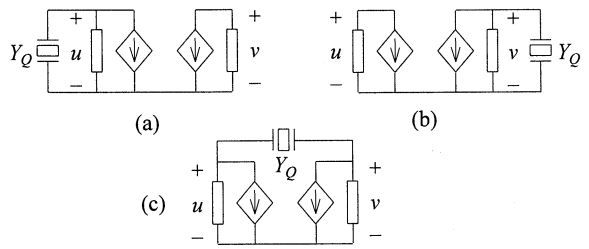


Fig. 3 : *Three basic positions of the resonator with respect to the amplifier*

By simple transformations, each of the three basic circuits reduces to the equivalent circuit shown in Fig. 4. Table I gives the expression of the equivalent admittances.

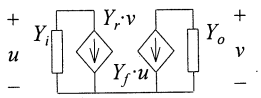


Fig. 4 : *Equivalent circuit of the oscillator*

By inspection of Fig. 4, we have

$$\begin{cases} Y_i\ u + Y_r\ v = 0 \\ Y_f\ u + Y_o\ v = 0 \end{cases} \tag{1}$$

The solution of this homogeneous system can be different from zero if and only if, its determinant is null

$$\Delta y = Y_i\ Y_o - Y_f\ Y_r = 0 \tag{2}$$

This equation represents the oscillation condition of the circuit. By using the expression of the equivalent admittances summarized in Table I, the oscillation condition of the three basic configurations can be expressed as a function of the resonator admittance Y_Q and the amplifier equivalent admittances y_{ia}, y_{ra}, y_{fa}, y_{oa}. So as to improve the simplification process when reducing the oscillator circuit, it turns out that it is more convenient to express the oscillation condition with respect to y_{ia}, y_{oa}, the determinant Δy_a and the matrix element sum y_{sa} defined by

$$\begin{aligned} \Delta y_a &= y_{ia}\ y_{oa} - y_{ra}\ y_{fa} \\ \text{and } y_{sa} &= y_{ia} + y_{ra} + y_{fa} + y_{oa} \end{aligned} \tag{3}$$

as shown in the last column of Table I.

Table I : *Equivalent admittances and oscillation condition (see Figs. 3 and 4)*

	Y_i	Y_r	Y_f	Y_o	Oscillation condition
(a)	$y_{ia} + Y_Q$	y_{ra}	y_{fa}	y_{oa}	$\Delta y_a + y_{oa} . Y_Q = 0$
(b)	y_{ia}	y_{ra}	y_{fa}	$y_{oa} + Y_Q$	$\Delta y_a + y_{ia} . Y_Q = 0$
(c)	$y_{ia} + Y_Q$	$y_{ra} - Y_Q$	$y_{fa} - Y_Q$	$y_{oa} + Y_Q$	$\Delta y_a + y_{sa} . Y_Q = 0$
			$y_{sa} = y_{ia} + y_{ra} + y_{fa} + y_{oa}$		$\Delta y_a = y_{ia} y_{oa} - y_{ra} y_{fa}$

3. CIRCUIT REDUCTION PRINCIPLE

The oscillation condition derivation presented in the previous section assumes that the amplifier can be reduced to the form represented in Fig. 2 and that the large signal admittance parameters are known nonlinear functions of the signal amplitude.

In fact, it is not convenient to measure the overall nonlinear behaviour of the amplifier parameters. It is simpler to start from the nonlinear admittance parameters of the active component (e.g. a BJT) which can be obtained either by measurement or by using an electrical simulator (e.g. SPICE) and to reduce the circuit to the required form by using a limited number of simple transformations represented in Fig. 5. At each step of the reduction process a new circuit component is introduced in the equivalent circuit and the parameters of the new equivalent circuit are expressed in terms of the former ones and the new component. Meantime, the input and output voltage of the new circuit are expressed in terms of the former ones.

Table II summarizes the expressions of the new circuit matrix elements and the expressions of the new input and output voltages as a function of the former circuit characteristics for each elementary transformation. It should be noted that the admittance matrix elements as well as the additional admittance Y represent equivalent circuit admittances. In the general case, these admittances are fractional functions of the circuit components so that the oscillation condition Eq. (2) takes the form of a polynomial in the Laplace's variable s :

$$\sum a_k\ s^k = 0 \tag{4}$$

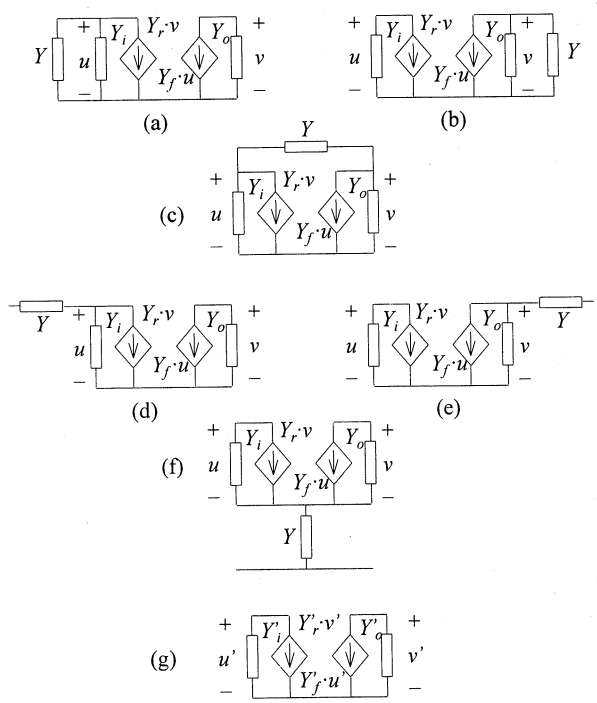


Fig. 5 : *Elementary transformations for circuit reduction*

Table II : *Elementary transformation relations (see Fig. 5)*

	Y'_i	Y'_r	Y'_f	Y'_o	Y'_s	$\Delta Y'$	P : $\begin{bmatrix} u' \\ v' \end{bmatrix} = P \begin{bmatrix} u \\ v \end{bmatrix}$
(a)	Y_i+Y	Y_r	Y_f	Y_o	Y_s+Y	$\Delta Y+Y_oY$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
(b)	Y_i	Y_r	Y_f	Y_o+Y	Y_s+Y	$\Delta Y+Y_iY$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
(c)	Y_i+Y	Y_r-Y	Y_f-Y	Y_o+Y	Y_s	$\Delta Y+Y_sY$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
(d)	$\frac{Y_iY}{Y_i+Y}$	$\frac{Y_rY}{Y_i+Y}$	$\frac{Y_fY}{Y_i+Y}$	$\frac{Y_oY+\Delta Y}{Y_i+Y}$	$\frac{Y_sY+\Delta Y}{Y_i+Y}$	$\frac{\Delta Y.Y}{Y_i+Y}$	$\begin{bmatrix} \frac{Y_i+Y}{Y} & \frac{Y_r}{Y} \\ 0 & 1 \end{bmatrix}$
(e)	$\frac{Y_iY+\Delta Y}{Y_o+Y}$	$\frac{Y_rY}{Y_o+Y}$	$\frac{Y_fY}{Y_o+Y}$	$\frac{Y_oY}{Y_o+Y}$	$\frac{Y_sY+\Delta Y}{Y_o+Y}$	$\frac{\Delta Y.Y}{Y_o+Y}$	$\begin{bmatrix} 1 & 0 \\ \frac{Y_f}{Y} & \frac{Y_o+Y}{Y} \end{bmatrix}$
(f)	$\frac{Y_iY+\Delta Y}{Y_s+Y}$	$\frac{Y_rY-\Delta Y}{Y_s+Y}$	$\frac{Y_fY-\Delta Y}{Y_s+Y}$	$\frac{Y_oY+\Delta Y}{Y_s+Y}$	$\frac{Y_sY}{Y_s+Y}$	$\frac{\Delta Y.Y}{Ys+Y}$	$\begin{bmatrix} \frac{Y_i+Y_f+Y}{Y} & \frac{Y_r+Y_o}{Y} \\ \frac{Y_i+Y_f}{Y} & \frac{Y_r+Y_o+Y}{Y} \end{bmatrix}$
	$\Delta Y = Y_i Y_o - Y_r Y_f$			$Y_s = Y_i+Y_r+Y_f+Y_o$			

4. AUTOMATIC CIRCUIT REDUCTION

Two components play a major role in the oscillator circuit : the resonator and the active component. The resonator mainly determines the oscillation frequency. It is modeled by the classical equivalent circuit shown in Fig. 6.

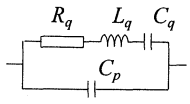


Fig. 6 : *Equivalent circuit of the resonator*

The resonator behaviour is characterized by the series resonant frequency f_q, the inductance L_q, the series resistance R_q and the parallel capacitance C_p. Because of the amplitude-frequency effect the actual resonant frequency f_o depends on the drive level [5] according to the law :

$$f_o = f_q\ (1 + a.P) \qquad (5)$$

where P is the active power in the crystal and a the isochronism defect parameter, the capacitance C_q is calculated from f_q and L_q.

As outlined in section 2, the four y-parameters of the amplifier circuit are function of the input or output voltage amplitude. They can be obtained by an independent set of simulations with larger and larger amplitude performed in the same bias conditions as in the oscillator circuit.

The symbolic calculation capability of commercially available software like MAPLE [6] has made possible the development of a topological analysis module. Starting from the oscillator circuit described under the form of a SPICE netlist, this module automatically performs the formal circuit reduction as demonstrated in the previous sections, and derives the oscillation condition. In fact, the analysis module gives the polynomial coefficients a_k in Eq. (4). They are edited in high level language (FORTRAN Code), each coefficient being formally expressed in terms of the circuit admittances including the transistor large signal y-parameters.

By processing the circuit step by step as described in section 3, it is possible to "encapsulate" the equivalent y-parameters at a given step into symbolic expressions which are then used in the subsequent steps thus avoiding explosion of terms as often met with when handling symbolic expressions [7]. Furthermore, by using the four variables Y_i, Y_o, Y_s and ΔY as defined in Table II, the topological analysis module leads to a characteristic polynomial of minimum degree. Another advantage of this procedure is that it starts from the active element y-parameter equivalent circuit whose nonlinear behaviour is known as a function of the component input and output voltage amplitude.

Thus, at each encapsulation step, the relationship between new and old equivalent circuit input and output voltages are expressed and saved as shown in Table II. Eventually, the circuit is reduced to one of the three forms shown in Fig. 3 where each admittance parameter is expressed as a function of the initial circuit admittances.

It should be remembered that the polynomial coefficients in Eq. (4) depend either on the active component input or output voltage or on the voltage across the resonator. By performing the matrix product of the successive transformations, the voltage across the resonator can be expressed as a function of the active component input and output voltages which can then be used to solve the oscillation condition Eq. (4).

5. TRANSIENT EQUATION

The steady state oscillation features can be obtained by replacing the Laplace's variable s by the harmonic variable $j\omega$. Equation (4) then splits into real and imaginary part equations. Solving these equations amounts to find the pair (amplitude, frequency) satisfying simultaneously both equations. A software based on this method has already been developed and described in the literature [2, 3].

Although the steady state oscillation amplitude and frequency and their sensitivity to the circuit component value or to environmental condition variation are of the upmost interest for the designer, it does not give him all the information on the oscillator behaviour. In particular, it should be interesting to describe the oscillation start up. To this end Eq. (4) can be used to derive the nonlinear differential equation of the oscillation signal by replacing the Laplace's variable s by the differential operator d/dt.

Although this operator can be applied to any one of the circuit voltage or current variables, it appears to be more convenient to use the active component input voltage as main variable. Equation (4) then looks like a nonlinear ordinary differential equation which can be solved by using usual numerical methods, unfortunately these algorithms generally fail to converge in a reasonable computing time because of the high quality factor of the oscillator circuit. Nevertheless it should be observed that the higher the quality factor the closer the output signal is to a sinusoid whose frequency is determined by the resonator. This remark induces to put Eq. (4) under the form :

$$\ddot{x} + \omega_q^2\, x = \lambda\, f(x, \dot{x}, \ddot{x}, \ldots \underset{\cdot}{x}, \underset{\cdot\cdot}{x}, \ldots) \qquad (6)$$

where x is the chosen variable (e.g. transistor input voltage), $\dot{x}$, $\ddot{x}$ are derivatives with respect to time : $\dot{x} = s.x = dx/dt$,
$\underset{\cdot}{x}$, $\underset{\cdot\cdot}{x}$ are primitives with respect to time $\underset{\cdot}{x} = x/s = \int x\, dt$.
λ is a so-called "small parameter" which determines the closeness of the system to a linear conservative system.

So as to change from Eq. (4) to Eq. (6) it is necessary to isolate the terms dealing with the resonator admittance. This is one of the reasons for which the resonator has been kept outside the reduction process (see Fig. 3) and the oscillation condition has been expressed under the three basic forms listed in the last column in Table I. In all cases the oscillation condition has the general form :

$$\Delta y_a + y_{ma} \cdot Y_a = 0 \qquad (y_{ma} = y_{oa} \ , \ y_{ia} \ or \ y_{sa}) \tag{7}$$

where the admittance Y_Q of the resonator equivalent circuit shown in Fig. 6 can be put under the form :

$$Y_Q = \frac{s}{L_q \cdot T(s)} \ , \quad T(s) = s^2 + \frac{\omega_q}{Q_q} s + \omega_q^2 \ ,$$
$$\omega_q^2 = \frac{1}{L_q C_q} \ , \quad Q_q = \frac{L_q \omega_q}{R_q} \tag{8}$$

Thus, the left hand side of Eq. (6) is contained in the trinomial $T(s)$ defined in Eq. (8). By putting Eq. (8) into (7) the trinomial $T(s)$ can be expressed as

$$T(s) = \frac{-s \cdot y_{ma}}{L_q \cdot \Delta y_a} = -\frac{N(s)}{D(s)} \tag{9}$$

So that the relevant differential operator takes the form :

$$T(s) \cdot D(s) + N(s) = 0 \tag{10}$$

The polynomial $D(s)$ can be expressed as

$$D(s) = \sum_{i=0}^{M} D_i \ s^i \tag{11}$$

The steady state analysis previously performed permits to obtain the steady state frequency and amplitude. These values are used to calculate the modulus of each term of the sum (11). By factorizing the monomial of maximum modulus : $D_k \ s^k$, Eq. (11) takes the form :

$$D(s) = D_k \ s^k \left[1 + \frac{1}{D_k} \sum_{\substack{j=-k \\ (j \neq k)}}^{M-k} D_{k+j} \ s^j \right] \tag{12}$$

So that the differential operator (10) takes the form :

$$T(s) = -\frac{N(s)}{D_k \ s^k} - \frac{T(s)}{D_k} \sum_{\substack{j=-k \\ (j \neq k)}}^{M-k} D_{k+j} \ s^j \tag{13}$$

Under the form (13), the differential operator leads to the integro-differential equation shown in Eq. (6). The factorization of the maximum modulus term ensures that the right hand side terms of the differential equation have a smaller modulus than the left one so that a small parameter can be factorized.

6. SLOWLY VARYING AMPLITUDE AND PHASE METHOD

This method was developed several decades ago [8-10] so as to get the amplitude and phase transients of Eq. (6) without having to solve the equation itself. Because of the smallness of the right hand side of Eq. (6), the solution of this equation should be close of the solution of the linear equation obtained when $\lambda = 0$. This assumption leads to seek a solution under the form

$$x = y(t) \cos [\omega_q \ t + \phi(t)] = y(t) \cos \psi(t) \tag{14}$$

If λ were null, $y(t)$ and $\phi(t)$ would be constant with respect to time, λ being small, $y(t)$ and $\phi(t)$ are no longer constant but slowly varying with respect to time so that their time derivatives are small. The quantity

$$\psi(t) = \omega_q \ t + \phi(t) \tag{15}$$

is called total phase of the motion whose derivative with respect to time is the instantaneous angular frequency :

$$\omega = \frac{d\Psi}{dt} = \omega_q + \frac{d\phi}{dt} = \omega_q + \dot{\phi} \tag{16}$$

In fact $\dot{\phi}$ represents the frequency perturbation with respect to the natural frequency ω_q. It can be easily shown that the variable change defined by Eq. (14) transforms the second order differential equation (6) into a first order differential system in the variables y and ϕ :

$$\begin{cases} \dot{y} = \dfrac{-\lambda}{\omega_q} f(y, \psi) \sin \psi \\ \dot{\phi} = \dfrac{-\lambda}{\omega_q \ y} f(y, \psi) \cos \psi \end{cases} \tag{17}$$

Because y and ϕ are slowly varying quantities it can be shown that, in the first order approximation in λ these quantities keep the same value for one period of the signal so that Eqs. (17) are approximately equivalent to

$$\begin{cases} \dot{y} = \dfrac{-\lambda}{\omega_q} F(y, \psi) \\ \dot{\phi} = \dfrac{-\lambda}{\omega_q \ y} G(y, \psi) \end{cases} \tag{18}$$

where

$$F(y, \psi) = \frac{1}{2\pi} \int_0^{2\pi} f(y, \psi) \sin \psi \ d\psi$$
$$G(y, \psi) = \frac{1}{2\pi} \int_0^{2\pi} f(y, \psi) \cos \psi \ d\psi \tag{19}$$

Usually, functions F and G are quite simple to calculate as they only involve trigonometric functions whose average over one period is either zero or a fractional number.

Amplitude and phase transients are then obtained by numerically solving the associated system, Eqs. (18).

7. A SIMPLE EXAMPLE

So as to demonstrate how the slowly varying amplitude and phase method can be used to obtain the oscillation transient of a circuit, let us consider the circuit shown in Fig. 7 in which components are reduced to their simplest equivalent form.

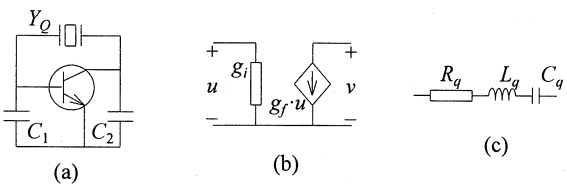


Fig. 7 : *Simple oscillator, (a) circuit, (b) transistor equivalent circuit, (c) resonator equivalent circuit*

The dynamical equivalent circuit is represented in Fig. 8.

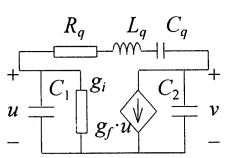


Fig. 8 : *Dynamical equivalent circuit*

As outlined in section 2, g_i and g_f are function of the voltage amplitude y. By comparison with the general case, this circuit is analogous to Fig. 3c so that Eq. (9) takes the form :

$$T(s) = -\frac{N(s)}{D(s)} \tag{20}$$

where

$$\begin{cases} T(s) = s^2 + \dfrac{\omega_q}{Q_q} s + \omega_q^2 \\ N(s) = (C_1 + C_2)\, s^2 + (g_i + g_f)\, s \\ D(s) = L_q\, C_1\, C_2\, s^2 + L_q\, g_i\, C_2\, s \end{cases} \tag{21}$$

Comparison between the terms in $D(s)$ shows that the first one is dominating, thus, putting $D(s)$ under the form shown in Eq. (12) leads to

$$D(s) = L_q\, C_1\, C_2\, s^2 \left[1 + \frac{g_i}{C_1\, s} \right] \tag{22}$$

The differential operator (13) then takes the form

$$T(s) = -\frac{(C_1 + C_2)\, s^2 + (g_i + g_f)\, s}{L_q\, C_1\, C_2\, s^2} - \frac{T(s)\, g_i}{C_1\, s} \tag{23}$$

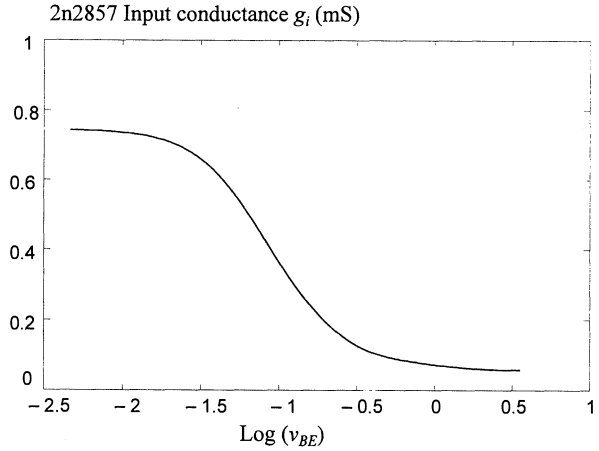
By keeping in the left hand side only the conservative part of $T(s)$ and after arranging the right hand side, the obtained differential operator applied to the transistor input voltage u gives

$$\ddot{u} + \omega_q^2 u = -\left(\frac{\omega_q}{Q_q} + \frac{g_i}{C_1} \right) \dot{u} - \left(\frac{C_1 + C_2}{L_q\, C_1\, C_2} + \frac{\omega_q\, g_i}{Q_q\, C_1} \right) u - \left(\frac{g_i + g_f}{L_q\, C_1\, C_2} + \frac{\omega_q^2\, g_i}{C_1} \right) \dot{u} \tag{24}$$

Application of the slowly varying amplitude and phase method leads to the associated system Eq. (18) which becomes here

$$\begin{cases} \dot{y} = -\dfrac{y}{2} \left[\dfrac{\omega_q}{Q_q} - \dfrac{C_q\, (g_i + g_f)}{C_1\, C_2} \right] \\ \dot{\phi} = \dfrac{1}{2} \left[\dfrac{C_q\, \omega_q\, (C_1 + C_2)}{C_1\, C_2} + \dfrac{g_i}{C_q\, Q_q} \right] \end{cases} \tag{25}$$

Remember that in Eq. (25), $g_i(y)$ and $g_f(y)$ are the large signal transistor parameters which depend on the amplitude y of the base-emitter voltage u (see Sect. 4), they are available under the form of numerical functions (Fig. 9) which can be used to numerically solve Eq. (25).



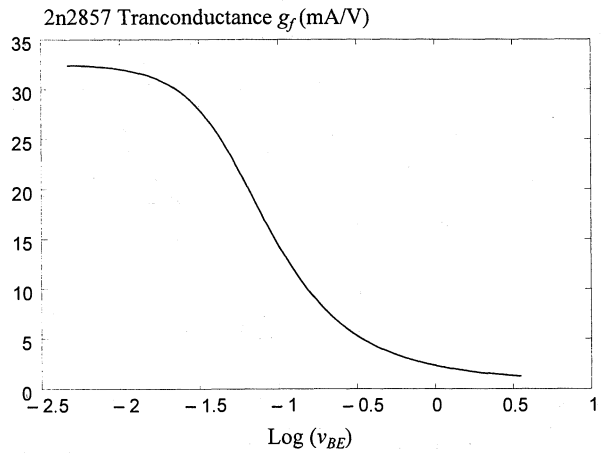


Fig. 9 : *Input conductance and transconductance of a transistor 2n2857 as a function of the base-emitter voltage amplitude*

Figure 10 represents the amplitude transient i.e. the envelope of the signal transient while Fig. 11 represents the frequency transient.

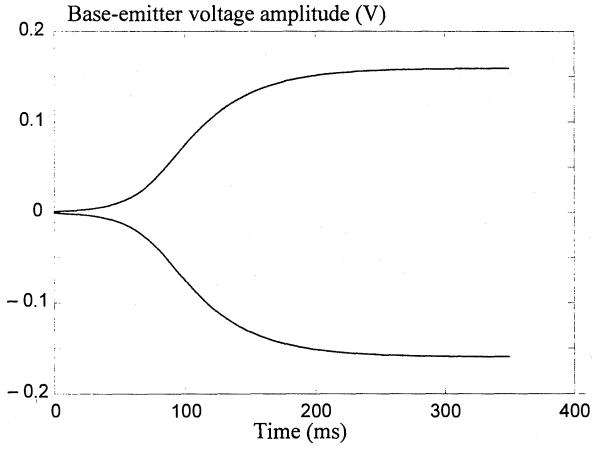


Fig. 10 : *Oscillator amplitude transient*

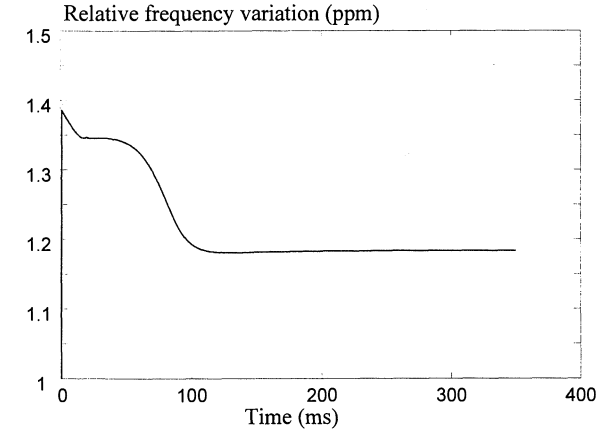


Fig. 11 : *Oscillator frequency transient*

It should be outlined that all steps in the analysis process can be automated so that the derivation of the differential equation and the associated system can be performed by using a symbolic analysis program. Numerical integration of system (18) calls for classical numerical algorithms like the 4th order Runge-Kutta method used to obtain the curves shown in Figs. 10 and 11 [11].

8. CONCLUSION

By exploiting with care a symbolic analysis program, it is possible to automatically derive and code the nonlinear characteristic polynomial describing the oscillation condition of an oscillator circuit. This polynomial not only provides the steady state amplitude and frequency but also gives the nonlinear differential equation whose solution is the oscillation transient. Then the slowly varying amplitude and phase method is used to overcome the problem of transient analysis in high Q-factor circuits. The initial differential equation is transformed in an associated system giving the amplitude and phase or frequency transients within a short computing time by using classical numerical algorithms. All steps of the oscillator analysis process have been automated so as to give the user an original and efficient simulation tool dedicated to the oscillator circuits.

Although the oscillation transient analysis may give the designer interesting features on the oscillator behaviour, it leads the way for the nonlinear analysis of the noise mechanism in the oscillation loop.

ACKNOWLEDGEMENTS

This work has been supported by CNES and DGA under contracts : # 832/CNES/88/5374/00, # 962/CNES/91/1476/00, # 962/CNES/94/1230/00.

REFERENCES

[1] R. Brendel, G. Marianneau, E. Robert, "Nonlinear modelling techniques for quartz crystal oscillators", Proc. of the 5th European Freq. and Time Forum, pp. 377-384 (1991).

[2] R. Brendel, G. Marianneau, T. Blin, M. Brunet, "Computer aided design of quartz crystal oscillators", IEEE Trans. on UFFC, vol. 42, n° 4, pp. 700-708 (1995).

[3] N. Ratier, R. Brendel, T. Blin, G. Marianneau, P. Guillemot, "Nonlinear simulation of quartz crystal oscillators", Proc. of the 9th European Freq. and Time Forum, pp. 88-92 (1995).

[4] M.E. Frerking, *Crystal oscillator design and temperature compensation*, Van Nostrand Reinhold Co (1978).

[5] J.-J. Gagnepain, R. Besson, "Nonlinear effect in piezoelectric quartz crystals", Physical Acoustics, vol. XI, Academic Press (1975).

[6] B. Char, K. Geddes, G. Gonnet, B. Leong, M. Monagan, S. Watt, *Maple V Language Reference Manual*, Springer Verlag, 1991.

[7] N. Ratier, R. Brendel, P. Guillemot, "Quartz oscillators : deriving oscillation condition by symbolic calculus", Proc. of the 10th European Freq. and Time Forum, pp. 442-446 (1996).

[8] N. Kryloff, N. Bogoliuboff, *Introduction to nonlinear mechanics*, translated from the Russian, Princeton (1943).

[9] N.W. McLachlan, *Ordinary nonlinear differential equations in engineering and physical sciences*, Oxford Clarendon Press, 2nd ed. (1958).

[10] A.A. Andronov, A.A. Vitt, S.E. Khaikin, *Theory of oscillators*, Pergamon Press (1966).

[11] W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vatterling, *Numerical recipes*, Cambridge University Press (1989).

TESTS OF TWO 10 MHz CRYSTAL OSCILLATOR DESIGNS BASED ON AN ANALOG A.S.I.C.

S. Michel, S. Galliou

Laboratoire de Chronométrie Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques
26, Chemin de l'Epitaphe, 25030 BESANCON cedex
Tel: 33.381.40.28.41. email: smichel@ens2m.fr

ABSTRACT

Nowadays, electronics needs more and more integration in order to improve abilities as well as to minimize consumption and dimensions. In this point of view, oscillator applications can be developped with A.S.I.C. technologies. We have designed an integrated circuit using a bipolar analog technology. This oscillator can work with a SC or an AT cut B.V.A. 10 MHz resonator.
This circuit has enabled us to make two prototypes : an hybrid ovenized control crystal oscillator and a dual-mode oscillator. Designs, realizations and main performances of both prototypes are presented.

Keywords: O.C.X.O., Analog A.S.I.C., Hybrid technology, Dual-mode, Internally heated resonator.

1. INTRODUCTION

In the field of quartz crystal oscillators, it is necessary to make compromises between stability, power consumption, sizes and cost. A large part of this market corresponds to stabilities between 10^{-8} and 10^{-11} , which is not covered by T.C.X.O.s since their performances are not sufficient, and not by high stability O.C.X.O.s because of their price [1].
The growth of demand for portable applications, like telecommunications for example, leads the design of small, cheap and low power devices.
Thus, in this article, we will present two ways of investigations in order to carry out these different goals.
First, we have designed an integrated oscillator using a bipolar analog technology. This oscillator can work with a SC or an AT cut B.V.A. 10 MHz resonator. A temperature sensor , designed with the same technology, has enabled us to make an hybrid O.C.X.O. prototype. An Aluminium Nitride substrate is used for this application. This material is recently used in hybrid technology and has a good thermal conductivity. Chips are directly sticked on this substrate and connected by wire bonding.
The second prototype is a dual-mode oscillator. This solution has demonstrated its interest in the field of digital T.C.X.O. (stability of $\pm 2.10^{-8}$ over a temperature range of -30°C to +80°C [1]). For an O.C.X.O., this method has an interest only if the power elements are close on or on the resonator. In this way it minimizes the thermal lag between the temperature sensor and the heat source. So, we have made a B.V.A. resonator with internally heated components sticked on the condensers. This crystal is excited on both C mode third overtone (10 Mhz) and B mode fundamental (3.7 Mhz).
This article will present designs, realizations and main performances of both prototypes.

2. DESIGN OF THE A.S.I.C.

2.1 A.S.I.C. technology

We have used a family (POLYUSE THOMSON TMS) of prediffused bipolar « high frequency » analog arrays available in both commercial and military temperature ranges [2].
The silicon process, HF2C is a bipolar technology which provides high performance devices for high frequency and low noise applications:

- NPN transistor features: Ft = 3GHz, β = 100 (Ic = 1mA), En = $4nV/\sqrt{Hz}$

The first metal layer is predefined (cross-unders, power supply distribution, bulk and tub polarisation) and only the second layer is customized with the following advantages:

_ cost and cycle time reduction
_ layout flexibility of dual layer metal structure
_ parasitic effects are minimized due to customizing with second layer metal

2.2 Integrated oscillator

The oscillator has been designed to work with an AT or a SC cut resonator. The integrated functions are shown on figure 1.

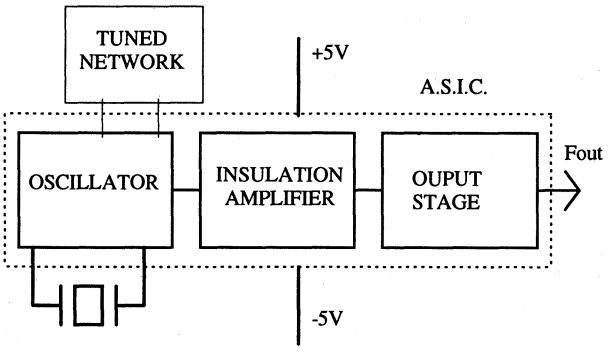


Fig.1: Integrated oscillator block diagram.

The capacitors available on the array are MOS type and their values are 3 and 6 pF. Thus, it was necessary to make direct link between stages by employing resistors polarised by current mirrors.
A Butler type oscillator loop with two transistors has been chosen (Fig.2). An oscillator loop with one transistor couldn't be integrated due to the fact that it needs feedback capacitors of several ten pF.

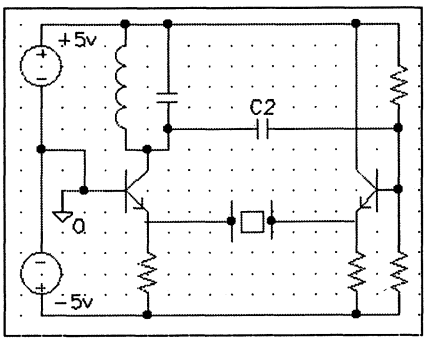


Fig.2: Oscillator loop.

If an AT cut fifth overtone is used,the external tuned network is not required, so the link between the two transistors is direct. In the case of a SC cut C mode third overtone is used, the external tuned network is necessary to ensure right frequency's oscillations and to reject the B mode, and thus, capacitive link must be used between the two stages.
The motional resistor is about 95Ω for AT fifth overtone and SC third overtone cuts. By a brief calculus, we remark that the resonator is guaranted to have a loaded Q-factor higher than the unloaded Q-factor divided by two.
The insulation amplifier is a cascode type, which consists of a common-emitter stage (Q1) followed by a common-base stage (Q2). The load resistance seen by Q1 is simply the input resistance of Q2. This low load resistor of Q1 considerably reduces the Miller multiplier effect and thus extends the upper cut off frequency.
The output stage is a common collector type which is not suitable for low load impedances (below 250Ω). This solution was decided in order to minimize consumption. It is necessary to add an external output stage to make an output impedance of 50Ω.

2.3 Temperature sensor

This network was designed by W. Simonneau in our laboratory, with the same integrated technology. It is based on the temperature effect on a transistor PN junction. It is well known that a junction voltage is near proportional to the temperature by the relationship: $\frac{\Delta v}{\Delta T} \approx -2.5$ mV/°C. Using this property, a temperatutre sensor was developped (fig 3.) and linked with a variable gain amplifier, this structure could be adjust to have an important sensitivity in the range of temperature : 70-90°C which is suitable for an O.C.X.O. The operating point is set by an external resistor.

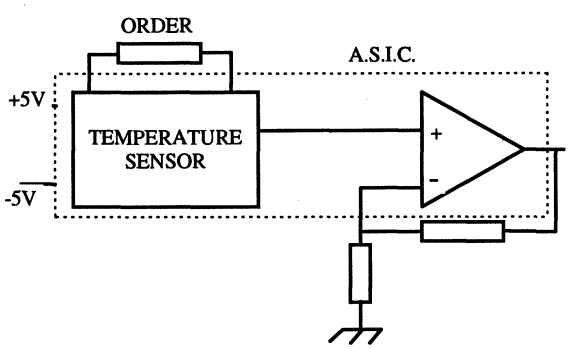


Fig.3: Temperature sensor block diagram

3. DESIGN OF AN HYBRID OVEN CONTROLLED CRYSTAL OSCILLATOR.

The two developped integrated circuits are well suitable for an O.C.X.O. application with low power features and miniature sizes. We describe the design (fig.4) and performances of such application.

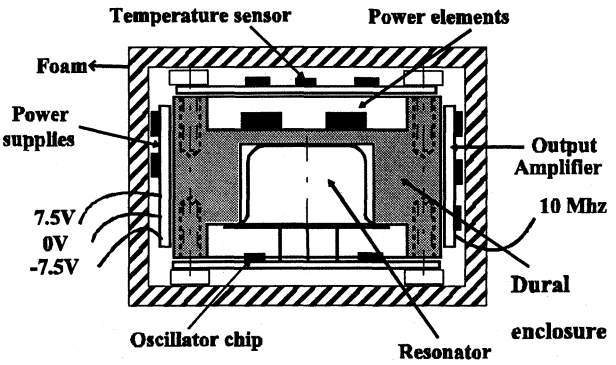


Fig.4 Hybrid O.C.X.O.

3.1 Electrical configuration

The device is divided into four basic blocks: regulator, oscillator, output stage and thermal control (fig.5).

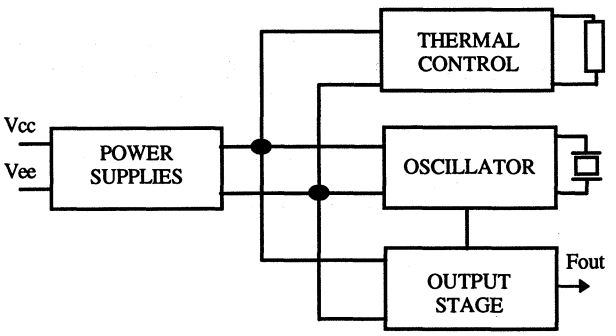


Fig.5: Electrical configuration of the H.O.C.X.O.

The above described integrated oscillator drives a SC cut, C mode 10 MHz B.V.A. resonator. It drives an output amplifier, a Push-Pull type which provides current gain and impedance matching to 50Ω.
The temperature of the resonator and oscillator components is controlled by the integrated temperature sensor which drives a current through the power transistor and resistor as the ambient temperature changes.

3.2 Mechanical structure

The four blocks are materialized by four boards. Power supplies and output buffer are made by S.M.D. components soldered on Epoxy substrate. Oscillator and thermal controller are made by Aluminium Nitride substrate on which Silver Palladium conductive tracks are printed.
Aluminium Nitride is very interesting for its thermal conductivity, which is about 150W/m°K at 100°C. For example, alumina has a ten times less thermal conductivity value. Using AlN in a thermal controlled application seems

very attractive, and could improve oven performance as well as minimize time constants.
But, AlN is a relatively new material in hybrid technology, so it requires an especially designed conductive paste. Adhesion and solderability are not really effective. The drying and firing temperature cycle are more accurate than a usal paste for alumina substrate. These points must be taken into account during the design of a prototype.
On these two boards, the chips are sticked with a silver paste and are connected by wire bonding. This operation has been made in the Laboratoire de Physique et Metrologie des Oscillateurs (L.P.M.O.) by M. Marianneau.
Several S.M.D. components are soldered around the chips. On the oscillator board, the resonator is also soldered. The connections between the boards are provided by simple soldered wires.

3.3 Thermal structure

The thermal structure is composed of an oven machined in Dural. A power transistor and a power resistor are sticked on one side, and are driven by the thermal controller board which is just screwed above (Fig.4).
The resonator package is brought into contact with the oven. The oscillator board and the resonator are held on by four screws.
Power supplies and output buffer are simply sticked on two another sides.
All these elements are insulated into foam.

3.4 Oscillator performances

The typical frequency-temperature curve for this oscillator, without oven controlled, is shown on figure 6. The turnover point is clearly 80°C.
The oven temperature must be set to 80°C so that the oven would still be stable at an ambient of 70°C. This correction must be made in order to set the turnover point of the oscillator in the middle of the temperature range [-30, +60 °C] figure 7. The stability, over this range, is about 7.10^{-9}.
The dissipated power of this unit varies over temperature. Figure 8 shows that at -20°C the oscillator consumes less than 3 W, and less than 2 W at the ambient temperature. The dissipated power of the electronics, without power component's one and with the present thermal insulation, is 400 mW. This power limits the maximum external working temperature range at 75°C. The oven wouldn't still be stable at this temperature.
Another useful feature of such small oscillator is the warm-up stabilization time. Figure 9 shows the typical warm-up performance of this oscillator. The typical warm-up performances are stabilization times to 1.10^{-8} of 14 mn, and stabilization times to 1.10^{-9} of 18 mn. The warm-up stabilization time depends on the volume and the density of the materials used in the oscillator structure. So, this parameter could be improved by the reduction of the sizes and by the reduction of the heated volume. For example, it's possible to make holes in the dural oven.
Figure 10 illustrates the short term stability. It is calculated by the Allan variance method. The main results are 2.10^{-11} for a τ of 1 s and 7.10^{-12} for a τ of 10 s.
Another characteristics are: 10 Mhz sine wave output, 50Ω output impedance, size: 100 cm^3, level +5 dBm.

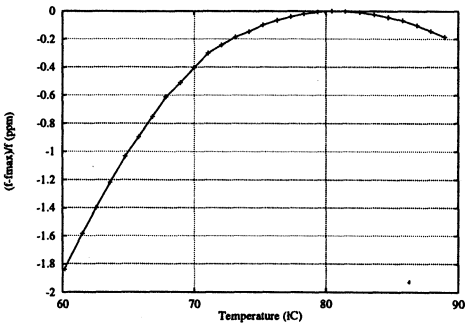


Fig.6

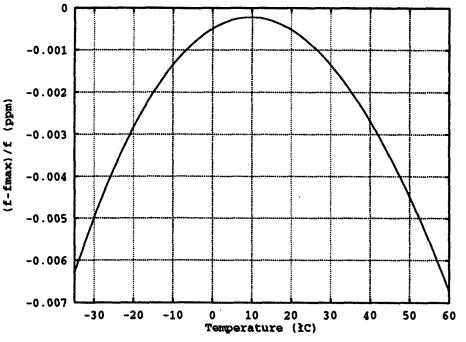


Fig.7

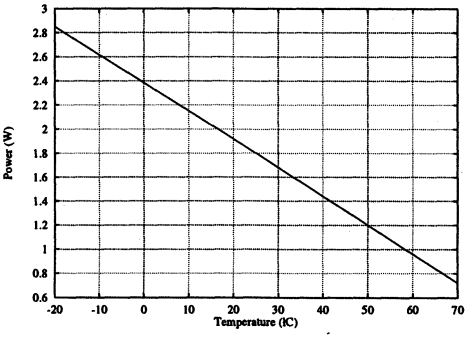


Fig.8

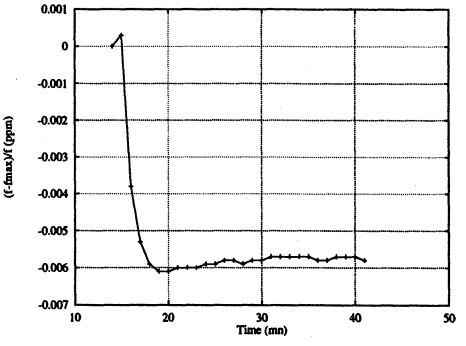


Fig.9

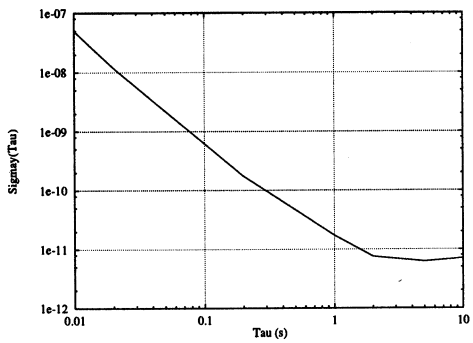


Fig.10

This first prototype has demonstrated the interest of making integrated circuit applications in the field of oven controlled crystal oscillator.
In fact, this prototype is not definitive. With the present architecture, the volume could be easily divided by two. The oven volume could be minimized by an openwork design. Thus, dissipated power and warm-up stabilization time will be divided by more than two with the same thermal insulation.
The integration could be more important. For example, this A.S.I.C. technology has been used in a complete integrated T.C.X.O. by SOREP [3]. It is possible to implement a chip which contains the oscillator, the temperature sensor, the output buffer and even a power element. Likewise, mechanical dimensions and the thermal network could be minimized and then temperature performances will be optimized. This kind of oscillator seems very attractive for a range of applications requiring low cost and performances between T.C.X.O. and ultra stable O.C.X.O..

4. DESIGN OF A DUAL MODE OSCILLATOR.

A dual mode oscillator consists of exciting a resonator simultaneously on two frequencies: a metrologic one and a thermometric one which provides for the resonator to determine its own temperature. This method reduces the thermal lag stemming from a difference in crystal and thermistor effective thermal time constants, thermal gradients and thermistor aging. As many properties which have different effects in both T.C.X.O. and O.C.X.O. applications. For this second kind of oscillator, the dual mode solution is really interesting linked with a control element placed close on or on the resonator plate. Thus, the dissipated power, the warm-up stabilization time and the size can be reduced.
Two methods have been investigated yet. In 1978, Kusters, proposed using the B mode of a SC cut for temperature sensing in a dual B and C mode oscillator [4]. But, excessive activity dips associated with the B mode, reported by Vig [5], and the difficulty of separating the two frequencies have compromised this method. However, in our laboratory, two prototypes of T.C.X.O. and O.C.X.O. based on a dual mode modulation type oscillator were developped in the late 80's [6,7]. In Russia, studies on two mode quartz resonator and automatically balanced modes dual frequency oscillators were developped for D.T.C.X.O. applications [8,9,10].
In 1989, several works using a new method were published [11,12,13]. Schodowski proposed a resonator self-temperature sensing method using a pair of harmonically related C modes in a dual C mode oscillator [11]. This solution combines three advantages: a) the C mode is guaranteed to be free of activity dips, b) the difference between the fundamental and the third or fifth overtone make easier the separation of these frequencies, c) the thermometric beat frequency has a linear dependance on the temperature range. Its curve absolute slope is about -15Hz/°C (-330Hz/°C for the B mode).
In fact, we decided to use a solution combining the two methods described above. This dual-mode oscillator works with a SC cut resonator excited on the C mode third harmonic (10MHz) for the metrologic frequency and on the B mode first harmonic (3.7MHz) for the thermometric frequency. With this method, ability for separating frequencies, good linearity and good sensibility of the thermometric mode are combined. Activity dips are not significant for an O.C.X.O., because the resonator temperature is always around a working point.

4.1 Resonator structure

For this experiment, we have used a special B.V.A.type resonator with S.M.D. resistors sticked on the condensers (Fig. 11).
The goal of this design is to use a dual-mode solution for making an O.C.X.O. application. So, a good thermal servo-control is provided by putting in very closely the temperature probe and the power elements. In our case, resonator, temperature probe and power elements are on the same piece of quartz.

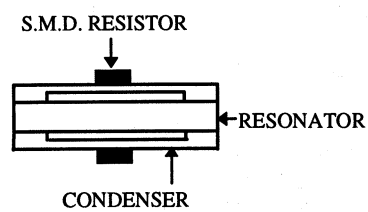


Fig.11: Dual-mode resonator structure

4.2 Oscillator structure (Fig. 12)

A common crystal is shared by two oscillators: one generating the C mode third overtone and using the integrated oscillator presented in the first part of this paper. For the other oscillator, generating the B mode fundamental, we have added an amplifier stage in the loop. This is due to the fact that the B.V.A. resonator we have used is developped to minimise the effects of both C mode and B mode fundamental (Ro=400Ω, Q=42000). A LC tuned network provides the right frequency oscillation in each oscillator loop.
In each oscillator, it is important to decouple the adjacent frequency to provide a very good spectral purity. This is made by the insertion of two LC notch filters of the coupled frequency in each loop.
We have illustrated this property by an A.C. SPICE simulation. We have simulated both oscillator loop for a frequency sweeping around 10 Mhz and a frequency sweeping around 3.7 MHz. In the C mode 3rd overtone oscillator, B mode fundamental is rejected about -100 dB. In the B mode fundamental oscillator, C mode 3rd overtone is rejected about -40 dB. So, the rejection is theoretically less important in the B mode fundamental oscillator. But the stability of this mode will not be changed because of the high stability of the C mode 3rd overtone.

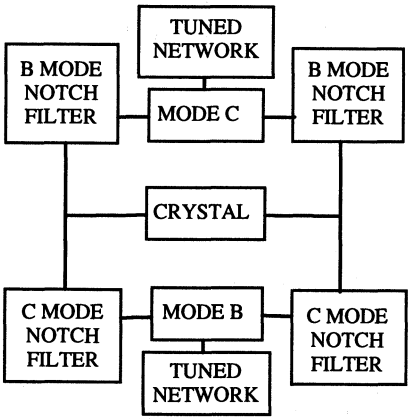


Fig.12: Dual-mode oscillator structure

4.3 Experimental results

Figure 13 shows the frequency temperature curve for the B mode oscillator, in the range 50°C 85°C. The B mode oscillator provides a -100 Hz/°C absolute slope, which corresponds to a -27 ppm/°C relative slope. We can remark that there isn't any significantly activity dips with this resonator.

Figure 14 shows the C mode frequency vs the B mode frequency.

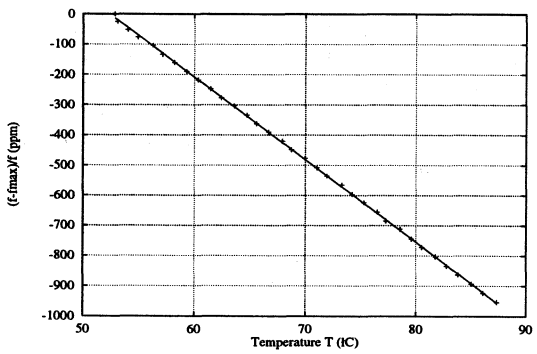


Fig.13

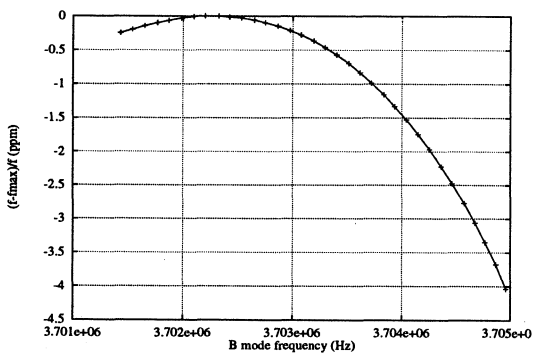


Fig.14

4.4 Prospects

Yet, we have realized the resonator with its integrated power elements. We have tested it in a dual-mode oscillator structure. The main goal of the future work will be to design the thermal servo-control. The more accurate solution seems to be a digital one (Fig.15). This process requires to count the frequency depending on the temperature, to make a comparison with an order which corresponds to the turnover point of the resonator. This variation is used to modulate the power dissipated into the resistors enclosed in the resonator package.

In term of power, we have already evaluated that only 250 mW will be necessary to hold our resonator at its turnover point at an ambient of -30°C. So, the global power dissipated for this application will be reduced in a ratio of 5 to 1, compared to the first prototype. The warm-up time will be reduced in the same ratio (3mn to 10^{-8}).

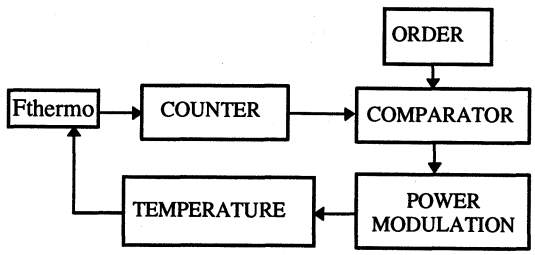


Fig.15: Digital thermal servo-control

5. ACKNOWLEDGEMENT

The authors wish to acknowledge D. Thiebaud for the assembly of the internally heated B.V.A. resonator.

6. REFERENCES

[1] J.J. Gagnepain, « Quartz crystal oscillators from their design to their performances », 4th E.F.T.F. (1990).

[2] Thomson Composants Militaires et Spatiaux, Reference manual T.S.F.L series.

[3] Y. Richard, J.M. Tarot, « Analog A.S.I.C. design for miniature temperature compensated crystal oscillators », 9th E.F.T.F. (1995).

[4] J.A. Kusters, M.C. Fischer, J.G. Leach, « Dual mode operation of temperature and stress compensated crystals », 32nd A.F.C.S. (1978).

[5] J.R. Vig, R.L. Filler, J.A. Kosinski, « SC-cut resonators for temperature compensated oscillators », 36th A.F.C.S. (1982).

[6] S. Galliou, M. Mourey, « Compensation statique et dynamique en temperature d'un oscillateur à quartz bi-mode », 1st E.F.T.F. (1987).

[7] S. Galliou, M. Mourey, « Dual-mode temperature controlled quartz resonator », 3rd E.F.T.F. (1989).

[8] I.V. Abramson, « Two-mode quartz resonators for digital temperature compensated quartz oscillators », 46th A.F.C.S. (1992).

[9] A.V. Kosykh, I.V. Abramson, V.P. Bagaev, « Dual-mode crystal oscillators with resonators excited on B and C modes », 48th A.F.C.S. (1994).

[10] A.V. Kosykh, S.A. Zavyalov, « Modulation type dual-mode oscillator intended for micro-chip realisation », 49th A.F.C.S. (1995).

[11] S.S. Schodowski, « Resonator self-temperature-sensing using a dual-harmonic-mode crystal oscillator », 43rd A.F.C.S. (1989).

[12] A. Benjaminson, S.C. Stalling, « A microcomputer-compensated crystal oscillator using a dual-mode resonator », 43rd A.F.C.S. (1989).

[13] M. Bloch, M. Meirs, J. Ho, « The microcomputer compensated crystal oscillator », 43rd A.F.C.S. (1989).

CORRELATION OF PREDICTED AND REAL AGING BEHAVIOUR OF CRYSTAL OSCILLATORS USING DIFFERENT FITTING ALGORITHMS

O. Leibfried, B. Neubig
TELE QUARZ GROUP
D-74924 Neckarbischofsheim; Germany

1. Abstract

This paper reports on the test results determinating the ageing behaviour of quartz oscillators more precisely as it was possible before. Different methods are presented for the calculation and curve fitting of measured ageing data. It was tried to minimize the time necessary for correct predictions of the OCXO in the ageing system.

The OCXO ageing system of TQ allows a simultaneous measurement of up to 1600 OCXOs for any period of time. The measurement data of all OCXOs produced are put into archives since the end of 1995. Thus as data basis one could go back to extensive archives of ageing data in which there is now the data of over 50000 OCXOs. OCXOs were sometimes left in the measuring system for tests during new developments; some of them longer than 250 days. Those data were used for the tests.

2. Introduction

The prediction of long-term ageing behaviour of high frequency OCXO is always connected with the factor of uncertainty whether the measured ageing curves of the oscillator are really kept. For the calculation of the coefficients for aging prediction several functions can be applied:

The exponential function [1]

$$(1) \qquad \Delta f / f_0 = a_0 + a_1\left(1 - \exp\left(-a_2 t\right)\right)$$

The polynomial function [2]

$$(2) \qquad \Delta f / f_0 = a_0 + a_1 t + a_2 \sqrt{t}$$

The pure logarithmic function [1]

$$(3) \qquad \Delta f / f_0 = a_0 + a_1 \log t \qquad \text{or}$$

$$(4) \qquad \Delta f / f_0 = a_0 + a_1 \ln t \qquad \text{with } t \geq 1$$

and the modified logarithmic function [3],[4]

$$(5) \qquad \Delta f / f_0 = a_0 + a_1 \ln\left(a_2 t + 1\right)$$

This function, proposed in MIL-O-55310 [3] is the most recently used function for aging estimations. [4].

TELE QUARZ is using the modified logarithmic function (3) and the polynomial function (2) depending on which of the two functions describes the measured frequency points better.

The coefficients for the aging functions result from [3] through the curve fitting of the measurement data of the OCXOs collected over 30 days. After this relatively short time related to the period of prediction the coefficients were adapted to the measuring data with least square algorithms. The ageing values for the prediction are now derived from these coefficients.

3. Problem description

At present TELE QUARZ is adapting the coefficients of the logarithm function to the measuring data through the nonlinear least square fitting algorithm proposed by Filler [5]. The measuring time is 10 ... 20 days depending on the OCXO type. The coefficients obtained are directly used for predicting the ageing behaviour. Long series of measurement revealed that in almost 80% of all cases the combination of modified log function and Filler algorithm provide too bad results. That means that the real ageing rates of over 80% of all measured OCXOs are better than the prediction. Figure 1 shows a typical measuring curve with the corresponding fitted curve. The real ageing curve has a clearly flatter course than the fitted function. Thus the predicted values are higher than the actual expected ones.

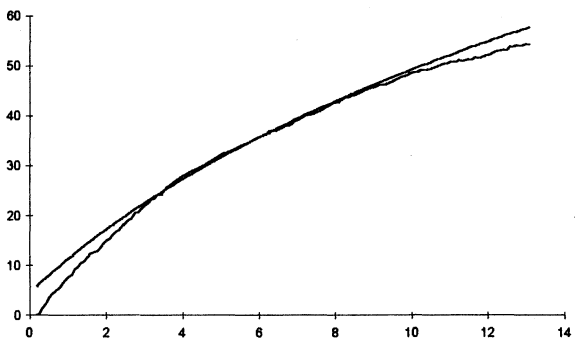


Figure 1 Bad fitting with Filler algorithm

This fact which is at first positive for the customers is again relativated because of higher oscillator prices caused by the internal rejection of good pieces.
All OCXOs used for the tests have to be measured for the period of predictions to be able to compare ageing curves with the curves precalculated. The real ageing curves can now be compared with the precalculated curve which is obtained from the measuring data of the first days

4. Curve discussion

Due to high oscillator frequencies in the MHz range and the essential measuring and processing accuracy the function proposed in MIL Standard MIL-O-55310B for calculating the ageing behaviour

$$(6) \qquad f(t) = b_0 + b_1 \ln(b_2 t + 1)$$

produces high quantities of data. To keep the necessary storage capacity of the measuring computer small it is possible to use a modified formula which gives the ageing values .....

$$(7) \qquad frel(t) = \frac{f(t) - f(t_{Start})}{f(t_{Start})} \qquad \Longrightarrow$$

$$(8) \qquad frel(t) = a_0 + a_1 \ln(a_2 t + 1)$$

$$\text{with} \quad a_1 = \frac{b_1}{f(0)} \qquad a_2 = b_2$$

$$(9) \qquad a_0 = \frac{f(0) - f(t_{Start})}{f(t_{Start})}$$

With this modified aging formula the ageing measuring data can directly be given in parts per billion. Due to the small numerical values the data volume also remains small as the absolute frequency has to be stored only once. Due to the OCXOs of the long-term measurement it can be shown that the modified log function (5) is correlating well with reality (figure 2).

Filler mentions in [5] that the ageing of quartz crystals and oscillators is influenced from many different mechanisms. Some influence factors can be minimized. The temperature dependence of oscillators for example in air-conditioned rooms is only of minor importance. Nevertheless many oscillators show a strange ageing behaviour. Above all in a complex ageing system such as used at TELE QUARZ the OCXOs seem to influence each other through injection locking, thermical influence and mechanical vibration or bumps.

Thus in some cases several OCXOs show frequency jumps at the same time or change their current consumption simultaneously. In these cases a precise estimation of ageing is nearly impossible as the cause usually is not reproducable. In case of OCXOs with continuous ageing behaviour the prediction accuracy can be considerably improved by suitable actions.
Figure 2 shows a typical aging behaviour of an AT OCXO representative for all measured oscillators. The curve was fitted over the whole measuring time. With over 70% of the measured ageing curves there is a similar good correlation. About 20% of the OCXOs show frequency jumps. They were not considered in this test.

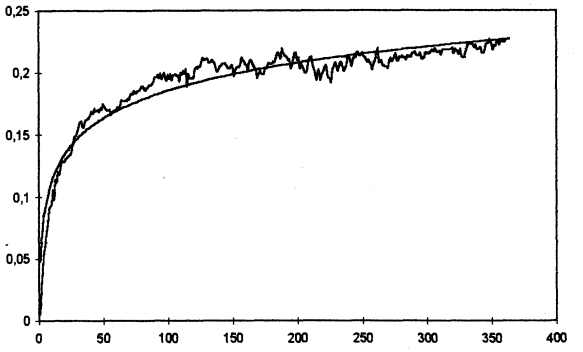


Figure 2 Measured ageing curve with fit over one year

The other 10% are very stable OCXOs with SC cut crystals, where other effects which are stronger than aging (like temperature stability, influence of adjacent oscillators, supply voltage variations, injection locking etc.) influence the ageing measurement. Thus the fitting algorithm can often find no satisfactory solution which partly leads to extreme results. (figure 3)

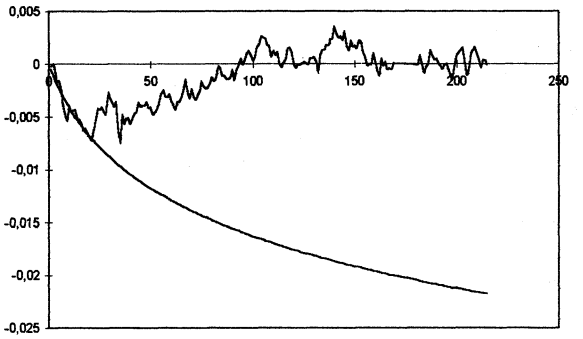


Figure 3 Very stable SC OCXO with bad fitting

The time derivative of the ageing formula is as follows

$$(10) \qquad frel'(t) = \frac{a_1 a_2}{a_2 t + 1}$$

for large values of t it is:

$$(11) \qquad frel'(t) \approx \frac{a_1}{t}$$

for small values of t it is:

$$(12) \qquad frel'(t) \approx a_1 a_2$$

Therefore, for large t the annual ageing is nearly independent from a_2. It can be described for large values of t in limited time intervals by a straight line with the slope a_1 / t. This is the case after 30 to 250 days depending on the oscillator type.The curve has the biggest slope of $a_1 a_2$ at small values of t. In this time interval usually the curve fitting is done. So for the algorithm the long term coefficient a_1 has nearly the same influence on the least square result than the short term coefficient a_2. Because of this the choice of algorithm to the coefficient extract has great influence on the long term gradient a_1 of the function. Some methods and algorithm are described to increase the accuracy of prediction.

5. Results

The correlation of the fitted measuring curves with the real measured ones can be improved, if the term a_0 in equation (5) is forced to zero. Through this restriction the adaptability of the fit to the measuring values is hardly changing. Figure 4 shows a fitted curve with a_0=15. A second one with a_0=0 was fitted to this curve by least suare algorithm.

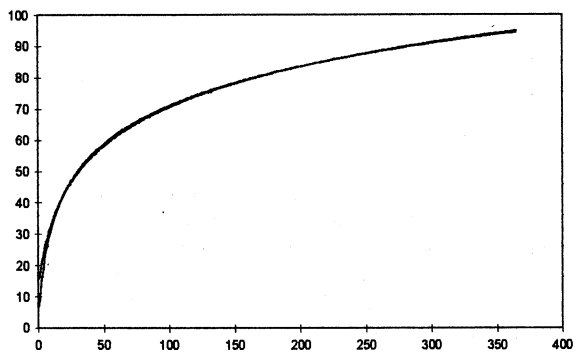


Figure 4 Differences of the curves a_0=0

The diffences are minimal and only important in the first days of operation. The long-term behaviour does not change through this measure as a_0 is only a mathematic offset for the curve. For the fitting algorithm, however, a_0. appears as a series of square sums. Therefore several solutions are possible for the fitting algorithm. So if a_0 is kept at zero during the fitting the algorithm often can provide better solutions (figure 5).

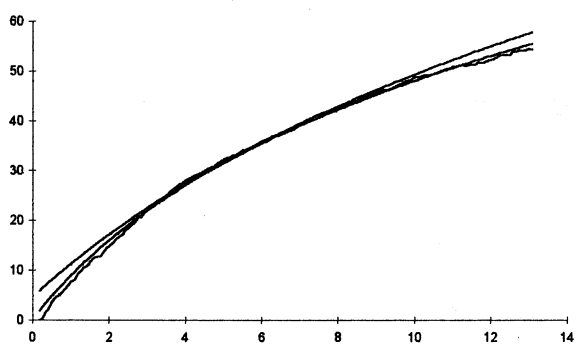


Figure 5 Improvement of the fittings with a_0=0

In the following tests only the offset free Mil Logarithm function was used due to the positive results with a_0=0.

After the first power on OCXOs often show a strong initial aging while the long term aging behaviour shows good values after initial phase. This behaviour seems to depend on the oscillator type as it usually can be observed at all OCXOs of the same type to nearly the same extent.Curve fitting over the whole measurement time of one year shows that there is good correlation with the logarithm function even in case of strong initial aging. Due to the strong initial drift, the fitted curves show poor correlation when calculated after the first 20 days.

This effect occured at all tested OCXOs of this type. Thus it was supposed that the strong initial drift is not correlated to the long-term stability but other factors. As the long-term behaviour seems to be independent from the initial drift it was tried to give the measured aging data at the end of the measurement interval a higher weight on the result. A simple possiblity of realisation is the use of a weighted fitting procedure where the measurement values are evaluated with a variable significance. The significance of the measurement values has to increase the closer a measurement point is located to the end of the measuring interval. For weighting the measured values the solution of a differential equation of 1rst order was used. In practice this can be achieved relatively simply. The fault squares of the individual diffences are muliplied with the weighting equation before suming up the function. This gives the following expression whose value has to be minimized. (13)

$$lsq = \sum_{t=0}^{t\max}\left[\left(\frac{\Delta f}{f_0}(t) - a_1 \ln(a_2 t + 1)\right)^2 * \left(1 - e^{-bt}\right)\right]$$

Through the variation of the coefficient b the influence of the weighting function can be „tuned" continuously. At high values for b the weighting reaches the asymptotic boundary (=1) even at small time values. Thus the weighting fit has no influence anymore and algorithm works as before. Small values of b, however, are more interesting. Here each series of measurement can be weighted which set priorities at the end of the measurement interval.

The coefficient b, however, is different for each OCXO-type and has to be evaluated separately for each type during the evaluation process. This method provides better results as the simple fit of the measurement data (figure 6). The measured values, however, have to show monotonic behaviour as frequency jumps and other factors are also weighted.

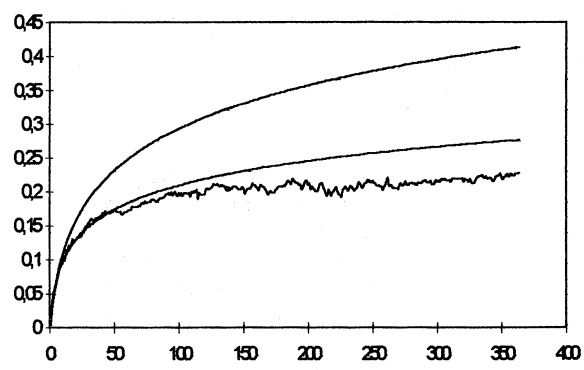


Figure 6 weighted least square fit

A lot of physical processes are balancing procedures which approach the state of equilibrium by an exponential function. It was tried to consider this compensation function in the ageing calculation. Figure 7 shows a measured curve (C) with a fitted modified logarithm function (B) over 365 days. Curve (A) is the conventional fit over 20 days. If the

measurement data (C) are subtracted from the fitted curve (B) and the difference of both curves at the time t=0 is added as offset to each point of the result function, an error curve (G) is achieved. Through the calculated points of the error function a function (F) can be fitted which corresponds to the solution formula of a differential equation of first order.

$$(14) \qquad df / f = k\left(1 - e^{-bt}\right)$$

The coefficients k und b are again determined through least square algorithms. This equation now includes the influences of the oscillator and the initial quartz crystal drift. If each point of this function is subtracted from the originally measured series of data (C) a new measurement curve (E) is achieved which is reduced by the initial influences. If the modified logarithmic function according to theMIL standard with a_0=0 is now fitted through this modified series of data (curve C) the calculated curve (D) shows a far more better correlation to the real aging response.

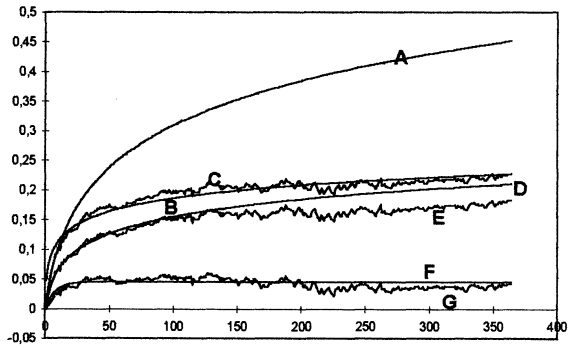


Figure 7 subtracted initial frequency drift

With all tested OCXOs of this type similar coefficients for k and b could be found. With the coefficients once calculated for the error function the prediction accuracy of all oscillators of one type can be improved (figure 8).

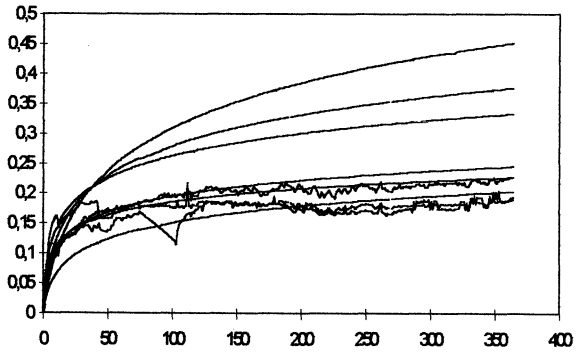


Figure 8 Improvement of fitting accuracy OCXO

Unfortunately this procedure is difficult to integrate in production as the ageing data of each OCXO type is required for the first year. All of the described methods allow an improvement in prediction accuracy. In series production, however, they can only be used with great expenditure. Due to the enormous variety of TELE QUARZ products this method can be used only in exceptions for high production volumes.

The major disadvantage of the described methods is that they do not eliminate the causes but only includes them in the calculation. The long-term behaviour of OCXOs can be determined earlier but the initial aging drift for the OCXO with tight ageing specification must be well within the measured interval.

In the following figure (9) the same OCXOs are shown as above. The measurement values of the first 20 days were ignored and the measurement value of the 21rst day was taken as reference point for the relative frequency deviation. Thus a 20 days' preaging of the oscillators can be mathematically added before the actual ageing measurement. Figure (9) shows a considerably better correlation of the fits with reality. OCXOs with less initial drift even have a better prediction accuracy.

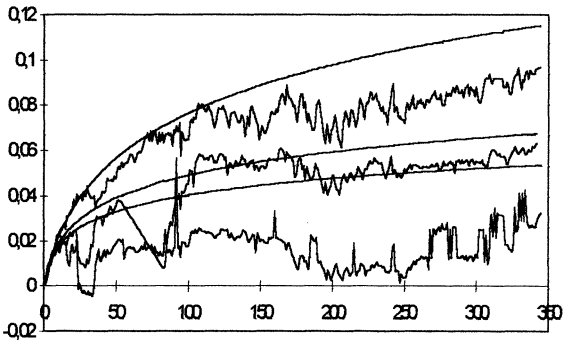


Figure 9 Improvement of fits obtained through preageing

6. Summary

It could be shown that least square algorithms being applied for ageing measuring data provide better results if the term a_0 of the logarithmic function is kept zero. The calculated curves correlate better with the aging observed in reality.

The prediction accuracy of oscillator ageing rates can be considerably increased by mathematic methods which consider the initial aging drift effects.

All mathematic procedures, however, cannot supply better results than simple preageing processes. The longer the oscillators are kept in operation, the better becomes the prediction accuracy.

Especially oscillators with SC cut crystals which show excellent ageing behaviour are significantly influenced by effects like operating voltage fluctuations, injection locking, initial drift, TK etc so that an accurate aging statement is not possible after only 20 days of operation. The aging behaviour of such highly stable OCXOs can only be determined after longer time periods of 30 days and more.

7. Literature

[1] Miljkovic, M.R., Trifunovic,G.Lj., Brajovic,V.J.: Aging prediction of Quartz Crystal units; Proc. 42nd FCS (1988), p.404-411

[2] TELE QUARZ internal standard for Aging evaluation (1995)

[3] MIL-O-55310, Military specification, General specification for crystal oscillators

[4] Neubig, Bernd, Comparison of passive and active aging of SC-cut and AT.cut crystals, Conference publication 418, EFTF 96, p.37-43

[5] Filler, Raymond L., Aging specification, measurement and analysis, Proc 7th Quartz devices conference (1985), p.93-104

Improved BVA resonator - oscillator performances & frequency jumps

K.K. TULADHAR, G. JENNI, JP. AUBRY

Oscilloquartz SA, rue des Brévards - Neuchâtel - 2002 - Switzerland
tel : ++ 41 32 722 55 55 fax : ++ 41 32 722 55 56

1/ Introduction

BVA technology have proven to be, since its introduction by R.Besson[1] in the 70's, ever and ever the state of the art or crystal oscillators.

Performances of resonators and oscillators have been enhanced significantly during the last decade, through a better understanding of physical behaviour, from quartz material to electronic components, but the BVA concept, while getting advantages of the general progresses, is still a step further.

In the 70's, speaking, on an industrial basis, ageing of a few parts in 10^{-11} per day was just a nice dream. Nowadays, the question is « Is the yield better than 70% for a specification of 10^{-9} per year »....

The BVA concept itself has been intensively studied and improved during the last years, from mechanical knowledge of its structure, and the improvements made in the last decade on BVA deal with the parts assembly, the mounting structure, the mechanical behaviour and involve various FEM analysis.

The tools which are considered now are :

material, design, process of crystal resonator
minimisation in electronic influences in the oscillator

This paper presents mainly the results given by the use of :

SC cut 5 MHz BVA resonators
high oven and oscillator intrinsic stability

Perhaps for the first time, we are in position to compare BVA resonator at the same frequency (5 MHz) with AT or SC cut, with basically the same equivalent network, in the same electronic assembly.

Two different basic oscillators are studied, both using AT and SC resonators. Environmental influence and protection will be discussed.

2/ State of the art in BVA technology

In his 95'AFCS paper, J Norton[2], from Applied Physics Laboratory, shows the improvement of performances seen during the last four decades of time, since the pioneering work of Warner in the 50's. He pointed out that among a continuously step by step progresses, following the initial design of 2.5 and 5 MHz AT cut fifth overtone resonator, the main two events were the introduction of SC cut by E EerNisse and J Kusters and the introduction of BVA design by R Besson.

Phase noise of -110 dBc.Hz^{-1} @ 1 Hz from carrier was the ultimate performance in the 70's, and main steps of improvements since that time were :

SC in commercial field since the 80's

BVA, in Laboratory since late 70's, and the commercial production in Oscilloquartz starts in 1982.

Values of $\sigma_y(\tau)$ below 1 10^{-13} @ 10 or 100 s were measured in laboratory since the mid 80's, with latest minimum value of -137 dBC/Hz @ 1 Hz and 3.7 10^{-14} at 10s being measured for a special design[3].

In a paper entitled « future trends on oscillators stability »[30] given at the Seminar on « Etalons de fréquence » held in LPMO in 1981, the projected figures were around -122 or -126 dBc @ 1 Hz and between 5 or 2 10^{-10} vs. temperature, for so called commercial or laboratory units. One can see the improvements since that time....

3/ Oscillators requirements

Oscillators requirements are listed below for the 3 main areas of very high stability oscillators, with emphasis on the key points for each application.

3-a : Telecommunication applications

Main requirement are, at least for up to date networks, to keep the local time of a node in « loose synchronisation » conditions, usually called « hold over » mode.

Main topic of interest is stability versus temperature. The oscillator being locked before drop out of the synchronisation signal, the time error is :

$$\Delta\tau = \int_{t_0}^{t_1} \left(\frac{\Delta f_N(t)}{f}\right).dt$$

Δf_N being the frequency drift between the local oscillator and the reference one. This being rewrite as :

$$\Delta t(\tau) = \Delta\tau_0 + \frac{\Delta f_0}{f}.(\tau - \tau_0) + \int \left(\frac{\Delta f(t)}{f}\right).dt$$

where $\Delta\tau_0$ and Δf_0 are the initial error setting in time and frequency, « syntonisation and synchronisation offset »

$\Delta\tau_0$ and Δf_0 can be ignored in normal operation, since the oscillator is locked before the cancellation of external synchronisation signal. The only Δf_0 influence can come through the DAC resolution within the locking loop or $\delta f/f$ due to DC variation (usually less than 10^{-10}). The fact that the oscillator is already « on » before the beginning of the « hold over » mode also cancel any real « frequency retrace » needs.

$$\frac{\Delta f(t)}{f} = b.(t - t_0) + \sum P(t).S(t)$$

is the sum of perturbation effect on local frequency.

b is the linearised ageing slope of oscillator frequency and b.t is the oscillator natural drift during the time of examination,

* P(t) is the perturbation as a function of time applied, between time τ_0 and τ, on the oscillator which exhibit a sensitivity S(t) against this perturbation.

The following tables give the time error after 3 days in various assumptions of temperature drift or oscillator sensitivity:

* a 30 °C variation : abrupt or linear, i.e. P(t) = 30 (Cste) or P(t) = k t with k = 10°C/day

* applied on two oscillators :
a good classical : 3 10^{-9} in -30, 70°C, 5.10^{-11}/ day

a high stability one : 1.8 10^{-10} in 0, 60 °C 5.10^{-12}/day

The four calculation are given for :

#1 : 3.10^{-11} / °C, slow T variation along the 3 days
#2 : 3.10^{-12}/ °C, slow T variation along the 3 days
#3 : 3.10^{-11} per °C, abrupt T variation at t_0
#4 : 3.10^{-12} per °C, abrupt T variation at t_0

Time Error @ 3 days	3.10^{-11} / °C	3.10^{-12}/ °C
slow T variation	**116 µs**	**11.6 µs**
abrupt T variation	**232 µs**	**23.2 µs**

During the same 3 days of interest, the ageing contribution will be :

	5.10^{-11}/day	5.10^{-12}/day
time error	**20 µs**	**2µs**

During these 3 days, an initial offset (DAC or any DC effect...) of 1 10^{-11} will introduce a 2.5 µs time error.

The weight of each contribution confirms that the main topic to achieve in Telecom application is a very low frequency vs. temperature characteristic.

One can also argue, but we will not develop this point here, that a stable loop require a frequency vs. voltage slope as monotonic as possible, to maintain stability and margin during all the time life of the oscillator.

The last, but not the least, of the requirements to define a good telecom oscillator, is to provide a long life-time, i.e. 20 years, without having to change the active device of the oscillator.

The G-811 and G-812 norms require, among various types of oscillators, the so-called « transit node oscillator » to be within a few µs per day on hold-over mode.

Excellent timing accuracy is required in application such as CDMA and attempts to get such an accuracy through temperature and ageing compensation of GPS disciplined oscillators[4] are in competition with the direct accuracy of a « good enough » oscillator

3-b : Metrology and science applications

Lowest noise (below -130 dBc.Hz^{-1} @ 1 Hz offset from carrier) and the best short term stability (below 1 10^{-13} in Allan variance between 1 and 100 s) is generally a « must » as a local oscillator in atomic, cold ions,... clocks to get advantage.

A limitation in Maser stability which has been reported at 1 second by HE Peters & al.[5],
in its EFTF 94's paper devoted to Trapped ions standards, Lute Maleki[6], from JPL, reports an actual « average performance » of 7 $10^{-14}.\tau^{-1/2}$, on trapped ions or optical standards based on Hg, Ca, Yb, Sr
In Cesium fountain, G.Santarelli [7] gives in 94 a limitation of 6 $10^{-13}.\tau^{-1}$ which has been improved in 95 by S Ghezali[8] as 2 $10^{-13}.\tau^{-1}$ limitation, instead of the 7 10^{-14} expected law, the degradation being due to the local oscillator
In optically pumped Cs , E de Clerq & al[9]. gives a short term law as 6 $10^{-13}.\tau^{-1/2}$
We have to go to cryogenic sapphire resonators[10] to find stability better than some 10^{-14} below 1 second .

These published results show that the local oscillator phase noise or short term is a limitation for these applications. A target of 1 10^{-13} for τ = 1 s is a good goal for this field of application.

3-c : Orbitography and localisation

In a high precision system such as DORIS[11], whose target is the sub centimetre accuracy in orbitography, medium term frequency stability, i.e. mean value of slope of a linear fit among 90 samples (τ = 10s) around 10^{-13}/ minute, and a least square fit residue better than 10^{-13} are required.

The goal of the work we have started is then an oscillator which is able of :

1 10^{-9} per year
$\sigma_y(\tau) < 1\ 10^{-13}$ Allan from 1s to t > $100_{(s)}$ s
$\mathcal{L}(f) < -130$ dBc/Hz @ 1 Hz from carrier(target -135 dBc/Hz).

4/ Optimisation tools

One can act on three design parameters to try to achieve such a target specification : the resonator, the oscillator and a protection against environment.

4-a : Resonator

A crystal resonator is defined by its :

crystal cut : Three cuts can be used for this oscillator :AT, BT or SC (or any other doubly rotated cut).

Because of thermal stability and difficulties which can be observed in getting high loaded Q with BT cut, we restrict our scope to AT and SC

design : One can find in the literature[21,22,25,26] some interesting analysis to show how the design can interfere with the stability.

The acoustic wave distribution within the volume can be an origin of frequency fluctuation because of :

* amplitude symmetry being perturbed by clamp, gives a frequency sensitivity to any variation of the local geometry of the mount
* bevelled resonators show a significant acoustic sensitivity to mounts
* activity dips/frequency jumps, whose origin are, at least for some of them, in the design (see annex A)

material : The material influence the frequency stability through different ways[27] :

dislocations content

- chemical reaction on the surface creates etch channels or pits (<> dislocations), possibly responsible of noise
- concentration of electrical defects along dislocation lines
- dislocations influence mode shape (possible noise contribution[22]) and influence on mechanical stress sensitivity, i.e. in G sensitivity, because the acoustic distribution defines by itself the « coupling volume » between stress and acoustic vibration (mathematical contour of integration)[23]

OH content & Q factor

- device 1/Q no longer proportional to alpha IR for alpha values lower than 0.03 cm^{-1} [27], but material IR absorption below 0.02 cm^{-1} are available[28].
- high Q resonators are key points for low noise oscillators
- a fundamental question is still not answered yet: what is the physical phenomenon limiting quartz device Q (intrinsic phonon structure, other local absorbers, crystal geometry, crystal surfaces, impurities and dislocations are some candidates)
- one need to explore others crystal synthesis ways[29], at least for comparison purpose and identification of contributions

inclusions

they are responsible for local variation of material stress which are, in turn, responsible of local frequency discrepancy

chemical purity

The most commonly identified impurities are aluminium ions on Si sites (which generate perturbation of phonon spectrum), with interstitial monovalent ions Na^+, Li^+, K^+ coming along for charge compensation[13,14,15].
These impurities, in conjunction with electroelastic effect[12], which express a frequency variation of SC cut to DC electric field, and because of their own mobility under electric field, can generate frequency variation :

$$\delta f/f = K_{el} * (V_0 - V_D(t))$$

with K_{el} the electroelastic coefficient of SC cut (# 10^{-8} / V.mm^{-1}), V_D being the depolarisation field due to impurities migration under DC field.
These impurities are also involved in frequency variation under photon stress[18,19].

manufacturing process

The manufacturing process is also of main importance, from physical or chemical aspect :

- blank geometry (i.e. acoustic symmetry) can be perturbed by blank lapping and polishing[16,17,22]
- baking temperature
- base and can definition
- cleanliness, outgassing ...
- stress relief in plate (mounts, plating ...)
- frequency jumps prevention,

4-b / Oscillator

It is more than well known that a circuit is easy to oscillate, but it is far more difficult to design a good oscillator. The main aspect of interest are :

- network topology / circuit design, along with known criteria
- technological design / physical realisation
- loaded Q, crystal power, impedance loop...

Modern simulation tools can help to evaluate oscillator sensitivity to components variations, either on frequency drift[31,32], or to perform noise prediction[33].

4-c / Reduction of external influence

Resonators and oscillators frequency is sensitive to external perturbation in two ways :

⇨ Environmental sensitivities of frequency *measurements* to vibrations, mechanical stresses, pressure variation

⇨ factors of influence on frequency *instability* in oscillator, such as

mechanical stresses : The general form of frequency variation due to a mechanical stress applied to the resonator is given by :

$$\frac{\delta f}{f} = \int_V K * T$$

where T is the stress and K the frequency sensitivity to this stress.

The integral contour V is defined by the acoustic wave amplitude distribution. The stress distribution have to be evaluated within this volume. It explains sensitivities to lack of blank symmetry, non symmetric mass loading effect, crystal dislocations influence...
The modification of acoustic distribution (by electrical or local additional mass loading) was also used to reduce G-sensitivity of resonators[11].

The analytical form of the frequency deviation under a mechanical stress have been derived[12] as

$$\frac{\Delta f}{f} = \sum_i k_i T_i$$

where k_i are the sensitivity coefficients, derived for SC cut assuming low strain induced effects, in units of 10^{-12} m^2/N :

k_1 =	-2.38	k_4 =	39.85
k_2 =	-6.55	k_5 =	18.91
k_3 =	1.35	k_6 =	43.86

and T_i are the six principals stresses applied in the central part of the crystal. These stresses have to be evaluated under the active area of acoustic vibration by analytical or other mechanical tools. This mechanism is involved in G or pressure sensitivity[24], or in magnetic field sensitivity[20].

Any stress relief will be seen as a frequency variation.

Among different origins of frequency instabilities which has been pointed out in the last few years are the frequency jumps. Some of then have been described and identified along with different criteria to classify these jumps. Angle of cut mismatch or parallelism between resonator and condensor can be some of the origin of frequency jumps in BVA resonators (Annex A)

Parametric stress : At the oscillator level, external perturbation such as AC spikes, temperature, humidity, pressure variation, acts as frequency instabilities. Variation of sensitive components, such as inductors, under ageing, temperature, magnetic field, humidity... induce large frequency sensitivities. Technique like B-mode reduction[21] can help to reduce the degree of criticality of such components.

5/ Targets and Methodology

Crystal batches

Resonator batches were build (or are under process actually) with the following designs :

⇨ AT & SC resonators, in BVA technology
⇨ Quartz materials
Std Q (0.03 cm^{-1}) and High Q (0.02 cm^{-1})
Std Q - Swept and high Q swept
high Q - low dislocation content
⇨ Blank preparation : classical or chemical polishing
⇨ Various thermal treatment-outgassing conditions

Oscillator definition

⇨ a high performance oscillator (low noise such as OSA 8600) will be considered because of its special crystal loop design, to be compared to a more simple oscillator
⇨ two different oven will be considered, to evaluate the thermal contribution on noise : a Dewar and a standard oven

We will be in a position to compare :
classical vs. BVA resonators
AT vs. SC cut
2 electronic oscillator design (1&2)
2 different thermal structure (dewar & classical)

For better comprehension, these configurations are labelled the following way :

	AT	SC	BVA AT	BVA SC
Osc 1	8665/6	8662/3	8695/6	8692/3
Osc 2			8600	8607

Our targets can be expressed as :

- evaluate the different origin and amplitude of frequency instabilities, by comparison of different configuration
- evaluate the ultimate performance obtained by the use of the today's best oscillator in OSA and by the use of the best promising resonator, BVA SC cut at 5 MHz

6/ Experimental results

6-a) Ageing

Results collected 30 days after turn on, on mean value and standard deviation (absolute value of slope) for **BVA SC** for 80% of the batch

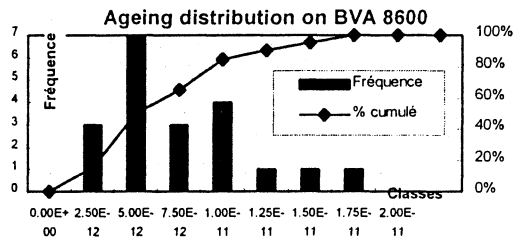
	8607
δf/f per day	3.5 10^{-12} 1σ=2 10^{-12}

The worst cases are + 6.39 and -5.76 10^{-12} par day.

Figure **1** gives typical ageing curve of BVA SC 5 MHz 8607 oscillator. One can see the stabilisation period of time, and the datas are collected at the end of this time.

For comparison , the following figure **2** gives the behaviour of a BVA AT batch of oscillators :

figure 2



These results, even if there are more AT BVA than SC BVA manufactured to day, show that the ageing is **slightly better with SC cut BVA than with AT BVA** resonators (80% lower than +/- 6 e-12 compared to 65% lower than +/-7.5 e-12 per day).

6-b) Phase noise

Figure **3 (a,b,c,d)** gives typical phase noise measurements obtained, with 860X and 869X oscillators, using AT or SC BVA resonators.

The following table gives the phase noise of 5 MHz BVA oscillators equipped with BVA AT and BVA SC resonators. Please note that a 3 dB reduction has been taken into account because of similarity of oscillators and because of the uniformity of results between pieces within a batch.

In dBc/Hz	8695	8600	8692	8607
crystal cut	*AT*	*AT*	*SC*	*SC*
Osc type	*1*	*2*	*1*	*2*
$\mathcal{L}$(f) @ 1 Hz	-120	-125	-120	-133
$\mathcal{L}$(f) @ 10 Hz	-140	-145	-140	-150
$\mathcal{L}$(f) @ 100 Hz	-148	-155	-150	-156
$\mathcal{L}$(f) @ 1 kHz	-155	-155	-155	-158
$\mathcal{L}$(f) @ 10 kHz	-155	-155	-155	-158

A major difference is observed, which will find an explanation in the oscillator design and topology.

For comparison purpose, Figure **4** gives phase noise measurements performed on an Europtest® test bench. Two 8607 oscillators are used, one under test, the other one as the reference. -134 dBc/Hz are measured

6-c) Short term stability (Allan variance)

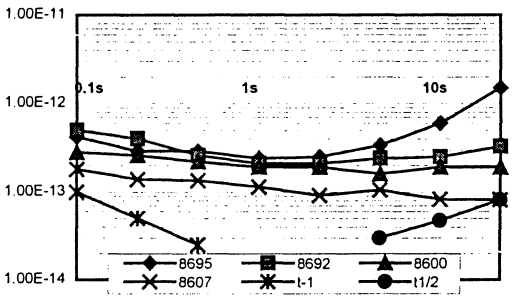
The following tables gives typical values of the short term stability (Allan variance) of pairs of oscillators of 8600 type or 869X type, with BVA AT and SC resonators. The straightforward results are given here without 3 dB reduction.

	8695	8692	8600	8607
crystal cut	AT	SC	AT	SC
osc type	1	1	2	2
$\sigma_y(\tau = 0.1\ s)$	4 10^{-13}	5 10^{-13}	2.7 10^{-13}	2 10^{-13}
$\sigma_y(\tau = 1\ s)$	2.4$10^{-13}$	1.5 10^{-13}	1.8 10^{-13}	1.1 10^{-13}
$\sigma_y(\tau = 10\ s)$	6 10^{-13}	1.6 10^{-13}	1.9 10^{-13}	9 10^{-14}
$\sigma_y(\tau = 30\ s)$	1.5 10^{-12}	1.8 10^{-13}	1.9 10^{-13}	9 10^{-14}
$\sigma_y(\tau = 100\ s)$		1.7 10^{-13}	2.4 10^{-13}	
batch size	*typ stat*	*1pair*	*typ stat*	*15 pces*

The environmental contributions, mainly temperature, humidity are observed for large tau values, in 8692 and 8695 oscillators, which use a simpler oven than the 860X family. This environmental contribution if larger for AT BVA than for SC BVA.

The next figure compare the short term stability results given in the different configuration studied :

figure 5



Curves in τ^{-1} and $\tau^{1/2}$ are plotted for slope comparison . One can see the different effects which have been discussed :

* environmental sensitivity ($\tau^{1/2}$) is observed for 8695 AT, the most sensitive, either for crystal or for oven
* lowest environmental sensitivity with SC than with AT in 869X
* Dewar oven is more stable than classical
* lowest intrinsic noise of SC
* τ^{-1} is not really seen on these oscillators between 1 - 10s : this can be due to the method of measurement (bench tau loop)

One can point out the 8607 results. The frequency stability and phase noise measurement appear to be well correlated, either on the flicker floor of frequency or on the white of phase ($\sigma_y(\tau)$ / 7 10^{-14} # -130 dBc/Hz @ 1 Hz). A slight contribution of 1/f of phase is observed around 10 Hz. The loaded Q_L cut off frequency is evaluated around 5Hz.

The following tables gives the short term stability of pairs of 8607 oscillators. The results are given here without 3 dB reduction.

	8607	8607	8607
pair of osc	A/C	A/B	B/C
$\sigma_y(\tau = 0.1\ s)$	2 10^{-13}	2.4 10^{-13}	1.8 10^{-13}
$\sigma_y(\tau = 1\ s)$	1.1 10^{-13}	1.3 10^{-13}	1.1 10^{-13}
$\sigma_y(\tau = 10\ s)$	9.8 10^{-14}	9.8 10^{-14}	8.25 10^{-14}
$\sigma_y(\tau = 30\ s)$	1.05 10^{-13}	1.05 10^{-13}	8.17 10^{-14}

The measurements reported here are obtained by the analysis of the dc voltage fluctuation in a frequency locking loop, converted through the phase comparator and the frequency-Vc coefficients. The noise of the bench have been evaluated as $1.510^{-14}/\tau + 5\ 10^{-14}$.

If we allow a $\sqrt{2}$ (3 dB « equi-weight ») reduction, the intrinsic short term stability of these oscillators can be *evaluated* as :

	8607	
	A or C	B or C
$\sigma_y(\tau = 0.1\ s)$	$1.4\ 10^{-13}$	$1.3\ 10^{-13}$
$\sigma_y(\tau = 1\ s)$	$7.7\ 10^{-14}$	$7.7\ 10^{-14}$
$\sigma_y(\tau = 10\ s)$	$6.9\ 10^{-14}$	$5.8\ 10^{-14}$
$\sigma_y(\tau = 30\ s)$	$7.4\ 10^{-14}$	$5.8\ 10^{-14}$

Because all of these measurements were not simultaneously obtained, we will not derive the « triangulation » results.

Some of these 8607 oscillators are currently under measurement using different conditions :

against a H_2 maser
under special environment protection

From these results, the 8607 oscillators are evaluated in the range of ***6 10^{-14}*** *for τ between 10 and 30s. Furthermore, it seems that the* ***1 10^{-13}*** *barrier for τ = 1 s is already achieved .*

These are probably the best result ever measured on an quartz oscillator industrially produced

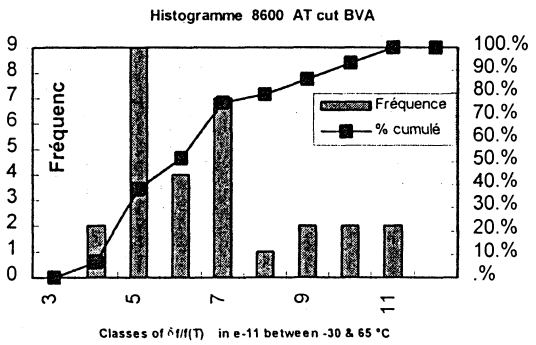
6-d) Frequency stability versus temperature

Typical results of frequency vs. Temperature drift (peak to peak) of The 869X and 860X oscillators, equipped with 5 MHz SC or AT Cut BVA resonators, are given in the following table :

	8695	8600	8692	8607
crystal resonator	AT	AT	SC	SC
oven (Dewar / Std)	S	D	S	D
$\delta f/f$ within -30 +60 °C	$4\ 10^{-10}$	$1\ 10^{-10}$	$8\ 10^{-11}$	$6\ 10^{-11}$

For comparison, the following figure gives the statistical behaviour of frequency vs. Temperature of AT BVA 8600 Oscillators.

Figure 6



The influence of the Dewar Oven and the use of SC vs. AT cut are reflected in the oscillator characteristics, the most sophisticated configuration (Dewar and SC) allows frequency stability of 6 10^{-11} in the temperature range.

6-e) G-sensitivity

Figure **7-a,b** give the 2G rotating test results for 2 samples of 8600 (AT) and 8607 (SC).

BVA SC confirms to be roughly 5 times better than BVA AT..

7/ Discussion on results

We can now compare AT and SC results, 5 MHz BVA, in both kinds of oscillators :

The following table gives this comparison on $\sigma_y(\tau)$ @ τ =10 s:

	AT	SC
8600-8607	$1.9\ 10^{-13}$	$9\ 10^{-14}$
8695-8692	$6\ 10^{-13}$	$1.6\ 10^{-13}$

and the next one gives this comparison on $\mathcal{L}(f)$ @ 1 Hz :

	AT	SC
8600-8607	-125	-133
8695-8692	-120	--120

A phase noise and short term stability fit have been tried either on 8607 or 8692 oscillators, with a classical power law model of frequency fluctuations :

$$S_y(f) = \Sigma_i\ h_i\ f^i \qquad (i : -2,+2)$$

restricted to i =2 (white of phase - external noise), i = 1 (1/f of phase - electronic origin) and i = -1 (1/f 'flicker floor' of frequency - resonator origin), using the usual relationships between phase and stability.

The next figures give a curve fit of the short term and of the noise behaviour of 8607 and 8692 oscillators. The three coefficients are adjusted to describe simultaneously the τ^{-1}, τ^0 and the f^3, f^1, f^0 areas.

For 8607, the fit is given by :

figure 8 a : stability 8607 BVA SC

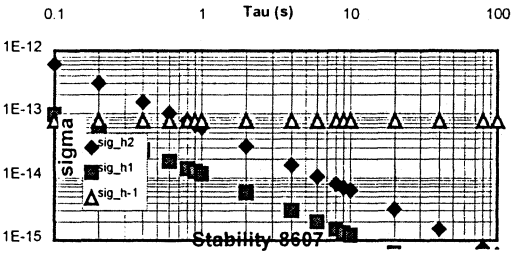
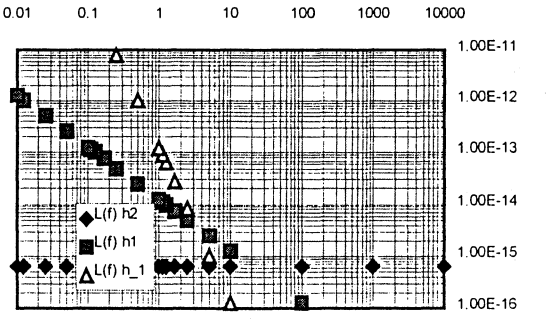


figure 8 b : phase noise 8607 BVA SC



For 8692, the curve fit is :

figure 8 c : stability 8692 BVA SC

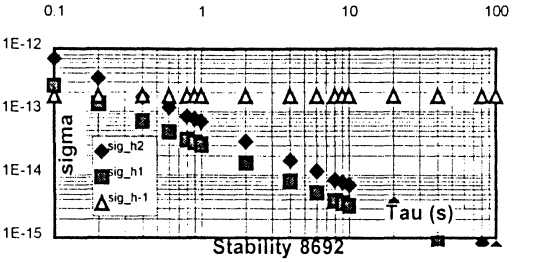
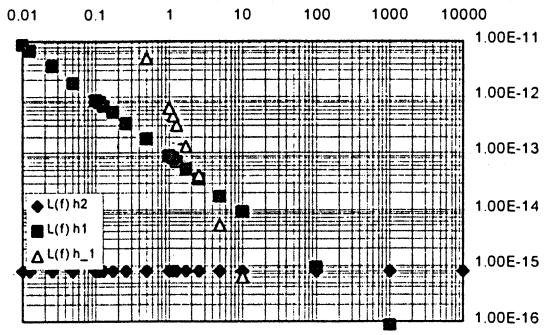


figure 8 d : phase noise 8692 BVA SC



The next table gives the fit coefficients for both oscillators :

	8607	8692
h_2	$5\ 10^{-29}$	$6\ 10^{-29}$
h_1	$1\ 10^{-27}$	$7\ 10^{-27}$
h_{-1}	$1\ 10^{-26}$	$4\ 10^{-26}$

The h_{-1}, (resonator related), reflects the crystal sorting, the h_2 value (external), appears to give similar results, not build the same way in both, but the large discrepancy in h_1 (oscillator loop) reveals the difference between 8607 and 8692 oscillators.

Identification and reduction of noise level in such oscillators will take into account different topics which has been described in the literature.

At the resonator level, the contribution of surface irregularity (scratch, chemical etch pits, ...) on drive level[22], and possibly on intrinsic noise[37], along with investigation on residual surface stresses and noisy contact [37] are all phenomenon involved in crystal noise. Amplitude frequency effect, even if it has been said that it is not the actual limitation of short term stability for well designed resonators and oscillators[38], can exhibit anomalies which contributes to frequency instabilities. Further more, so called frequency jumps cover various kind of anomalies (see annex and references therein) which leads to frequency instabilities.

Overall design contribution to frequency stability has been developed in different ways. The recent model introduced by F Walls[34,36] :

$$S_y(f) = \beta\ V_o\ /\ Q^4$$

gives a relation between stability and the active acoustic vibrating area. An important matching critical parameters is the loaded crystal Q_L in the oscillator (flicker floor given by [34,35] : $\sigma y(\tau) = (\ln 2\ A_{-1})^{1/2}\ /\ Q_L$), where A_{-1} is the 1/f coefficient of sustaining amplifier of the oscillator. The thermal design contribution has also been evaluated through the dynamic thermal coefficient of crystal resonators. Being in the 10^{-14} range require a thermal stability of 3nK/s with AT resonators, and only 30 nK/s with SC[34]. High thermal gain, high insulation and high mass has been described as key oven parameters.

Internal amplifier contribution to frequency instabilities has also been evaluated[34,35]. By mixing simple relation between phase and frequency of a resonator ($\Delta\nu/\nu = -\ \Delta\Phi\ /\ 2\ Q_L$), assuming a simple model of phase fluctuation of the sustaining amplifier $S_\Phi^{amp}(f) = 2\ (A_{-1}/f + A_0)$ where A_{-1} is called α_E and $A_0 = 2kFGT/P$ in a Leeson model, and assuming a crystal PM noise $S_\Phi^R = \alpha_R/f + 2kT/P$, the phase fluctuation density of the oscillator has been calculated. Figure 4 in reference 35 indicates the crystal loaded Q_L and amplifier flicker phase noise level L(f=1Hz) relation to be obtained to reach a flicker floor of frequency in the range of some 10^{-14}. The given example is an amplifier phase noise L(f=1Hz) = -148 dBc/Hz for a $Q_L = 0.65\ 10^6$ to reach $5\ 10^{-14}$. This literature help us to identify the stage we have reached and the direction we have to explore for further improvements.

The behaviour of the oscillators we have compare are understandable through

- the intrinsic noise lower in BVA SC than in BVA AT
- lowest motional resistance in BVA SC

- the design differences between 8607 and 869X
  - . loaded Q higher in 8607
  - . injection locking loop impedance higher
  - . crystal drive level higher in 8607
  - . different AGC configuration
  - . the lowest AM noise in 8607
  - . lowest parametric noise in 8607, due to Dewar oven and lower DC injected noise

One can also point out that the very high load insulation in 8607 can induce some additional external noise.

Further steps will act on the resonator (material, surface irregularities, acoustic volume...) and on the oscillator (noise selection, loaded Q...) and on the intrinsic stability of the oven. Better ageing seems to be accessible by further improvement on the material and process involved

Conclusion

We have presented results of 5 MHz BVA resonators and oscillators.

Two main results are obtained :

First, a comparison is achieved between various configuration of oscillators :
- 2 oscillator circuit types in two different oven structures
- BVA resonators AT and SC at the same frequency, 5 MHz

This work is far from its end, since the results in areas of investigation such as :
- high Q - low dislocation raw material
- improved resonator process and resonator design
- drive level effect
- electronic components sensitivity (and optimisation),

were not yet obtained, while under progress.

Nevertheless, we are able identify the various contributions of noise in these oscillators, determine their origins and then try to reduce them

Among these configurations the results today show that the use of BVA SC cut 5 MHz resonators in state of the art oscillators such as the 8600 concept, allows performances which were, till now, from the area of sophisticated laboratories or from the dream side...

The results presented here are, from what I know, the best industrial oscillator available on the open market. It probably confirms that the choice we made to stay at 5 MHz, is the right one, as far as the intrinsic resonator Q is concerned, as pointed out by J Norton[1].

Short term stability lower than $1\ 10^{-13}$ for τ = 1s, and around 6 to 8 parts in 10^{-14} are available with reproducibility and consistency good enough to speak of a « commercial » product. Phase noise @ 1 Hz below -130 dBc/Hz are also commercially available, even if some further improvements seem to be achievable.

Further step or actual goal is an ageing rate below $1\ 10^{-9}$ per year, to meet the $1.6\ 10^{-8}$ / 20 years target of telecom application,. The actual definition of such an oscillator (i.e. $3\ 10^{-9}$ / year, pulling range of +/- $1.5\ 10^{-7}$, linear slope of voltage control to get a stable PLL operation during the full life time, thermal stability below $2\ 10^{-12}$ / °C) allows to meet the hold over mode requirements and the long term stability one, without any change of the active device during the 20 years of usability. The only limit will come from the MTBF(calculated # 20 years for 8600 for oven operation around 80 °C), which is much more higher than in high temperature electronics (typ. 120 °C), as the one used in Rubidium clocks.

Bibliography

1/ R Besson - Proc. 31st Ann. Symp. On Frequency control - 1977
2/ J R Norton, JM Cloeren and PG Sulzer - Proc. 50th IEEE Freq. Control Symp Ann. Symp. On Frequency control - 1995
3/ JR Norton Proc 8th European Frequency and Time Forum-1994
4/ B.Penrod - Proc. 50th IEEE Freq. Control Symp - 1996
5/ HE Peters et al. Proc. 50th IEEE Freq. Control - 1996
6/ L Maleki Proc 9th European Frequency and Time Forum-1995
7/ G Santarelli et al. Proc. 49th IEEE Freq. Control - 1995 and Proc 8th European Frequency and Time Forum-1994
8/ S.Ghezali et al. Proc 9th European Frequency and Time Forum-1995
9/ E. de Clerq et al. Proc 8th European Frequency and Time Forum-1994
10/ DG Blair et al. Proc 8th European Frequency and Time Forum-1994
11/ M.Brunet et al. Proc. 50th IEEE Freq. Control Symp - 1996
12/ R Brendel Proc. 3rd European Frequency and Time Forum-1989
13/ C :Poignon et al. Proc of the 9th European Frequency and Time Forum-1995
14/ GH Frishat Berichte der Deutsh Keram Ges. 47 1970
15/ JP Bacheimer Proc of the 9th European Frequency and Time Forum-1995
16/ A el Habti et al. Proc of the 7th European Frequency and Time Forum-1993
17/ JP Aubry et al. Proc of the 7th European Frequency and Time Forum-1993
18/ PE Cash et al., Proc. Of the 50th IEEE Freq. Control Symp - 1996
19/ LD Vittorini & B Robinson, Proc. Of the 50th IEEE Freq. Control Symp - 1996 :
20/ R Brendel Proc of the 10th European Frequency and Time Forum-1996
21/ R Bourquin et al. Proc of the 10th European Frequency and Time Forum-1996
22/ E Eernisse Proc. Of the 50th IEEE Freq. Control Symp - 1996
23/ E EerNisse Proc. Of the 43rd Ann Frequency Control Symp. - 1989
24/ Deyzac Proc of the 7th European Frequency Time Forum 1993
25/ J.S Yang & H.F. Tiersten, Proc of the IEEE Freq. Control Symp - 1995
26/ J Detaint et al., Proc of IEEE Freq. Control Symp - 1996
27/ G Jonhson Proc IEEE Freq. Control Symp 1995
28/ SICN Commercial production of Low OH & private communication on low dislocation-low OH
29/G Demazeau LCS Bordeaux Private communication
30/ R Besson - Séminaire sur les Etalons de fréquence - LPMO-1981
31/ R Brendel et al ,Proc of IEEE Freq. Control Symp 1994
32/ N Ratier et al. Proc of the 9th European Frequency and Time Forum-1995
33/ R Brendel et al. These proceedings
34/ F.L. Walls, Proc of the 9th European Frequency and time Forum - 1995
35/ J Groslambert, Proc. of the IEEE Freq. Control Symp 1996
36/F.L. Walls et al., Proc. of the IEEE Freq. Control Symp.-1996
37/ E.S. Ferre-Pikal et al., Proc of the IEEE Freq. Control Symp 1996
38/ R Besson et al., Proc of the IEEE Freq. Control Symp 1996

Figures not in text

figure 1 : ageing characteristic of a BVA SC

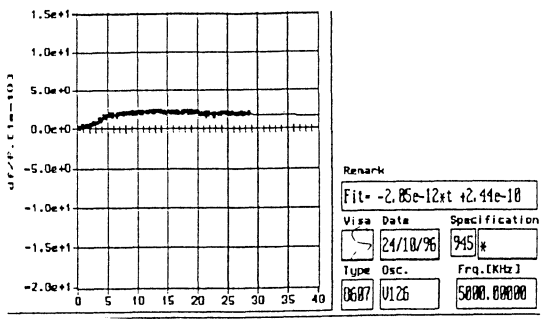


figure 3-a : Phase noise of 8607 BVA SC

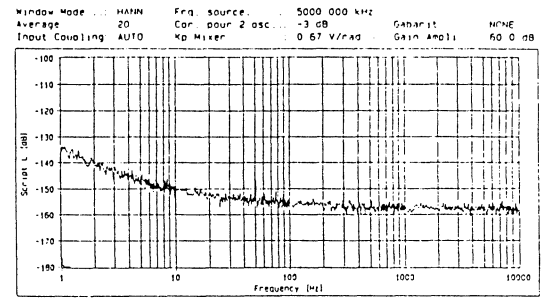


figure 3b : Phase noise of 8600 BVA AT

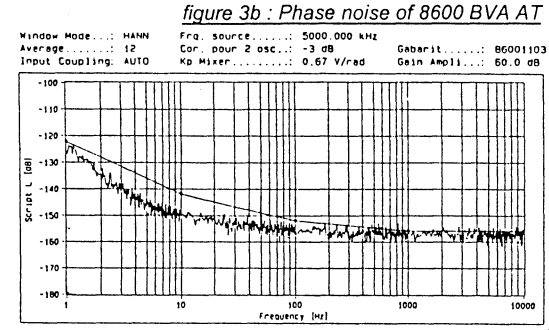


figure 3-c : Phase noise of 8692 BVA SC

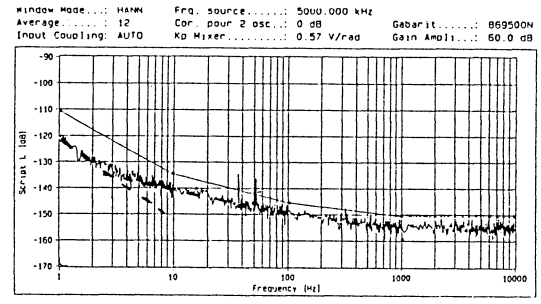


figure 3-d : Phase noise of 8695 BVA AT

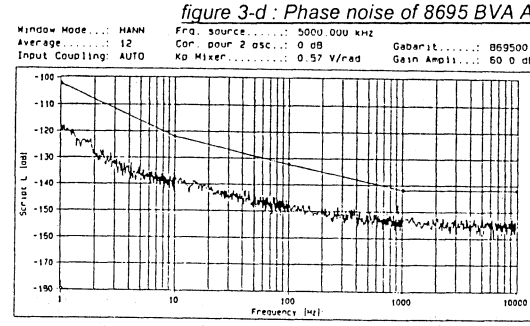


figure 4 : Phase noise of 8607 BVA SC on Europtest® bench

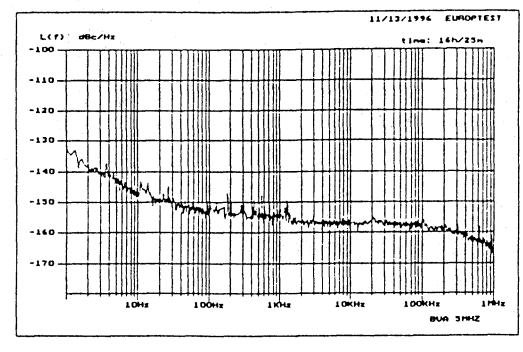


figure 7-a : G sensitivity on 8600 BVA AT

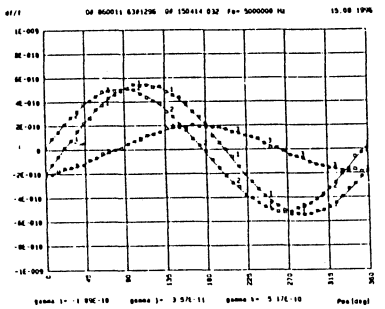
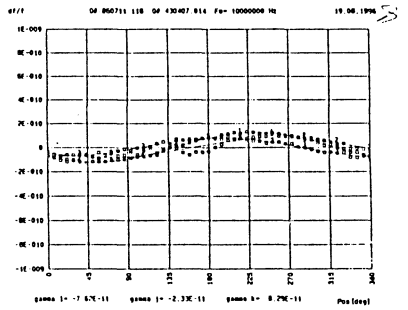


figure 7-b G sensitivity on 8607 BVA SC



Annex
FREQUENCY JUMPS ON BVA RESONATORS

Kanak K. TULADHAR & Gerald JENNI

Frequency Jump phenomena have been recently introduced, classified and described for a number of different types of quartz crystal resonators (refs. A1, A2). One interesting type of observed frequency jumps on BVA resonators has been detailed in fig. 11 of ref. A3. These frequency jumps are in the range up to 4 E - 10 and can be classified as discrete frequency jumps (ref. A1). They are mainly caused due to a variation in gap width between the condensers (C_{g1} and/or C_{g2} ; note that C_g of fig. 2, ref. A4, corresponds to the parallel combination of C_{g1} and C_{g2}) and resonator of the BVA assembly.

Conventional BVA sandwich assembly details are given in figs. 1 to 3 of ref. A5. One of the main problems encountered during assembly is due to the condenser electrode tabs, these in turn, firstly makes gap widths asymmetric on both sides of the resonators and secondly, again due to these tabs, resonator sandwich assembly produces undue stresses and strain on the condensers and resonator. During stress relaxation of these condensers, minute changes in gap widths take place and hence a discrete frequency jump will be produced.

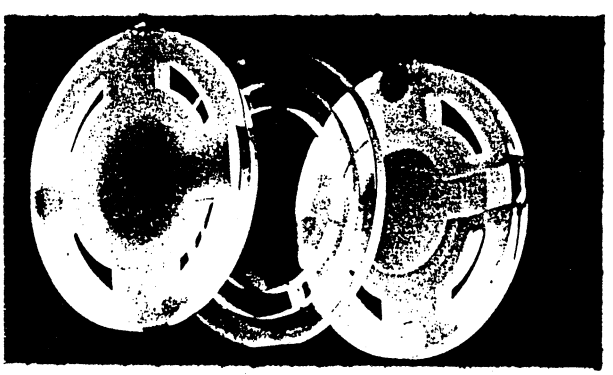
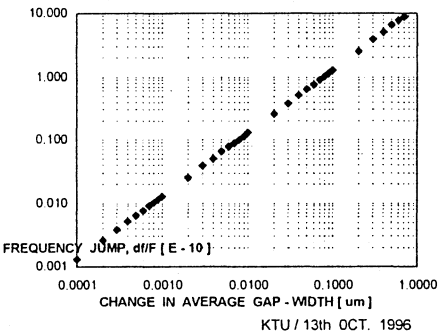


Fig. A1. BVA Resonator and new type of Condensers having three extra spots for improved performance

This problem can be eliminated by introducing three additional spots(tabs) on the condensers. These extra tabs can be deposited easily during vacuum deposition of condenser electrodes. The resulting condensers are shown in fig. A1. Note also that this modification further ruggedised the BVA assembly. Further more this newly introduced condensers also makes vacuum baking process more efficient and preliminary results have indicated a superior ageing behaviour relative to the conventional resonators.

Now consider a hypothetical fifth overtone mode BVA resonator at 5.000 MHz having motional capacitance, C_1 of 0.2 fF & static capacitance, C_0 of 5.0 pF . And assuming condenser electrode diameter of 10 mm and nominal gap - width of 10 µm between the condensers and the resonator. Frequency changes i.e. jumps due to changes in one of the gap - widths in the range 9 µm to 9.9999 µm from nominal value of 10 µm have been computed and given in fig. A2. This clearly demonstrate that frequency jump of 1 E - 11 is produced due to a relative change in gap - width by 0.008 µm (i.e. 10.000 µm to 9.992 µm). Note also that this computation was done after the suggestion made by Dr. A. Ballato (ref. A6).

Fig. A2. COMPUTED FREQUENCY JUMPS FOR 5.000 MHz BVA RESONATOR



REFERENCES

A1. Tuladhar, K.K, and Jenni, G, Proc. 10th EFTF , March 1996, pp. 111 - 115.
A2. Tuladhar, K.K, and Jenni, G, Proc. 50th AFCS, June 1996, pp. 339 - 345.
A3. Brunet , Proc. 6th EFTF, March 1992, pp. 125 - 136.
A4. Besson, R Proc. 30th AFCS, June 1976, pp. 78 - 83.
A5. Graf, E. P. and Peier U. R. Proc. 37th AFCS, June 1983, pp. 492 - 500.
A6. Ballato, A. " Private Discussion with one of the Authors(KKT) at the 10th EFTF " , Brighton , England, March 1996.

TIME, SPACE, AND RELATIVITY

Chairman: P. Kartaschoff

ACCURACY OF INTERNATIONAL ATOMIC TIME TAI

Claudine Thomas

Bureau International des Poids et Mesures

ABSTRACT

The accuracy of the International Atomic Time, TAI, is estimated by the relative difference, together with its uncertainty, between the duration of the TAI scale interval and the SI second as produced on the rotating geoid by primary frequency standards. The BIPM regularly receives measurements from a number of accurate primary frequency standards, among them: the BNM-LPTF caesium fountain FO1, the optically pumped caesium standard NIST-7 and the classical standards PTB CS2, PTB CS3 and SU MCsR 102. Each individual measurement, corrected for all known corrections, provides an estimation of the accuracy of TAI with an uncertainty equal, in most cases, to the frequency uncertainty of the standard. The individual measurements can also be treated in a global way in order to improve the knowledge of the accuracy of TAI. The relative departure of the TAI scale unit from the SI second was about $2{,}0\times10^{-14}$ for the period 1994-1996 and was known with uncertainties ranging from $0{,}5\times10^{-14}$ to $1{,}0\times10^{-14}$.

Keywords: Primary frequency standard, International Atomic Time, Accuracy.

1. INTRODUCTION

From a declaration of the Comité Consultatif pour la Définition de la Seconde, CCDS, in 1980 (Ref. 1) International Atomic Time, TAI, is

a coordinate time scale defined in a geocentric reference frame with the SI second as realized on the rotating geoid as the scale unit.

Since 1 January 1988, the Bureau International des Poids et Mesures, BIPM, has been responsible for the generation and dissemination of TAI, which is a world-wide reference time scale and thus should be as reliable, stable and accurate as possible.

TAI relies on measurements taken on a regular basis from commercial atomic clocks and primary frequency standards maintained in national timing centres spread world-wide. Reliability and stability are optimized in the first step of the procedure used for combining these data. In this, a free atomic time scale, EAL (échelle atomique libre), is computed as a weighted average of a large number of free running and independent atomic clocks (Ref. 2). The weight attributed to a given contributing clock increases with its stability which makes it possible to improve the stability of the resulting average. However, no attempt is made in this first step to ensure the conformity of the EAL scale unit with the SI second as realized on the rotating geoid.

In a second step, the duration of the scale unit of EAL is evaluated by comparison with the data of primary caesium standards which realize the SI second. TAI is then derived from EAL by adding a linear function of time with a convenient slope to bring the TAI scale unit close to the SI second on the surface of the rotating geoid. The frequency offset between TAI and EAL is changed when necessary to maintain accuracy (Ref. 3), the magnitude of the changes being of the same order as the frequency fluctuations resulting from the instability of EAL. This operation is referred to as the 'steering of TAI'.

During the three year period 1994-1996 the quality of the timing data used for TAI computation has rapidly evolved. One major change has been the extensive replacement of older designs of commercial clocks by the new HP 5071A clocks which show outstanding stability. This naturally led to an improvement of the stability of EAL estimated in 1996 at 2 parts in 10^{15}, in terms of Allan standard deviation, for averaging times between 1 and 3 months (Ref. 2). Such a level of stability limits the amplitude of the steering corrections to values less than 2×10^{-15}.

Another important point concerns the question of the black-body frequency shift experienced by primary frequency standards. Under the influence of the radiation from the walls surrounding the atoms inside the clock at temperature T, the clock transition frequency is reduced with respect to its value at $T = 0$ K. From the formula provided by Itano *et al.* (Ref. 4), the amplitude of the effect is $1{,}7\times10^{-14}$ for $T =$ 300 K, with an uncertainty conservatively estimated at 1×10^{-15}. At its 13th meeting, held in March 1996, the CCDS recommended that a correction by which the clock frequency is transferred at 0 K, i.e. compensating for the black-body frequency shift, should be applied to all primary standards measurements (Ref. 5). It follows that the duration of the SI second on the rotating geoid, as realized by primary frequency standards, abruptly became larger than before by about 2×10^{-14} s.

The situation also changed with the appearance of new primary frequency standards, such as the optically pumped frequency standard NIST-7 (Ref. 6), developed at the NIST, Boulder, Colorado, USA, and the caesium fountain FO1 (Ref. 7), in operation at the BNM-LPTF, Paris, France. These novel standards show an

outstanding accuracy with Type B uncertainties ranging in the low 10^{-15}.

Altogether these different elements have led us to question our knowledge of the accuracy of TAI and to ask how to improve it. These are the topics addressed in the following sections.

2. ESTIMATING TAI ACCURACY

From its definition, the duration of the TAI scale unit, u, should be as close as possible to the duration of the SI second, u_0, on the rotating geoid. The accuracy of TAI may thus be characterized by the relative departure, d, defined as:

$$d = \frac{u - u_0}{u_0}, \qquad (1)$$

and its uncertainty σ.
In (1) u is expressed in seconds and u_0 = 1 s is the period of the f_0 = 1 Hz signal provided by a primary frequency standard, after all frequency corrections, especially those compensating for the black-body shift and the gravitational red shift, have been applied. In the same way, u is the period of the frequency f (very close to 1 Hz) of TAI so d can be written as:

$$d \approx -\frac{f - f_0}{f_0}, \qquad (2)$$

and is close to the negative of the relative frequency departure (designed as 'rate') of TAI relative to a primary frequency standard.

The frequency f_0 = 1 Hz is delivered by standard i with a relative uncertainty σ_{Bi} in the laboratory where it operates (the so-called Type B uncertainty), but access to the quantity $f - f_0$ is not immediate. It goes through two frequency comparisons:

* First, one compares the frequency of standard i with the frequency of a local oscillator, often a hydrogen maser, over a given time interval designed as the calibration interval. The calibration interval, I_{Ci}, is defined by its length τ_{Ci} and its central date t_{Ci}.
This first step gives the rate of the local oscillator relative to standard i over I_{Ci}. Its uncertainty includes several components (see Table 1): the Type B uncertainty of standard i, σ_{Bi}, the Type B uncertainty of the local oscillator, $\sigma_{B,LO}$, the Type A uncertainties of standard i and of the local oscillator, $\sigma_{Ai}(\tau_{Ci})$ and $\sigma_{A,LO}(\tau_{Ci})$ which are their Allan standard deviations over τ_{Ci}, and the uncertainty of the local comparison, $\sigma_{LC}(\tau_{Ci})$ which is the frequency instability of the local transfer method over τ_{Ci}.
* Second, the frequency of the local oscillator is transferred to TAI over a given time interval designated as the estimation interval I_E, also defined by its length τ_E and its central date t_E. I_E depends only upon the dates of computation of TAI and thus does not include the index i. The shortest possible duration τ_E was 10 days until December 1995 and 5 days in 1996.
This second step gives the rate of TAI relative to the local oscillator over I_E. Its uncertainty is made up several components (see Table 1): the Type B uncertainty of the local oscillator, $\sigma_{B,LO}$, the Type A uncertainty of the local oscillator, $\sigma_{A,LO}(\tau_E)$, and the uncertainty σ_{TR} of the transfer method, frequency instability of the GPS common-view method over τ_E.

Table 1. Summary of the uncertainty components arising in the transfer of the frequency of a primary standard i to TAI.
LO, LC and TR stand for local oscillator, local comparison and GPS transfer.
Type A and Type B uncertainties are defined in *Guide to the Expression of Uncertainty in Measurement*, ISO 1993, item 2.3.1 page 3.

Interval	Measured quantity	Uncertainty components	
		Type B	Type A
I_{Ci} Length τ_{Ci} Cent date t_{Ci}	Rate of LO relative to standard i	σ_{Bi} $\sigma_{B,LO}$	$\sigma_{Ai}(\tau_{Ci})$ $\sigma_{A,LO}(\tau_{Ci})$ $\sigma_{LC}(\tau_{Ci})$
I_E Length τ_E Cent date t_E	Rate of TAI relative to LO	$\sigma_{B,LO}$	$\sigma_{A,LO}(\tau_E)$ $\sigma_{TR}(\tau_E)$

The sum of the rates obtained in these two steps gives access to a measurement d_i, valid over I_E, of d provided by standard i evaluated over I_{Ci}. The total uncertainty σ_i of d_i includes all components given in Table 1. They are not all independent. The systematic part of the uncertainty coming from the local oscillator, $\sigma_{B,LO}$, disappears in the difference (this quantity does not appear in what follows). A typical order of magnitude of $\sigma_{TR}(\tau_E)$ is 1×10^{-14} for τ_E = 1 day.

For most primary frequency standard measurements $I_E = I_{Ci}$, so the Type A uncertainties of the local oscillator disappear. It follows that only three components are actually important σ_{Bi}, $\sigma_{Ai}(\tau_{Ci})$ and $\sigma_{TR}(\tau_E)$. However, it sometimes happens that τ_{Ci} is shorter than the minimum 5 or 10 days of τ_E. It is then necessary to take into account the Type A uncertainty components arising from the instability of the local oscillator over τ_{Ci} and τ_E.

Each pair of numbers (d_i, σ_i) is an individual estimation of TAI accuracy. Examples covering the three year period 1994-1996 are given in Section 3.

The individual estimations can also be treated in a global way in order to deliver a more precise value of d for any interval I_E. This global treatment operates with preceding and following calibrations taking place over intervals I_{Ci} included or not in I_E, even partially overlapping with I_E, and for which $t_E - t_{Ci}$ may be positive or negative. It is thus necessary to transfer temporally, over $|t_E - t_{Ci}|$, the individual calibrations (d_i, σ_i). In the temporal transfer the values d_i are kept constant but the values σ_i are increased. In fact, the temporal transfer is carried out by the time scale itself, which, because it is continuous, can temporally link calibration and estimation intervals. Its instability over $|t_E - t_{Ci}|$ thus creates a new component of uncertainty.

After temporal transfer to the chosen interval I_E, calibrations are combined to obtain d and σ valid over a given I_E. In practice, this global treatment is not an easy task because the time scale is affected by different types of noise according to the length of time involved, and because the system is not stationary since the time scale stability improves with the passing of time. In addition, the combination of transferred calibrations should be optimized in order to obtain the best global estimate (with the smallest possible uncertainty). This problem was solved in 1977 by Azoubib, Granveaud and Guinot (Ref. 8) [The question of temporal transfer is also treated in Ref. 9]. Post-processing allows d to be estimated over a given I_E through a weighted average of all individual measurements which have occurred before or after t_E. The error made in the estimation process is given in Ref. 8 (page 89). It involves the weights and a number of other parameters:

* the uncertainty σ_i of each measurement d_i,
* a model for the stability of the time scale,
* the length of individual calibration intervals τ_{Ci},
* the length of the estimation interval τ_E,
* the length of the time interval $|t_E - t_{Ci}|$ separating the calibration intervals and the estimation interval.

Minimizing this error makes it possible to determine the weights (the sum of which is equal to 1) and to get σ which has the value corresponding to the computed minimum.

The method described in Ref. 8 is applied at the BIPM and provides the regular estimations of d and σ which are published in successive issues of *Circular T* and of the *Annual Report of the BIPM Time Section*. We give results of this treatment for the years 1994-1996 in Section 4 of this paper.

3. INDIVIDUAL ESTIMATIONS OF TAI ACCURACY

Over the period 1994-1996, the BIPM received measurements from six very accurate primary frequency standards, several of them being of different types. They are listed in Table 2 (meaning of acronyms given at the end of the text). The measurements provided by these six primary frequency standards are corrected uniformly for all known corrections, especially the gravitational frequency shift and the black-body radiation frequency shift, so they are fully coherent and can be compared over the complete three year period of this study.

Table 2. Characteristics of the primary frequency standards which provided measurements to the BIPM over the three years 1994-1996.

*Uncertainty of $0{,}5\times10^{-14}$ starting April 1996.

**PTB CS1 ceased activity at mid-1995.

Laboratory	Standard i	Type B uncertainty σ_{Bi}	Standard type
BNM-LPTF	FO1	$0{,}3\times10^{-14}$	Fountain
NIST	NIST-7	$1{,}0\times10^{-14}$ *	Optical Pumping
PTB	CS1**	$3{,}0\times10^{-14}$	Classical
PTB	CS2	$1{,}5\times10^{-14}$	Classical
PTB	CS3	$1{,}4\times10^{-14}$	Classical
SU	MCsR 102	$5{,}0\times10^{-14}$	Classical

For the classical standards PTB CS1, CS2, CS3 and SU MCsR 102, the intervals of calibration I_{Ci} are equal to the intervals of estimation I_E. They are the usual consecutive two-month intervals defined for TAI computation, January-February, March-April, ...*etc*, (Ref. 2) included in the period 1994-1996. The uncertainty σ_{TR} of GPS transfer is smaller than 1 part in 10^{15} over τ_E = two months and thus negligible compared to σ_{Bi} and σ_{Ai} for the four standards (see Table 3).

For PTB CS2, there is no local oscillator since it directly provides the input of the GPS receiver in operation at the PTB and $\sigma_{A,CS2}(\tau_{C,CS2}$ = 2 months) is less than 7×10^{-15} for the three years under study (see Table 3). It follows that PTB CS2 measurements provide direct estimations d_{CS2} with uncertainties $\sigma_{CS2}\approx\sigma_{B,CS2}$.

For PTB CS1 and CS3, the local oscillator is PTB CS2 to which they are internally compared with an uncertainty which is completely negligible. For averaging times of two-months the Allan standard deviations of PTB CS1 and PTB CS3 are respectively estimated at 1×10^{-14} and 5×10^{-14} (conservative value for PTB CS3 due the observation of frequency steps). It follows that $\sigma_{CS1}\approx\sigma_{B,CS1}$, but $\sigma_{CS3} \gg \sigma_{B,CS3}$.

For averaging times of two-months the Allan standard deviation of SU MCsR 102 is estimated to be 1×10^{-14} and the local oscillator is a hydrogen maser to which the standard is compared with an uncertainty which is completely negligible. It follows that $\sigma_{MCsR}\approx\sigma_{B,MCsR}$.

The NIST provides measurements of the rate of NIST-7 relative to their local hydrogen maser over calibration intervals of a few days. The dates of occurrence of the calibration intervals are carefully chosen at the NIST so that they nearly match the standard dates of TAI computations (each 5 days in 1996 and 10 days in 1994 and 1995). The instability of their local hydrogen maser thus has very little influence on the transfer to TAI. The uncertainty of GPS transfer can be estimated at $4x10^{-15}$ for averaging times of 5 days and $1x10^{-15}$ for averaging times of 10 days, and cannot always be considered as negligible when compared with the Type B uncertainty of NIST-7. The stability of the standard itself is less than $2x10^{-15}$ over 5 and 10 days (see Table 3). Measurements from NIST-7 have been sent regularly (each one or two months) to the BIPM starting mid-1994. Unfortunately, some of these calibrations were withdrawn: the six first points, covering the first nine months of operation of the standards, were declared wrong (Ref. 10) and there is some suspicion that the last three points for 1996 may be subject to a systematic error of $1,5x10^{-14}$. These nine evaluations have been withdrawn from the BIPM records and are not included in the calculations.

The calibrations provided by LPTF-FO1 are different from those from other primary standards because the length of the calibration intervals is generally 10 hours, the averaging time over which the fountain becomes more stable than the local hydrogen maser. The comparison between the fountain and the hydrogen maser is carried out with a very small uncertainty (see Table 3); however, the link to TAI is made with an additional step corresponding to the time comparison of the hydrogen maser to their local representation of UTC which serves as input to the GPS receiver. This artificially increases the uncertainty $\sigma_{TR}(\tau_E = 10$ or 5 d) to $1x10^{-14}$ (Ref. 11). In addition, since $\tau_{C,FO1} \ll \tau_E = 10$ or 5 d, the instability of the local hydrogen maser has an impact on the overall uncertainty σ_i, which, at the end, is much larger that the Type B uncertainty of the fountain.

The BIPM has received two sets of measurements from LPTF-FO1. The first covers the period September-December 1995 and the second, May 1996. During the first period the local hydrogen maser was globally much noisier than during the second (in fact it experienced a failure, which necessitated a repair, between them). We thus had little confidence in the stability of the local hydrogen maser. The transfer of the fountain measurements to TAI over that period was therefore carried out using another oscillator, the hydrogen maser used as master clock at the NPL, Teddington, United Kingdom. It was possible to compare the frequencies of the hydrogen masers, at the BNM-LPTF and the NPL, over durations of 1 d via GPS. In this case the BNM-LPTF hydrogen maser served only to ensure the transfer from 10 h to 1 day and the NPL hydrogen maser, which was less noisy, from 1 day to 10 days (minimum τ_E in September-December 1995).

Figure 1 shows the values of d_i obtained from the individual measurements provided over the period 1994-1996. All points are consistent taking account of their respective uncertainties as given in Table 3. Some remarks, however, are of interest:

* A systematic discrepancy is observable between PTB CS1 and PTB CS2 calibrations over their period of common operation.
* Points from PTB CS3 are higher than others and are coherent with them only if the stability of PTB CS3 is conservatively estimated at $5x10^{-14}$.
* Points from NIST-7 cover nearly half the period under study and are in good agreement with calibrations from PTB CS2.
* The first set of points from LPTF-FO1 shows a dispersion and gives values d_{FO1} which are smaller than those given by NIST-7, PTB CS2 and PTB CS3 while the second set of points is in good agreement with the other calibrations.

The simple average of these measurements over the period 1994-1996 gives a value of about $2x10^{-14}$ which means that the duration of the TAI scale unit is rather far from the SI second. This is easy to interpret. In the past, the only calibrations of the TAI frequency were provided by PTB CS1 and PTB CS2, and these were not corrected for the black-body radiation shift. Frequency steering corrections were applied to TAI in order to approximate the duration of its scale interval to the mean SI second provided by these two standards, which is smaller by about $5x10^{-15}$ than the SI second produced by PTB CS2. When the black-body correction was taken into account, all available measurements were shifted by another $1,7x10^{-14}$ giving a total difference close to that which appears here.

4. GLOBAL ESTIMATION OF TAI ACCURACY

Post-processing of the individual measurements shown in Section 3 was carried out according to the method described in Section 2.

The intervals of estimation I_E chosen for applying the global treatment described in Section 2, are the usual consecutive two-month intervals defined for TAI computation, January-February, March-April, *...etc*, (Ref. 2) in the period 1994-1996. The value $\tau_E = 2$ months was chosen because it corresponds to the averaging time for which TAI presents the best stability, and thus the duration on which the accuracy of TAI can be determined most precisely.

The time scale chosen to carry out the temporal transfer is EAL and its stability model includes three noise types: white frequency modulation (WFM), flicker frequency modulation (FFM) and random walk

frequency modulation (RWFM). For 1995 and 1996, the chosen noise levels are:

* $\sigma_{yWFM}(\tau = 1d) = 1{,}4x10^{-14}$,
* $\sigma_{yFFM} = 4{,}5x10^{-15}$,
* $\sigma_{yRWFM}(\tau = 1d) = 0{,}4x10^{-15}$.

For 1994, slightly higher values were chosen to take account of the natural improvement in the stability of TAI.

In the global treatment presented in this paper, all individual measurements are regarded as independent even if they are provided by the same continuously operating standard.

The estimations of d and σ over the period 1994-1996 are added in Fig. 1 (continuous line). They are always very close to $2x10^{-14}$ with a reduced uncertainty of $0{,}5x10^{-14}$ over time intervals when several points from NIST-7 or LPTF-FO1 are available. Compensation for the discrepancy induced by the uniform application of the black-body correction was initiated in March 1995: it takes the form of cumulative frequency steering corrections, applied on dates separated by two months intervals. Their individual relative amplitude was increased from $1x10^{-15}$ to $1{,}5x10^{-15}$ in August 1996. This procedure, however, has not succeeded in decreasing the offset between the duration of the TAI scale unit and the SI second on the rotating geoid. It has merely cancelled the natural drift of EAL, so it may become necessary to increase the relative amplitude of the frequency steering corrections at some time in the future.

An interesting point of this study concerns the weights of individual measurements in the global estimate. They are listed in Table 4 for each standard and for three different I_E intervals corresponding to March-April 1995, November-December 1995 (also shown in Fig. 2) and September-October 1996.

Table 4. Relative weight of each primary standard whose individual measurements have been used for the global estimate of d over three two month intervals.

Standard	Interval of estimation I_E		
	Mar-Apr 95	Nov-Dec 95	Sep-Oct 96
LPTF-FO1	3,1%	68,2%	22%
NIST-7	65,1%	19,8%	7,4%
PTB CS1	6,0%	0,2%	0%
PTB CS2	25,1%	10,3%	64,7%
PTB CS3	0%	0,7%	5,6%
SU McsR	0,7%	0,8%	0,3%

Theoretically, each time the BIPM receives a new measurement from a primary frequency standard, the estimations should be reprocessed for any time interval. Fig. 2, however, shows that the weights of those calibrations which occur about 9 months before or after t_E are smaller than 0,5% and thus have very little influence. It follows that there is no need to reprocess the data after 2 years (except when data are withdrawn after the event). Instead of including all calibrations in our global treatment it is thus sufficient to consider only measurements covering the 1½ year interval centred on the central date of I_E.

Another point is that measurements from very accurate primary frequency standards do not sweep along the estimations over years: the first set of LPTF-FO1 points (September-December 1995) has a total weight of only 3,1% on the estimate for March-April 1995 and 4,1% on the estimate for September-October 1996 (the other 17,9% quoted in Table 4 correspond to the 2nd set of LPTF-FO1 points in May 1996). The intrinsic stability of the time scale used for the temporal transfer does have a strong influence on the values σ_i which rapidly increase with time.

Fig. 2 also shows that those measurements taken at dates close to t_E and provided by the most accurate frequency standards have the largest weights in the estimation over I_E. For example LPTF-FO1 has a total weight of 68,2% for the estimate over November-December 1995. However, the average weight of the five LPTF-FO1 measurements actually included in this particular I_E is 11%, while the closest NIST-7 measurement has a weight of 10,5%. Although LPTF-FO1 is the most accurate primary frequency standard in the world, the uncertainty of its measurements is degraded by its transfer to TAI (local oscillator, local comparison, GPS transfer) with the effect that one 10 day calibration from LPTF-FO1 has no more weight than one simultaneous 10 day calibration from NIST-7.

At the end of 1996, the accuracy of TAI relied almost entirely on one single standard, PTB CS2 (total weight close to 65%). This situation is the worst we have experienced since years: even in 1993, we had two standards, PTB CS1 and PTB CS2. In addition this is inconsistent with Recommendation S1 (1996) of the Comité Consultatif pour la Définition de la Seconde (Ref. 12).

The consequence of all these different comments is that even if novel and very accurate primary frequency standards are operated and evaluated throughout the world, the accuracy of TAI can be established safely only if the measurements are sent regularly to the BIPM and if great caution is taken when making the transfer to TAI. Caution is required especially in the choice of the local oscillator and the choice of the calibration interval.

5. CONCLUSIONS

The accuracy of TAI is efficiently represented by a graph showing the individual measurements of its relative scale unit to the SI second as produced by

primary frequency standards, operated in various laboratories, after conversion to the rotating geoid, together with corresponding individual uncertainties. The global treatment of these individual measurements makes it possible to add to the graph a curve of the mean estimate together with its uncertainty. Such a graph demonstrates international equivalence between the primary frequency standards reported to the BIPM in the sense discussed at the meeting of directors of national metrology institutes, organized by the Comité International des Poids et Mesures on 17 and 18 February 1997.

Acknowledgements

Many thanks to Peter Wolf for helpful discussions.

Acronyms

BNM-LPTF	Bureau National de Métrologie, Laboratoire Primaire du Temps et des Fréquences, Paris, France
NIST	National Institute of Standards and Technology, Boulder, Colorado, USA
PTB	Physikalisch-Technische Bundesanstalt, Braunschweig, Germany
SU	VNIIFTRI, All Russian Research Institute for Physical, Technical and Radio-technical Measurements, Moscow, Russia

References

1. BIPM Com. Cons. Déf. Seconde, 1980, **9**, p S15.
2. Thomas C and Azoubib J, *Metrologia*, 1996, **33**, 227-240.
3. Differences between the normalized frequencies of EAL and TAI, 1996, *Annual Report of the BIPM Time Section*, 9, Table 5, p 35.
4. Itano W.M., Lewis L.L. and Wineland D.J., *Phys. Rev. A*, 1982, **25**, 1233-1235.
5. BIPM Com. Cons. Déf. Seconde, 1996, **13**, p S75.
6. Lee W.D., Shirley J.H., Lowe J.P., Drullinger R.E., *IEEE Trans. Instrum. Meas.*, 1995, **44**, 120-123.
7. Clairon A., Ghezali S., Santarelli G., Laurent P., Lea S.N., Bahoura M., Simon E., Weyers S. and Szymaniec K., Preliminary accuracy evaluation of a cesium fountain frequency standard, *Proc. 5th Symposium on Frequency Standards and Metrology*, Ed. James C. Berquist, World Scientific, 1995, 49-59.
8. Azoubib J, Granveaud M and Guinot B, Estimation of the scale unit duration of time scales, *Metrologia*, 1977, **13**, 87-93.
9. Douglas R.J., Boulanger J.-S, Cundy S., Gagné M.-C., Cazemier W., Hoger B., Pelletier R., Bernard J., Madej A.A., Marmet L., Siemsen K. and Whitford B.G., *Proc. 28th PTTI Meeting*, 1996, in press.
10. Bauch A., Heindorff T., Schröder R. and Fischer B., *Metrologia*, 1996, **33**, 249-259, reference 26 page 259.
11. A. Clairon, personal communication, January 1997.
12. BIPM Com. Cons. Déf. Seconde, 1996, **13**, p S74.

Table 3. Uncertainties involved in estimation of TAI accuracy from measurements provided by primary frequency standards.

All numbers given in this Table for Type A uncertainties are BIPM estimates deduced from literature or discussions with authors. All are subject to revision. Type B uncertainties are those of Table 2.

$\oplus$ stands for quadratic sum.

Standard i	I_{Ci}	I_E	$\sigma_{Ai}(\tau_{Ci})$	$\sigma_{A,LO}(\tau_{Ci})$	$\sigma_{LC}(\tau_{Ci})$	$\sigma_{A,LO}(\tau_E)$	$\sigma_{TR}(\tau_E)$	σ_i
LPTF-FO1								
1995	$\tau_{Ci} \approx 10$ h	$\tau_E = 10$ d	2×10^{-15}	2×10^{-15}	$\ll 1\times10^{-15}$	1×10^{-14}	1×10^{-14}	$\gg \sigma_{B,FO1}$
1996	$\tau_{Ci} \approx 10$ h	$\tau_E = 5$ d	1×10^{-15}	1×10^{-15}	$\ll 1\times10^{-15}$	1×10^{-14}	1×10^{-14}	$\gg \sigma_{B,FO1}$
NIST-7								
1994-1995	$\tau_{Ci} \approx 10$ d	$\tau_E = 10$ d	$< 2\times10^{-15}$	-	$\ll 1\times10^{-15}$	-	1×10^{-15}	$\sigma_{B,NIST-7}$
1996	$\tau_{Ci} \approx 5$ d	$\tau_E = 5$ d	$< 2\times10^{-15}$	-	$\ll 1\times10^{-15}$	-	4×10^{-15}	$\sigma_{B,NIST-7} \oplus \sigma_{TR}(\tau_E)$
PTB CS1	Usual TAI computation intervals		1×10^{-14}	-	$\ll 1\times10^{-15}$	-	$\ll 1\times10^{-15}$	$\sigma_{B,CS1}$
PTB CS2			$< 7\times10^{-15}$	-	-	-	$\ll 1\times10^{-15}$	$\sigma_{B,CS2}$
PTB CS3			5×10^{-14}	-	$\ll 1\times10^{-15}$	-	$\ll 1\times10^{-15}$	$\gg \sigma_{B,CS3}$
SU MCsR			1×10^{-14}	-	$\ll 1\times10^{-15}$	-	$\ll 1\times10^{-15}$	$\sigma_{B,MCsR}$

Figure 1. Individual values of d_i obtained from primary frequency standards over 1994-1996. The global estimation d (continuous line) and its uncertainty σ (dashed lines indicating $d + 1\ \sigma$ and $d - 1\ \sigma$) are added to the graph.

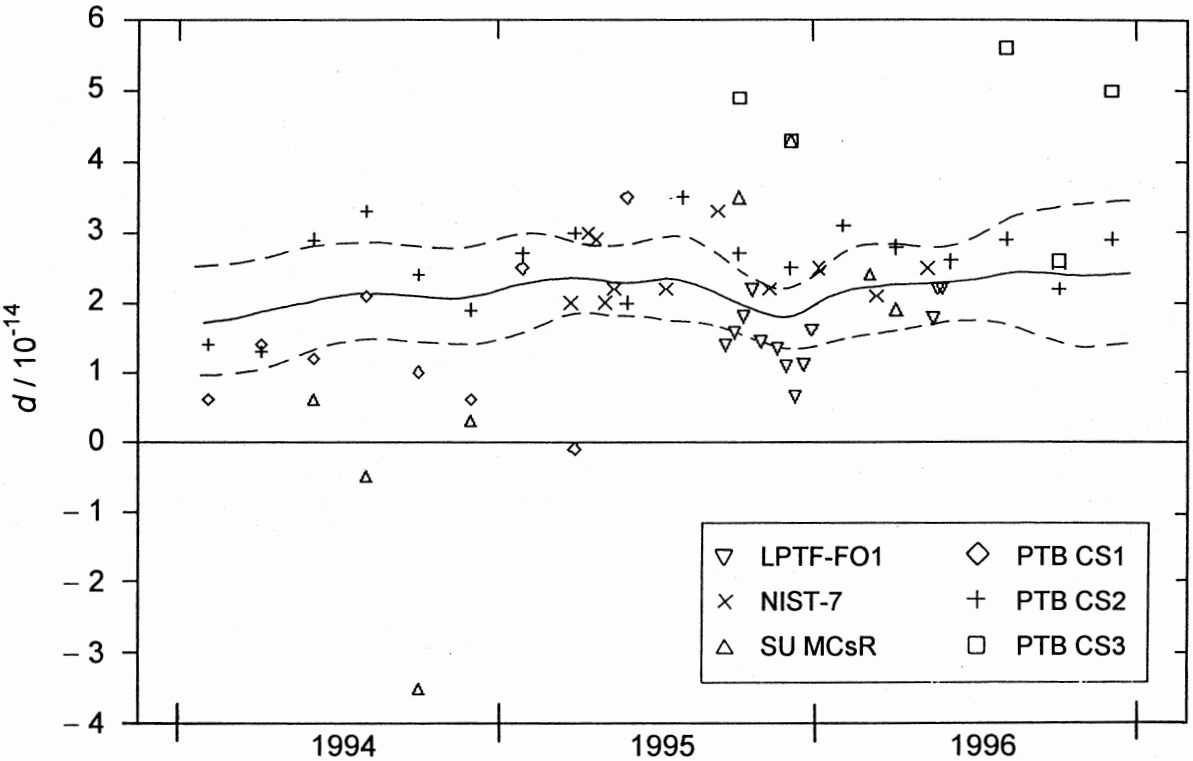
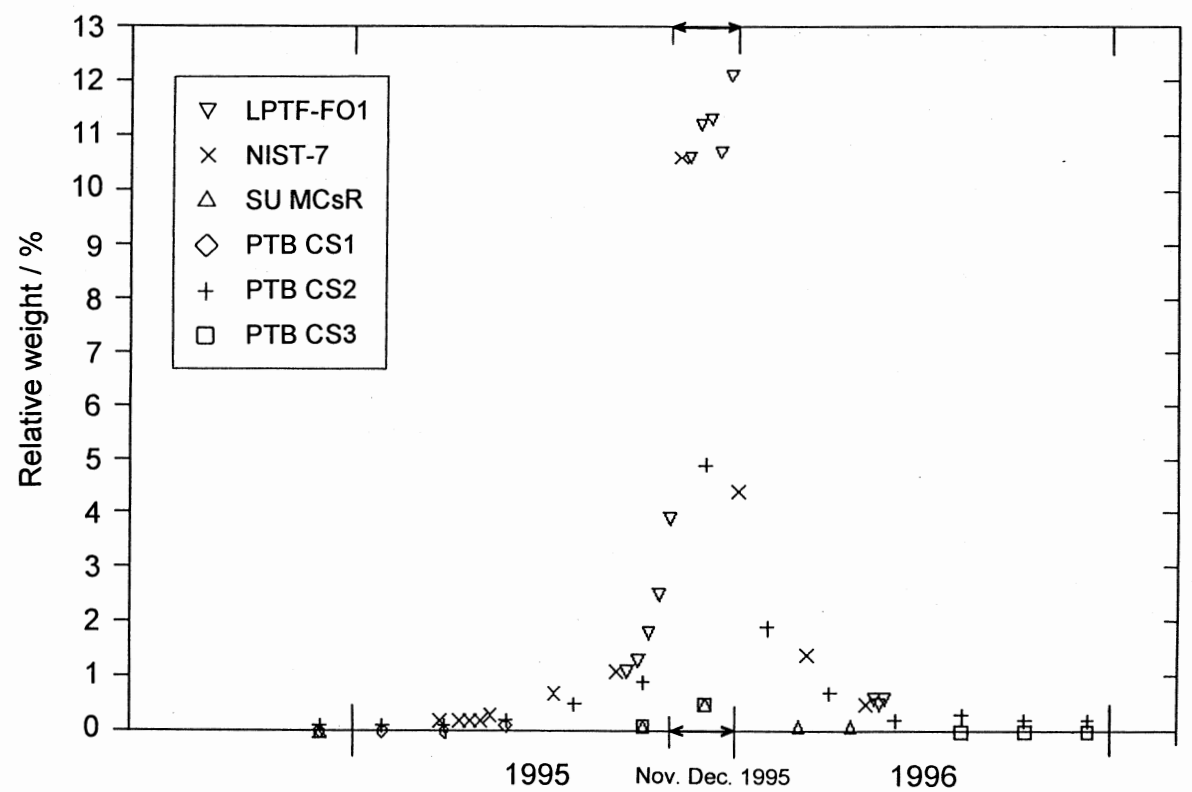


Figure 2. Relative weights attributed to individual measurements d_i provided by primary frequency standards in the global estimation of d over November-December 1995.



APPLICATIONS OF HIGHLY STABLE CLOCKS IN SPACE

Robert F. C. Vessot

Smithsonian Astrophysical Observatory
Cambridge Massachusetts 02138, USA

1. ABSTRACT

This paper describes experimental techniques for performing new tests of relativistic gravitation. These tsts involve earth orbiting clock tests, a close approach solar probe, detection of pulsed gravitational radiation, and Very Long Baseline Interferometry with spaceborne radio telescopes.

Keywords: General Relativity, atomic clocks, Doppler cancelling, Sagnac effect

2. INTRODUCTION

The uparalleled precision of scientific measurement made possible by atomic clocks has revolutionized metrology. The meter, the SI unit of length, is now defined in terms of the velocity of light, a pivotal quantity in our present concepts of space and time.

This discussion is intended to provide a picture of some available techniques for space experiments with clocks, and of new, technically feasible, space experiments.

3. TESTS OF RELATIVISTIC GRAVITATION WITH CLOCKS

3.1 Comparison of a Proposed Test of Relativistic Gravitation with a Clock in a 24-Hour Eccentric Earth Orbit with the 1976 SAO/NASA GP-A Redshift Test

In the late 1960s the original concept for testing the Gravitational Redshift called for operating a spaceborne clock in a spacecraft placed in a highly eccentric 24-hour earth orbit.[1,2]

While the original idea of an orbital test has yet to be implemented, a rocket borne test with an H-Maser was conducted in 1976. Figure 1 describes the 1976 SAO/NASA Gravitational redshift test (GP-A) and the phase-coherent system used to remove first-order Doppler and other propagation- related frequency shifts.

With this system, the fractional output frequency variations obtained by subtracting one-half of the two-way Doppler cycles from the one-way cycles received by the earth station is given in the expression:

$$\frac{f_s - f_e}{f_o} = \frac{(\phi_s - \phi_e)}{c^2} - \frac{|\vec{v}_e - \vec{v}_s|^2}{2c^2} - \frac{\vec{r}_{se} \cdot \vec{a}_e}{c^2} \qquad (1)$$

Here the total frequency shift is $f_s - f_e$, and f_0 is the clock downlink frequency. The term $(\phi_s - \phi_e)$ is the Newtonian potential difference between the spacecraft and earth station, $\vec{v}_e$ and $\vec{v}_s$ are the velocities of the earth station and the spacecraft, $\vec{r}_{se}$ is the vector distance between the spacecraft and earth station, and $\vec{a}_e$ is the acceleration of the earth station in an inertial frame. The first term in Equation 1 is the gravitational redshift resulting from the difference in the Newtonian gravitational potential between the two clocks, the second term is the second-order Doppler effect of special relativity, and the third term is the result of the acceleration of the earth station during the light time, r/c, owing to the earth's rotation. During the two-hour near-vertical flight, the first-order Doppler shifts were as large as $\pm 2 \times 10^{-5}$ and the noise from ionospheric and tropospheric propagation effects was about 1×10^{-12} at $\tau \sim 100$ sec. After the frequency variations predicted in Equation 1 were fitted to the data, the error in the fit of the data was within $(+2.5 \pm 70) \times 10^{-6}$ of Einstein's prediction.[3] The proposed orbital test is expected to give an improvement from 70 parts per million to about 1 part per million, partly from improved clocks, but mostly from the longer averaging intervals and from reduction of systematic bias errors by having time to adjust and tune the space maser before taking data.

3.2.1A Symmetrical Four-Link System to Provide Time-Correlated Doppler Data

By providing a transponded signal back to the spacecraft, as shown in Figure 2, we can record one-way, two-way, and Doppler-cancelled data at both stations of the system in terms of the proper time scale kept by the station's clock. When the light-time between stations is long compared to the intervals required for measurements, a dominant, spatially localized noise process can be cancelled *systematically* by time correlating the data obtained from the two stations. Figure 3 shows the space-time paths of the four signals in Figure 2. Here the dots signify the clocks, and the arrows, $E_1(t)$ and $E_2(t)$, signify signal outputs representing earth-based one-and two-way data. $S_1(t)$ and $S_2(t)$ represent one- and two-way data recorded in space. By time-correlating the Doppler responses we can systematically cancel a strong localized noise source such as from the earth's troposphere and ionosphere.[4,5]

3.3.2 Relativistic Doppler Shifts and Redshifts with the Four-Link System

The Doppler-cancelled signal outputs $E_0(t)$ and $S_0(t)$ in Figure 3 contain relativistic and gravitational information that can be time correlated. Equation 6 is repeated below as $A_0(t)$, along with its counterpart expression at the spacecraft, $S_0(t)$:

$$E_0(t) = \frac{\phi_s - \phi_e}{c^2} - \frac{|\vec{v}_s - \vec{v}_e|^2}{2c^2} - \frac{\vec{r}_{se} \cdot \vec{a}_e}{c^2} \qquad (2)$$

$$S_0(t) = \frac{\phi_e - \phi_s}{c^2} - \frac{|\vec{v}_e - \vec{v}_s|^2}{2c^2} - \frac{\vec{r}_{es} \cdot \vec{a}_s}{c^2} \qquad (3)$$

By adding these two time-ordered data sets we cancel the first term and double the magnitude of the second term representing the second-order shift. Conversely, if we subtract the data sets, we double the first-term representing the gravitational red (blue) shift and cancel the second term. In both instances we must account for the components of acceleration of the earth and space stations along the line of sight in a suitable inertial frame.

3.3 A Proposed Extension of the GP-A Experiment to Test General Relativity with a Solar Probe

Studies of tests of relativistic gravitation have been conducted for a clock in a space probe approaching within 4 solar radii of the sun's center in a polar orbit.[6] This test is intended to measure directly the second-order behavior of the redshift, $[\Delta\phi/c^2]^2$.

Figure 4 shows data during the 10 hours before and after perihelion. The value of Φ/c^2 varies from 5.3×10^{-7} at perihelion to 2.0×10^{-7} at times ±10 hours from perihelion. The second-order redshift $[\Phi/c^2]^2$ varies from 2.81×10^{-13} to 4×10^{-14} over the same interval. The Allan standard deviation of today's H-masers for 10 hours averaging time is 6×10^{-16}, so the inaccuracy of measurement imposed by the maser instability for the first-order measurement is 1.8×10^{-9}, the corresponding inaccuracy for the second-order measurement is 2.5×10^{-3}.

The sun's gravitational potential is complicated by having a number of multipole components. The largest of these is the solar quadrupole moment, J_2, which must be accounted for in the measurement of the second-order term in the redshift. Estimates of J_2 have been made from solar oscillations,[7] and we can estimate the effect of the uncertainty in these measurements from the behavior of the J_2 signature in the data. At order c^{-2} the first-order redshift has the following behavior:

$$\frac{\phi}{c^2} = \frac{\mu}{r} + \frac{\mu}{r^3} J_2 R^2_{sun} \frac{3\cos^2\vartheta - 1}{2} = \frac{\Phi_1}{c^2} + \frac{\Phi_2}{c^2} \qquad (4)$$

where $\mu = GM_{sun}/c^2$. Assuming $J_2 = 1.7 \times 10^{-7} \pm 0.17 \times 10^{-7}$, then at perihelion, $r = 4R_{sun}$, $\vartheta = \pi/2$, and $\Phi_2/c^2 = -2.8 \times 10^{-15}$. At about ±7 hours from perihelion, $\vartheta = 0$, $r = 8R_{sun}$, and $\Phi_2/c^2 = +7.0 \times 10^{-16}$.

The frequency variation caused by the J_2 contribution, has an estimated peak-to-peak magnitude of 3.5×10^{-15} and is also shown in Figure 4. If we have an error of 10% in the estimate of J_2, the uncertainty in its contribution to the redshift over this interval is about 3.5×10^{-16}, comparable to the instability of the clock. The error contribution will have a distinctive $(3\cos^2\vartheta - 1)/r^3$ signature in contrast to the very smooth 1/r dependence of the first-order redshift and the $1/r^2$ dependence of the second-order redshift.

A crucial feature is the ability to take Doppler-cancelled data at the probe. In Figure 4, we have shown how the localized noise near earth can be cancelled by correlation. While the spacecraft Doppler cancellation system can systematically remove the effect of the tropo-iono noise when it produces the $S_0(t_i + 2R/c)$ data, this is not the case for the earth station Doppler-cancelled output E_0. In the earth station cancelled Doppler there would be about 1000 sec of time delay between the uplink transmission and reception from the transponder. During this interval the combined atmospheric and ionospheric delay could have varied considerably.

The feasibility of this experiment has been studied jointly by SAO and NASA's Jet Propulsion Laboratory.[8]

3.3 Detection of Pulsed Gravitational Radiation using Doppler Techniques

The long travel time to the sun provides an opportunity to search for gravitational radiation using Doppler techniques.[9,10] In Einstein's General Theory of Relativity (GRT), gravitational radiation results whenever a massive body is accelerated. Rotating binary stars radiate energy and will eventually collapse together. While evidence for such radiation has not been observed directly, the orbital decay of a binary pulsar has been observed since 1975 and its rate continues to follow closely the predicted behavior for loss of energy by gravitational radiation.[11]

According to GRT, gravitational radiation is a wavelike distortion of space-time travelling at the speed of light. When a gravitational wave intercepts bodies connected by electromagnetic signals, at the receiver the transmitted frequency is changed by $h = \Delta f/f$. At some level, gravitational waves should be observable in Doppler shifts of a signals from a highly stable microwave (or laser) transmitter detected by a receiver located at a distance greater than about one-half the wavelength of the gravitational wave.

Figure 5 shows an example of Doppler detection[12] of a pulsed gravitational wave using the four-link Doppler measurement system. Here the wavefront of the gravitational pulse is assumed to intercept the earth-probe line at an angle, θ = 60 degrees. The effect of the pulse would be observed three times in the Earth 2-way Doppler trace as follows:

$$\frac{df}{f} = \frac{(1-\mu)\,\psi(t_R)}{2} - \mu\psi\left[t_R - L\frac{(1+\mu)}{c}\right] + (1+\mu)\,\psi\left(t_R - 2\frac{L}{c}\right) \qquad (5)$$

- by a Doppler shift of the gravitational wave disturbing the earth station at $t = t_1$, while it is receiving a signal transmitted earlier by the spacecraft.
- by its "echo" when the earth station receives a transponded signal at $t = t_1 + 2R/c$.
- when the gravitational wave arrives at the spacecraft at time t_2 and is reported at earth at $t = t_2 + R/c$.

The spacing of the pulses, their time signature designated by the parameter $\psi(t)$, and the relative magnitude and sign of the signature are described by a single parameter, $\mu = \cos\theta$ [13,14]. Here t_R signifies arrival time at the first station.

The one-way transmission from the spacecraft would show the pulses at t_1 and at $t_2 + R/c$.

A similar set of five observations of the same gravitational pulse is available at the spacecraft. In this case $\mu = \cos(\pi+\theta)$ and five manifestations of the pulse appear in the spacecraft data. While only four of the ten pulses, i.e.,those from the two one-way Doppler signals, are unique, the other six are follow other paths through the electronics system and provide redundancy to distinguish noise as gravitational wave signals. If one of the stations is on earth, noise from the earth's troposphere and ionosphere would be the main limitation to the sensitivity of detection. Tropospheric noise will substantially degrade the stability of a signal. Studies show that the Allan deviation of the tropospheric noise for signals passing vertically has a $\tau^{-2.5}$ behavior for intervals between 20 and 200 sec, with σ_y (100 sec) = 8×10^{-14} [15]. While it is possible to model the tropospheric frequency shifts using other data, such as the columnar water vapor content and the local barometric pressure, tropospheric propagation variations will nevertheless severely limit the detection of gravitational radiation with transponded two-way Doppler signals. Estimates have been made that the sensitivity with ideal tropospheric conditions at night in the desert will be at a level of $h = \Delta f/f \sim 3 \times 10^{-15}$ for gravitational waves in the millihertz region, which is one to two orders of magnitude above the levels estimated by astrophysicists.[16,17]

Tropospheric noise can be removed systematically by simultaneously recording Doppler data from a clock in a spacecraft and at the earth station and combining these data. Simulations of this process[5] show nearly complete rejection of such spatially localized sources as near-earth tropospheric and ionospheric variations and earth station antenna motion noises. With both clock systems operating in space and at frequencies where the noise from the solar corona ionization is not significant, the principal nongravitational noise sources will likely be from the buffeting of the space vehicles by nongravitational forces such as light pressure, particle collisions, and sporadic outgassing of the spacecraft. Here, again, since these disturbances are localized at the ends of the system, the time signatures of the noises are separated by R/c, and can be distinguished from the patterns expected from pulsed gravitational waves, which have signatures that depend only on the parameter μ of equation 3.

4. HIGH RESOLUTION VERY LONG BASELINE INTERFEROMETRY (VLBI) ASTRONOMY IN SPACE

4.1 The Effect of Oscillator Instability on the Measurement of Angles

High precision measurement of the angle between the propagation vector of a signal and the direction of a baseline, defined as the line between the phase centers of two widely separated radio telescope antennas, can be made with VLBI techniques[18] shown in Figure 6. Here, two radio telescopes, separated by a distance, L, each detect the arrival of radio noise signals from a distant radio star. After heterodyning to a lower frequency the noise signals are recorded as a function of time and the two sets of noise data are subsequently time-correlated. The observable quantities from the correlation process are the correlated amplitude and the relative phase of the signals detected at the widely separated points on the wavefront. VLBI measurements have been used in light deflection tests of relativistic gravitation.[19]

The stability limit on the successive measurements of angle imposed by the oscillator instability on successive measurements of angle taken τ seconds apart is

$$\sigma_{\Delta\theta}(\tau) \sim \frac{c\,\tau\,\sigma_y(\tau)}{L \sin\theta}, \qquad (6)$$

where q is the angle between the propagation vector and the baseline. The result of correlating the noise data obtained from a common source by the two stations is the production of fringes analogous to those observed from two-slit optical diffraction. The spacing between the fringes is $\lambda/L \sin\theta$), where λ is the average wavelength of the signals arriving at the antennas. The visibility of the fringes depends on the extent to which the signals arriving at the antennas are correlated.

The angular resolution of the interferometer is given by the change of fringe phase, ϕ, with θ

$$\frac{d\phi}{d\theta} = \frac{2\pi L}{\lambda} \qquad (7)$$

The error in successive angular measurements owing to the instability of the clocks in a terrestrial system with L = 6000 km, assuming σ_y (10^3 sec) = 1×10^{-15}, and $\theta = \pi/2$, is given by

$$\sigma_{\Delta\theta}\,(10^3 \text{ sec}) = 5 \times 10^{-11} \text{ rad or } 2\,\mu \text{ arcsec.}$$

This is far smaller than the present actual limit of 100 μ arcsec level from terrestrial stations with an 8000-km baseline operating at 7 mm wavelength.[20] The effects of tropospheric and ionospheric fluctuations impose limits that are far more serious than clock instability.

By operating VLBI stations in space, limits in angular resolution owing to tropospheric and ionospheric propagation and baseline distance, imposed by the size of the earth, can be overcome . A successful demonstration of a spaceborne radio telescope operating as a VLBI terminal was made in 1986[21] using NASA's orbital Tracking and Data Relay Satellite System (TDRSS) system as a spaceborne radio telescope in conjunction with a number of radio telescopes on earth.

As an example of the limits that a spaceborne two station system could achieve, let us consider a spaceborne system where $L = 5 \times 10^6$ km, σ_y (10^4 sec) = 4×10^{-16}, and $\theta = \pi/2$. In this case, $\sigma_{\Delta\theta}$ (10^4) = 2×10^{-13} rad or 0.05 μarcsec .

For $\lambda = 1$ mm we have $\lambda/L = 2 \times 10^{-13}$ rad and we see that the limit imposed by clock stability with 10^4 sec integration time is capable of resolving fringes at 1-mm wavelengths in a spaceborne system with baseline distances of 5×10^6 km.

Terrestrial VLBI systems are now used to record polar motion and rotation of the earth and to monitor the movements of the earth's tectonic plates. While *relative* positions of radio stars, and features of their brightness distribution, can be made with a precision of a few tenths of a milliarcsecond, the *absolute* directions in space of the baselines between VLBI stations depend on the choice of a frame of reference. This reference is usually taken from the position of very distant radio sources.

4.2 A Spaceborne Four Terminal (VLBI) Array to Establish an Inertial Reference Frame

Let us assume an array of four stations, separated by about 5×10^6 kilometers and in the form of a tetrahedron defining a three-dimensional figure in space. Figure 7 shows such an array. Each station contains a clock and is connected to its three neighbors by the four link system shown in Figure 2. [22] The six baseline distances can be precisely measured, using one-way and two-way Doppler techniques discussed earlier, and thus precisely define the *configuration* of the array as a function of time.

Defining the *orientation* of the array of six baselines poses an interesting problem. Distant radio stars define the conventionally used inertial frame. By invoking the Sagnac effect,[23] we can define the orientation of this array in a local inertial frame based on the velocity of light. It should thus be possible to compare the inertial frame from the Sagnac effect with the frame defined by distant radio stars.

The Sagnac process involves measuring arrival times of light signals sent in opposite senses about a closed path on a surface rotating at Ω rad/second. If the projected area, perpendicular to the rotation axis, on that surface is A, then the difference in the arrival times of light signals going around the path in opposite senses is $\Delta\tau = 4\Omega A/c^2$. By measuring the difference in arrival times of signals going in opposite senses about a triangle defining one face of the tetrahedron, we can obtain the component of rotation normal to that face. From the four triangles that define the tetrahedron we have an overdetermination of the rotation and can make an estimate of the accuracy of its measurement.

For the array shown in Figure 8, the sensitivity of angular rotation measurement, $\Delta\Omega$, is on the order of , is 1.2×10^{-15} rad/sec. If the array were located at 1 AU from the sun, such a rotation measurement should include the Einstein–deSitter precession of 3×10^{-15} rad/sec owing to the bending of space-time by the sun's gravity.

This array could compare, with very high precision, the frame of reference defined by the most distant radio sources with an inertial frame defined by the local velocity of light. Perhaps this system might provide a way to observe some aspect of the behavior of the missing matter in the universe.

5. CONCLUSION

The atomic frequency standards now serve as clocks and meter sticks and this basis of modern metrology is consistent with present concepts of relativistic gravitation. We challenge our imaginations with measurements of astronomical and astrophysical objects near the edge of our universe that depend on clocks governed by quantum processes involving staggeringly smaller scales of dimensions! As the performance of atomic clocks improves and their uses are extended in astrophysical measurements, perhaps we will see some surprises about the nature of our universe.

6. REFERENCES

1 D. Kleppner, R.F.C. Vessot, and N.F. Ramsey, "An orbiting clock experiment to determine the gravitational redshift," *Astrophys. Space Sci.* , vol. 6, pp. 13-32, 1970.

2 L.L. Smarr. R.F.C. Vessot, C.A. Lundquist, R. Decher, and T. Piran, "Gravitational waves and redshifts: A space experiment for testing relativistic gravity using multiple time-correlated radio signals," *Gen. Rel. and Grav.* vol. 15, no. 2, pp. 129-163, 1983.

3 R.F.C. Vessot, M.W. Levine, E.M. Mattison, E.L. Blomberg, T. E. Hoffmann, G.U. Nystrom, B.F. Farrell, R. Decher, P.B. Eby, C.R. Baugher, J.W. Watts, D.L. Teuber, and F.D. Wills, "Tests of relativistic gravitation with a space-borne hydrogen maser," *Phys. Rev. Lett.*, vol. 45, pp. 2081-2084, Dec. 1980.

4 R.F.C.Vessot and M.W. Levine, "A time-correlated four-link Doppler tracking system," in A Closeup of the Sun (M. Neugebauer, and R.W.Davies, Eds.) JPL Publication 78-70, NASA, 1978.

5 T. Piran, E. Reiter, W.G. Unruh, and R.F.C. Vessot, "Filtering of spacecraft Doppler tracking data and detection of gravitational radiation," *Phys. Rev. Lett., D*, 34, no. 4, p. 984, 1986.

6 M. Neugebauer and R.W. Davies, *A Closeup of the Sun*, JPL Publications 78-70, NASA, 1978.

7 T. M. Brown, J. Christensen-Dalsgaar, W.A. Dziembowski, R. Goode, D.O. Gough, and C.A. Morrow, "Inferring the sun's internal angular velocity from observed P-mode frequency splitting," *Astrophys. J.*, vol. 343, p. 526, 1989.

8 NASA contract numberNAS-918 with J.D. Anderson of JPL and E.M. Mattison and R.F.C. Vessot of SAO.

9 F. B. Estabrook, "Gravitational wave searches with ground tracking networks," *ACTA Astronautica*, vol. 17, no. 5, pp. 585-587, 1988.

10 J. W. Armstrong, J.D. Anderson, and E.L. Lau, "Application of hydrogen maser technology to the search for gravitational radiation,"in *Proc. of Twenty-first Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting,* Redondo Beach, CA. Nov. 28-30, 1989, pp. 259-263.

11 J. H. Taylor and J.M. Weisberg, "Further experimental tests of relativistic gravity using the binary pulsar PSR 1913 + 16," *Astrophys. J.*, vol. 345, pp. 434-450, Oct, 1989.

12 A. J. Anderson, "Probability of long period (VLF) gravitational radiation," *Nature*, vol. 229, pp. 547-548, 1971.

13 F. B. Estabrook and H.D. Wahlquist, "Response of Doppler spacecraft tracking to gravity waves," *Gen. Rel. Grav.*, vol. 6, pp. 439-447, 1975.

14 K. S. Thorne and V.B. Braginsky, "Gravitational-wave bursts from the nuclei of distant galaxies and quasars: proposal for detection using Doppler tracking of interplanetary spacecraft," *Astrophys. J. Lett.*, vol. 204, L1, 1976.

15 A.E.E. Rogers, A.T. Moffet, D.C. Backer, and J.M. Moran, "Coherence limits in VLBI observations at 3-millimeter wavelength," *Radio Sci.*, vol. 19, no. 6, pp, 1552-1560, 1984.

16 J. W. Armstrong, J.D. Anderson, and E.L. Lau, "Application of hydrogen maser technology to the search for gravitational radiation,"in *Proc. of Twenty-first Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting,* Redondo Beach, CA. Nov. 28-30, 1989, pp. 259-263.

17 F. B. Estabrook and H.D. Wahlquist, "Response of Doppler spacecraft tracking to gravity waves," *Gen. Rel. Grav.*, vol. 6, pp. 439-447, 1975.

18 A.E.E. Rogers and J.M. Moran, Chapter 5, "Interferometers and Arrays," Chapter 5 in *Methods of Experimental Physics,* Astrophysics, M.L. Meeks, Ed., vol. 12. New York: Academic Press, 1976.

19 E. B. Formalont and R.A. Sramek, "Measurement of the solar gravitational deflection of radiowaves in agreement with general relativity," *Phys. Rev. Lett.*, vol. 36, p. 1475, 1976.

20 N. Bartel, V. Dhawan, T. Krichbaum, D.A. Graham, I.I.K. Pauliny-Toth, A.E.E. Rogers, B.O. Rönnäng, J.H. Spencer, H. Hirabayashi, M. Inoue, C.R. Lawrence, I.I. Shapiro, B.F. Burke, J.M. Marcaide, K.J. Johnston, R.S. Booth, A. Witzel, M. Morimoto, and A.C.S. Readhead, "VLBI imaging with an angular resolution of 100 microarcseconds,"*Nature,* vol. 334, no. 6178, pp. 131-135, 1988.

21 G.S. Levy, C.S. Christensen, J. F. Jordan, R.A. Preston, C.D. Edwards, R.P. Linfield, S.J. DiNardo, L. Skjerve, J.S. Ulvestad, T. Hayashi, T. Nishimura. T. Taskano, T. Yamada, T. Shiomi, H. Kunimori, N. Kawaguchi, M. Inoue, M. Morimoto, H. Hirabayashi, B.F. Burke, A. Whitney, D.L. Jauncey, C.H. Ottenhoff, K. Blaney, and W. Peters, "Results and communications considerations of the very long baseline interferometry demonstration using the tracking and data relay satellite system," *Acta Astron.*, vol. 15, no. 6/7, pp. 481-487, 1986.

22 R.F.C. Vessot, "Clocks and spaceborne tests of relativistic gravitation," in*Advances in Space Research,* Eds. (R.D. Reasenberg and R.F.C. Vessot Eds.) Pergammon Press, vol.9, No. 9, 1989.

23 E.J. Post, "Sagnac Effect," *Rev. Mod. Phys.* , vol. 39, p. 475, 1967.

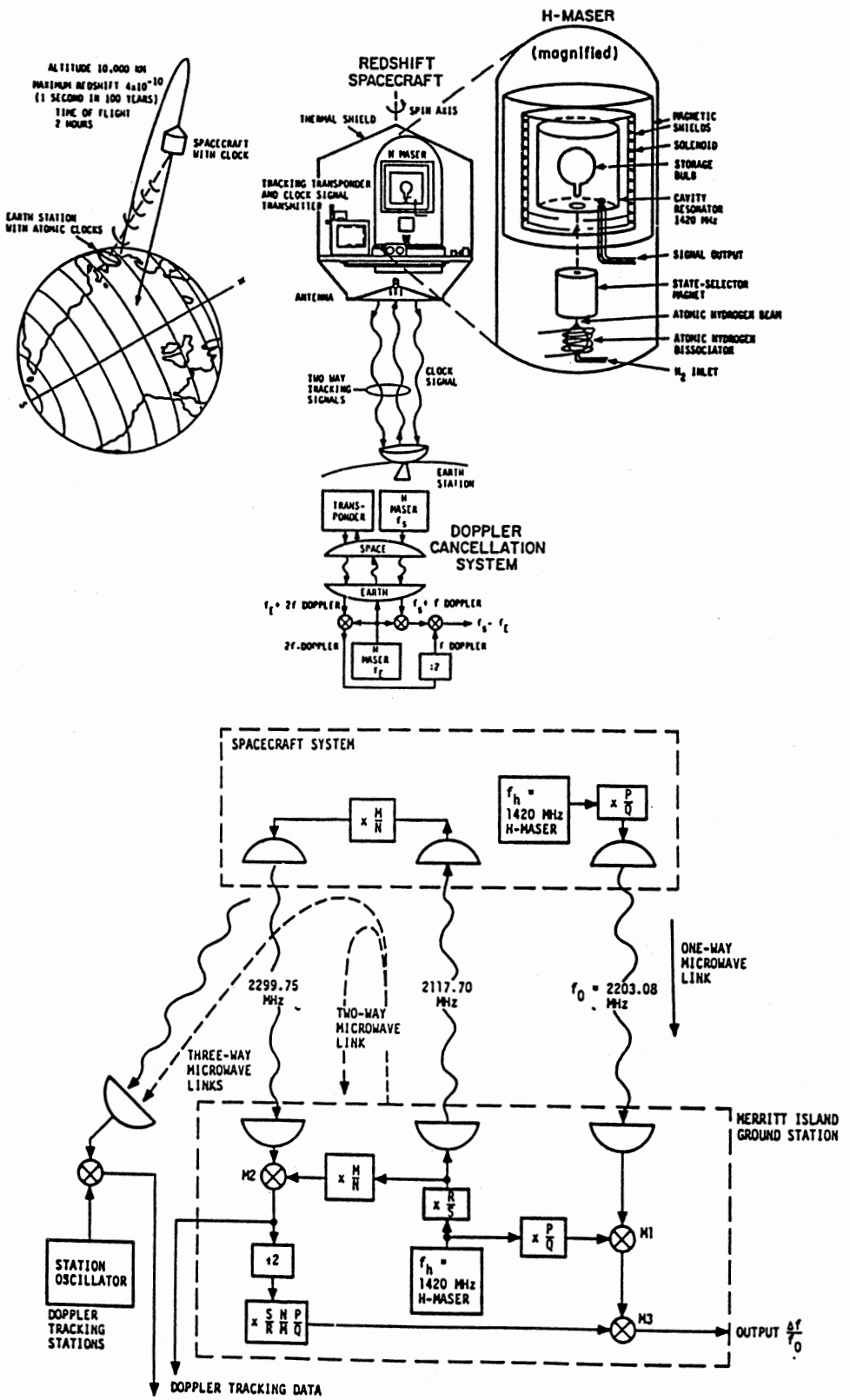


Figure 1
The 1976 GP-A (Red Shift) test and the Doppler cancelling system
M/N = 240/221; R/S = 82/55; P/Q = 76/49

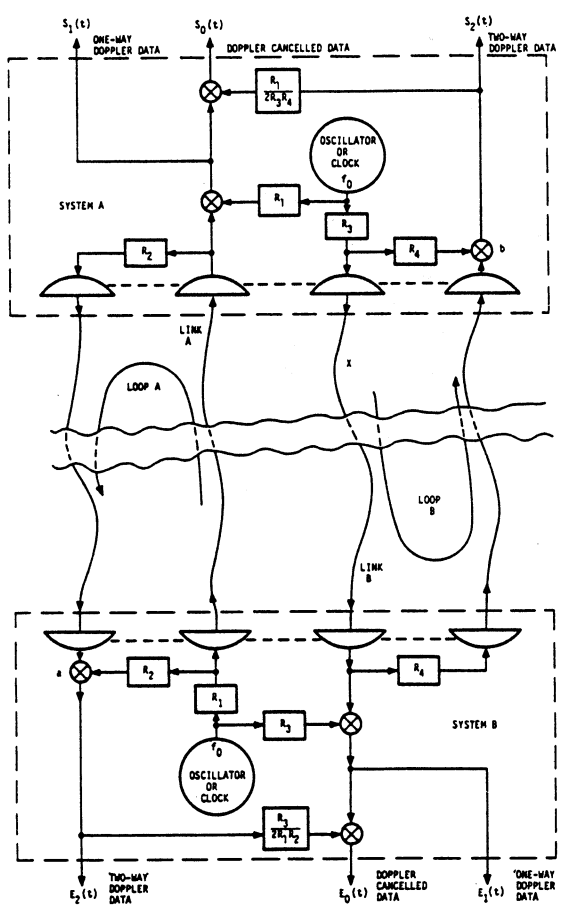


Figure 2
Concept of the four-link time-correlated Doppler cancelling system

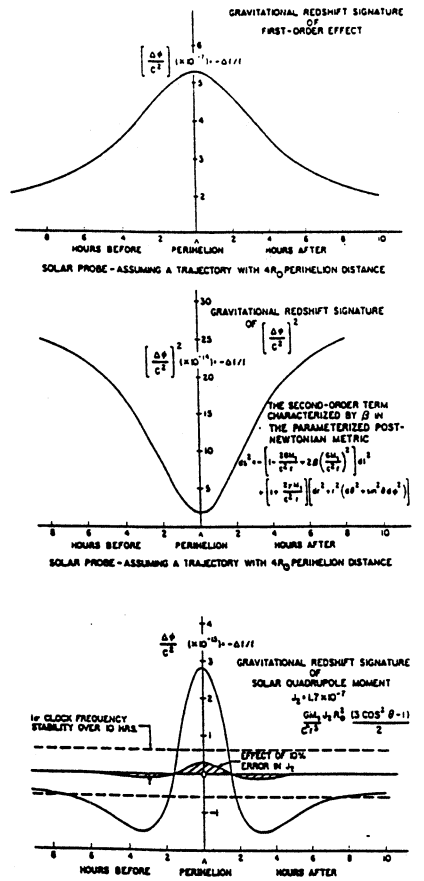


Figure 4
Gravitational frequency signatures in solar probe data

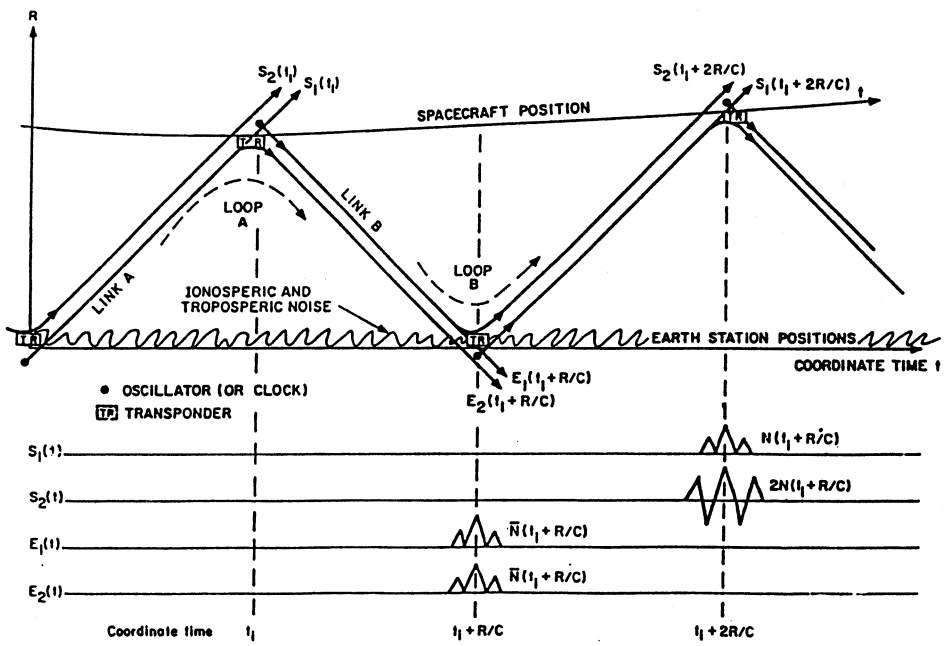


Figure 3
Space-time diagram of signals the four-link system showing cancellation of near-earth disturbances

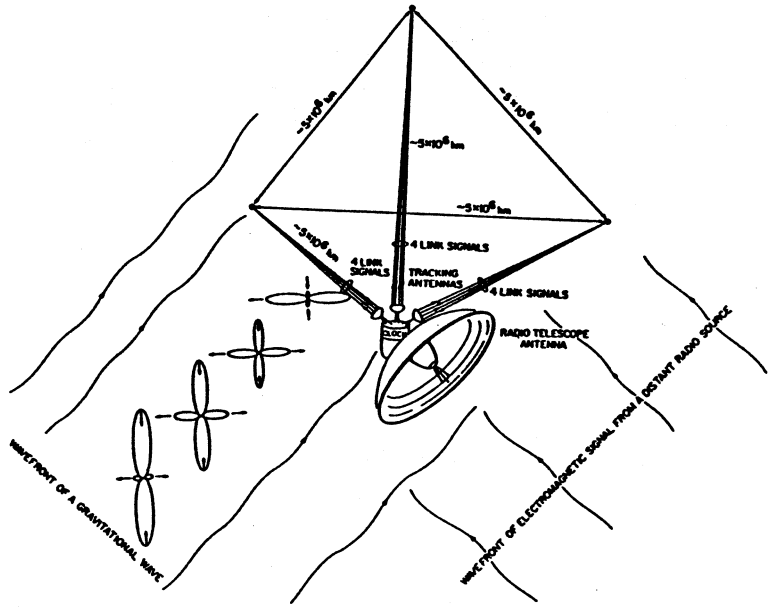


Figure 7
A spaceborne array of radiotelescopes interconnected by four-link systems

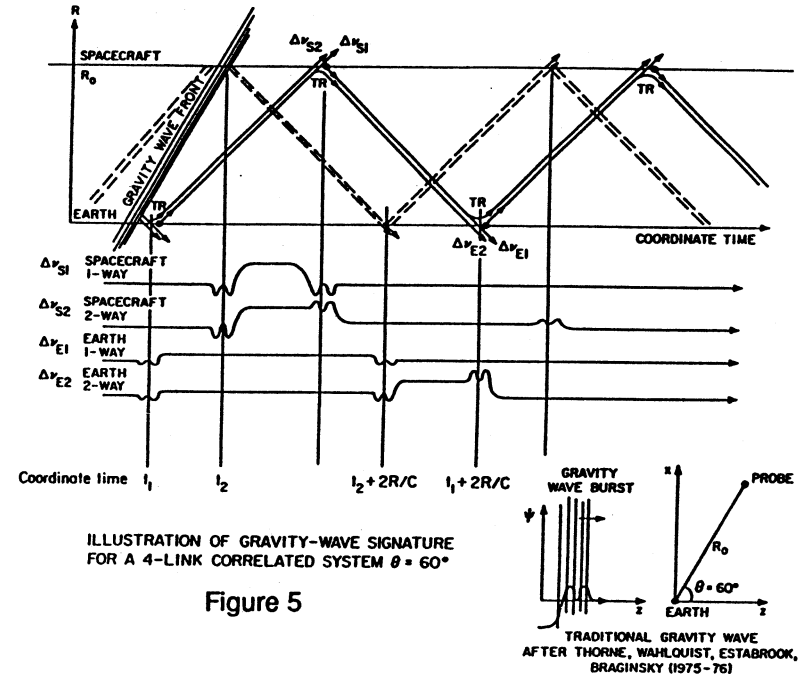
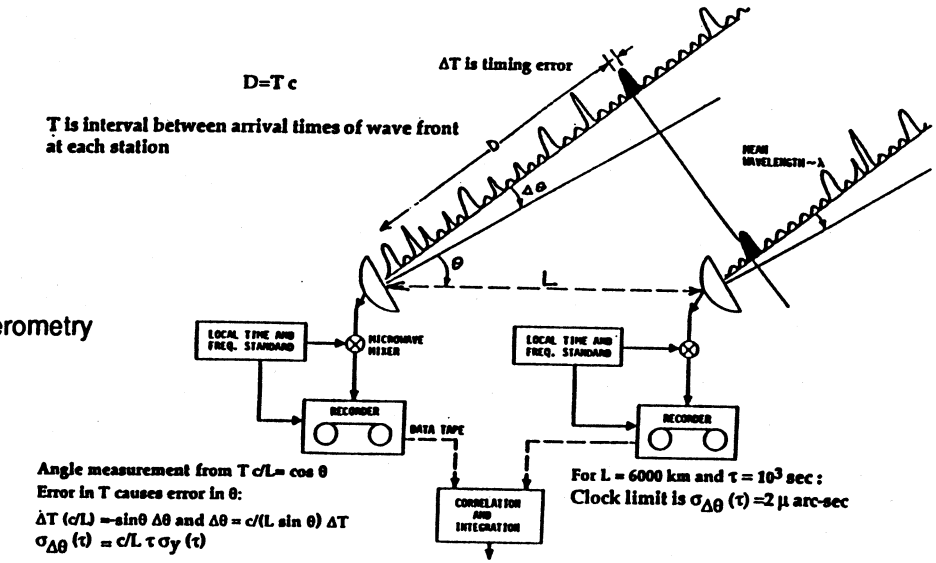


ILLUSTRATION OF GRAVITY-WAVE SIGNATURE FOR A 4-LINK CORRELATED SYSTEM $\theta = 60°$

Figure 5

TRADITIONAL GRAVITY WAVE AFTER THORNE, WAHLQUIST, ESTABROOK, BRAGINSKY (1975-76)

Figure 6
Measurement of angle by Very Long Baseline Interferometry



VLBI enables the highest angular resolution of any astronomical technique (100 μ arc-sec)
Operation in space removes limitations imposed by atmospheric disturbances and finite baseline length

SATELLITE TEST OF SPECIAL RELATIVITY USING THE GPS

Peter Wolf[1,2] and Gérard Petit[1]

[1]*Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92312 Sèvres CEDEX, France*
[2]*Queen Mary and Westfield College, School of Mathematical Sciences, Mile End Road, London E1 4NS, Great Britain*

ABSTRACT

A novel test of the second postulate of special relativity (the universality of the speed of light) has been carried out using data of clock comparisons between hydrogen maser clocks on the ground and caesium and rubidium clocks on board 25 GPS satellites. The clocks were compared via carrier phase measurements of the GPS signal using AOA Rogue and Turbo-Rogue geodetic receivers at a number of stations of the IGS (International GPS Service for Geodynamics) spread world-wide. A violation of the second postulate can be modelled by an anisotropy of c along a particular spatial axis with experiments setting a limit on the value of $\delta c/c$ along this axis. Within this model our experiment is sensitive to a possible anisotropy in any spatial direction, and on a non-laboratory scale (baselines $\geq$ 20000 km). The results presented here set an upper limit on the anisotropy of c of $\delta c/c < 5\text{x}10^{-9}$ when considering all spatial directions and $\delta c/c < 2\text{x}10^{-9}$ for the component in the equatorial plane. These are the most stringent limits for this parameter reported up to date.

1. INTRODUCTION

Einstein's second postulate, affirming the universality of the speed of light for inertial frames, is fundamental to the theories of special and general relativity. It can be tested directly by comparing the one-way propagation times of light signals along known paths, but in different spatial directions (often referred to as a test of the isotropy of the one-way speed of light). The only such test was carried out by Krisher et al. [1], who compared the phases of two hydrogen masers separated by a distance of 21 km and linked via an ultrastable fibre optics link of the NASA deep space network. The sensitivity of this experiment, expressed as a limit on the anisotropy of the speed of light, was $\delta c/c < 3{,}5\text{x}10^{-7}$, where c is the velocity of light in vacuum. Riis et al. [2] tested the isotropy of the first order Doppler shift of light emitted by an atomic beam (and indirectly thereby the second postulate) using fast-beam laser spectroscopy obtaining what is currently the smallest limit for the anisotropy, $\delta c/c < 3\text{x}10^{-9}$. Both of these experiments relied on the rotation of the Earth to change the direction of signal transmission and were therefore only sensitive to a component of the anisotropy that lies in the equatorial plane. The GP-A rocket experiment [3] can be interpreted as testing the isotropy of the first order Doppler shift of the link between the ground and on-board masers, giving a limit of $\delta c/c < 3{,}2\text{x}10^{-9}$ in one particular spatial direction. The only experiment sensitive to anisotropy in any spatial direction was carried out by Turner and Hill [4] who tested the isotropy of the first order Doppler shift in a Mössbauer rotor, obtaining a limit of $\delta c/c < 3\text{x}10^{-8}$. We present here the results of a test of Einstein's second postulate sensitive to an anisotropy of c in any spatial direction. Using the clocks on-board the Global Positioning System (GPS) satellites (providing baselines $\geq$ 20000 km) we obtain a limit of $\delta c/c < 5\text{x}10^{-9}$ when considering all spatial directions and $\delta c/c < 2\text{x}10^{-9}$ for the component of the anisotropy that lies in the equatorial plane. These results, together with those obtained by previous experiments are summarised in Tab. 1. In Sec. 2 the principle of the experiment is explained, Sec. 3 provides details about the experimental procedure and data treatment, and Sec. 4 shows the results. We consider possible systematic effects in Sec. 5 with a final discussion and conclusion in Sec. 6.

2. PRINCIPLE OF THE EXPERIMENT

Satellites of the GPS constellation are distributed in six orbital planes, at an inclination of 55° in near circular orbits with a period corresponding to 0,5 sidereal days (718 min) [5]. Each satellite is equipped with an on-board atomic clock and a dual-frequency signal transmission system.

The emission time of a signal as measured by the on-board clock τ_e and its reception time as measured by the ground-clock τ_r are recorded. The difference $T = \tau_r - \tau_e$ represents the transmission time of the signal plus some initial phase difference between the clocks. Note that no synchronisation convention or procedure is assumed. Defining D as the distance along a straight line from the satellite (at the moment of emission) to the ground station (at the moment of reception) in a geocentric, inertial (non-rotating) coordinate system one can write

$$T - \frac{D}{c} = \Delta_0 \qquad (1)$$

where Δ_0 is a constant characterising the initial phase difference of the two clocks. Einstein's second postulate requires that, for a series of measurements in different directions (e.g. during a complete passage of the satellite), T-D/c should remain constant, after correction for the relative rate of the two clocks due to the gravitational redshift, second order Doppler shift and the intrinsic (proper) frequency difference of the clocks.

In the theoretical framework generally used for the interpretation of experiments that test the second postulate of special relativity, the speed of light is anisotropic along a particular spatial direction (in an inertial frame) i.e. its component along this preferred axis is $c+\delta c$ in one direction and $c-\delta c$ in the opposite direction. The experiments then determine whether the special relativistic postulate $\delta c = 0$ is confirmed within the uncertainty of the experiment and set an upper limit on the anisotropy of the speed of light, that is, on the parameter $\delta c/c$. A more sophisticated theoretical approach to all tests of special relativity (including those testing the second postulate) was developed by Mansouri and Sexl [6]. A detailed interpretation of our experiment in this framework is presented elsewhere [7]; only the results are given here (see Sec. 6).

The effect of a possible anisotropy of magnitude $\delta c/c$ on an individual link would be $(\delta c/c)(D/c)\cos\alpha$, where α is the angle between the direction of the anisotropy and of the transmitted signal, resulting in a measurable variation of T as a function of direction. However, such an anisotropy might also affect the determination of the satellite ephemerides, and therefore the value of D, leaving the difference T-D/c unchanged. A meaningful test of special relativity using the above principle therefore requires a method of satellite orbit determination which is insensitive to a non-zero value of $(\delta c/c)$.

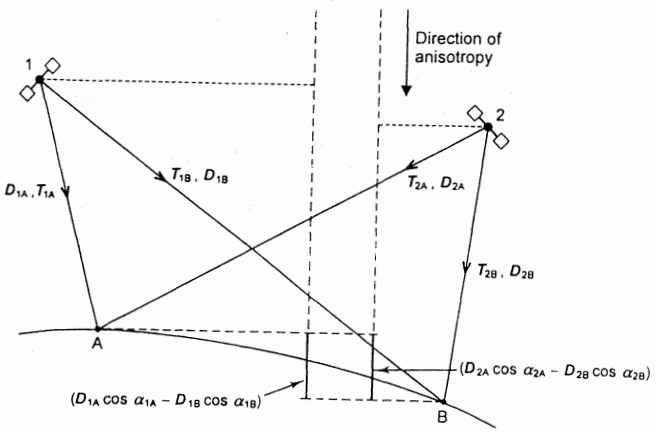


Fig. 1: Double difference (Y) for pairs of stations (A and B) and satellites (1 and 2), $Y=(T_{1A}-T_{1B})-(T_{2A}-T_{2B})$. The effect of an anisotropy on Y given by $[(D_{1A}\cos\alpha_{1A}-D_{1B}\cos\alpha_{1B})-(D_{2A}\cos\alpha_{2A}-D_{2B}\cos\alpha_{2B})]\delta c/c^2$, where α is the angle between the directions of the anisotropy and of the transmitted signal, vanishes.

This is the case for the GPS ephemerides obtained by the IGS-CODE processing centre [8]. The method used adjusts a post-Keplerian, non-relativistic orbit model to doubly differenced GPS timing data (see Fig. 1). The effects of anisotropy on the individual links cancel (to first order) when the double differences are formed (see Fig. 1). The IGS-CODE method is used to simultaneously adjust a number of parameters, including the satellite ephemerides and the ground station coordinates, thereby providing values of D which are unaffected by anisotropy of c.

Additionally one has to ensure that corrections applied to the raw timing data used for orbit determination and the measurement of T do not presuppose the second postulate. In fact two corrections are routinely applied to GPS timing data which are of relativistic origin and therefore do imply the isotropy of c [5]: the correction for the gravitational redshift and the second order Doppler shift of the rate of the satellite clock with respect to coordinate time, and the correction for the so called Sagnac effect which is due to the rotation of the Earth during signal transmission. Both of these are small corrections of order c^{-2} hence the effect of an error in these corrections due to an anisotropy would be negligible with respect to the first order effect on T (for a more detailed theoretical analysis see e.g. [12]).

Observation of GPS satellites in varying spatial directions thus provides a meaningful test of the second postulate of special relativity via relation (1) with D obtained using IGS-CODE ephemerides and station coordinates.

3. EXPERIMENTAL PROCEDURE

The IGS is a global network of ground stations that continuously observe the GPS satellites for civil, geodetic purposes [8]. From the raw data the IGS processing centres calculate (among other parameters) precise satellite ephemerides and ground station coordinates. These, together with the raw observations, are freely available through the internet via anonymous ftp [8]. We use data from eight ground stations for our experiment: Brussels (Belgium), Algonquin (Canada), Yellowknife (Canada), Fairbanks (Alaska, USA), Kokee Park (Hawaii, USA), Fortaleza (Brazil), Santiago (Chile) and Hobart (Australia). The motivation for this choice of ground stations is to ensure global coverage whilst providing maximum ground clock stability (for averaging times $\approx$ 6 h (one passage)) by using only stations which are equipped with hydrogen-maser clocks. The GPS receivers used are all AOA Rogue or Turbo-Rogue geodetic receivers providing raw phase measurements of the two GPS carrier frequencies at a sample interval of 30 seconds. The data sets cover six days (1994 September 18 to 23) and contain observations of all 25 GPS satellites available at the time. During this period (coinciding with the military intervention in Haiti) all GPS signals

were free of the intentional degradation (Selective Availability, SA) which is imposed by the US military. In general, this affects all but two satellites, making them unusable for the experiment described here. Of the 25 satellites used 19 are equipped with caesium clocks, and 6 with rubidium clocks.

From the raw data the differences *T-D/c* are formed, taking into account corrections for the variable part of the gravitational redshift and second order Doppler shift, the Sagnac effect, the ionospheric delay (using an ionosphere-free combination of the two frequencies), and the tropospheric delay (using the STANAG tropospheric model [5]). The amount of data is reduced by averaging over nine 30 s points, in order to save computer space and to make the data more manageable. Fig. 2 shows a typical example of a resulting data set for a station-satellite pair.

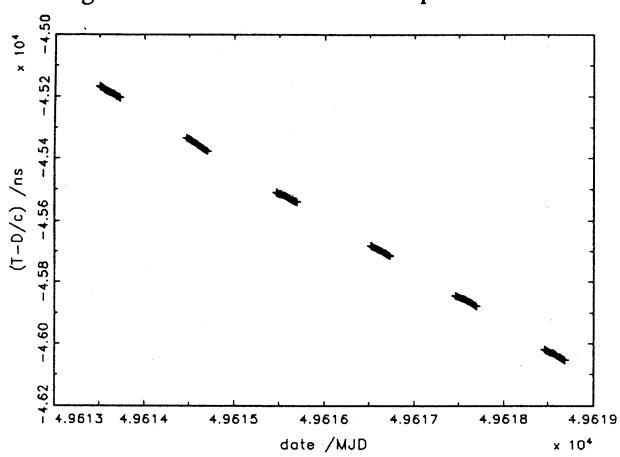


Fig. 2: Measurements of (*T-D/c*) for 6 passages of PRN04 over Algonquin(CAN).

For a test of the isotropy of *c*, one is interested in the variation of the difference *T-D/c* during individual passages of the satellite over the ground station i.e. variations over time scales of roughly 6 hours. Therefore we first filter the data, excluding all long-term variations (greater ≈ 5 days) effectively subtracting the relative rate of the two clocks. Then an arbitrary offset per passage is adjusted as only the *variation* of *T-D/c* during the passage is of interest.

As is well known, measurements of the GPS carrier phase are subject to an unknown phase ambiguity error of an integer number of cycles. This does not present a problem for our purposes as long as the induced error remains constant during each passage, which is the case if the receiver stays locked onto the satellite over the complete passage. Therefore all passages that are incomplete (data gaps indicating a possible loss of the satellite) were excluded.

The 0,5 sidereal day period of the GPS satellites implies that a station sees each passage of a particular GPS satellite at the same time of day (in sidereal days) and in the same directions (in a geocentric non-rotating frame). So, to see the data more clearly, all passages can be projected onto the same day, by shifting them individually by an integer number of sidereal days. Figure 3 shows the residuals of a typical data set after filtering and adjustment of an offset per passage, and with all passages shifted onto the same day. The standard deviation of these residuals for the complete data set (all stations and satellites) is 2,2 ns.

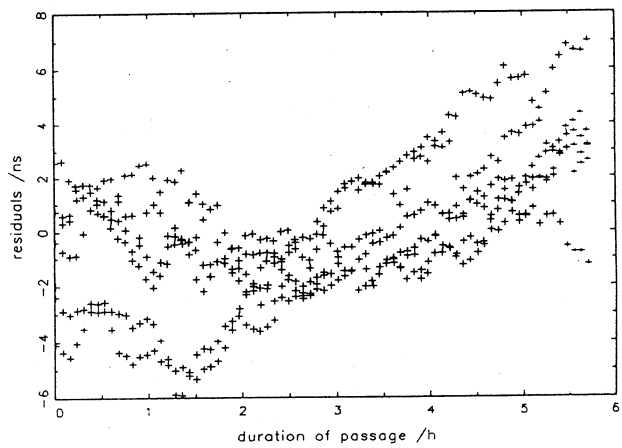


Fig. 3: Residuals of (*T-D/c*) after filtering and adjustment of an offset per passage. The graph shows 6 passages of PRN22 over Brussels(B) shifted onto the same day.

4. RESULTS

Figure 4 shows the spatial directions of signal transmission for the individual links in a non-rotating geocentric frame. There are no links at colatitudes below ≈20° and above ≈163° which is due to the 55° inclination of the satellite orbits and implies that the experiment was least sensitive in the N-S direction.

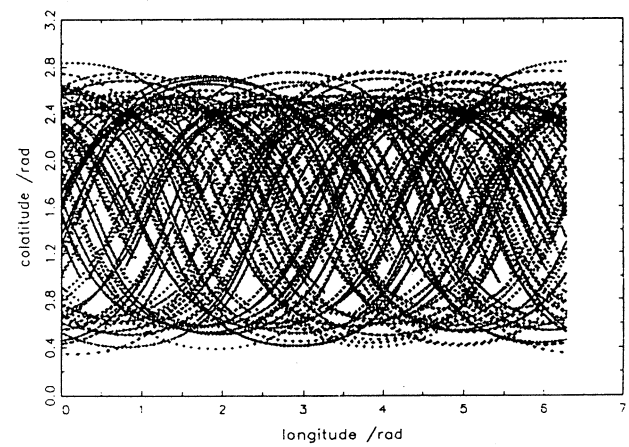


Fig. 4: Directions of signal transmission, in a geocentric non-rotating frame, for all satellite-station links.

The effect of an anisotropy of *c* on the transmission time *T* for a particular link is given by $(D/c)\ (\delta c/c)\ \cos\alpha$ where α is the angle between the direction of signal transmission and the direction of the anisotropy. This model was fitted to the data using the least squares method, adjusting the magnitude of the

anisotropy ($\delta c/c$) and an offset per satellite-station pair. The adjustment was performed for a range of anisotropy directions, spanning colatitudes and longitudes from 0 to π in a grid of 0,1 rad x 0,1 rad. It is sufficient to cover half of all possible spatial directions, as opposing directions correspond to the same anisotropy with the opposite sign. Directions are given in the non-rotating geocentric frame that is coincident with the ITRF Earth fixed frame at MJD 49755 (March 7, 1995) 0h 00 (UTC).

Figure 5 shows the adjusted magnitudes of the anisotropies as a function of their direction. The extremum value of ($\delta c/c$) is $4{,}9\text{x}10^{-9}$ at a colatitude of 2,9 rad and a longitude of 0,5 rad. In the equatorial plane (colatitude of 1,6 rad) the extremum value of ($\delta c/c$) is $1{,}6\text{x}10^{-9}$ at a longitude of 0,6 rad. The formal statistical standard uncertainties of these values are below $3\text{x}10^{-10}$.

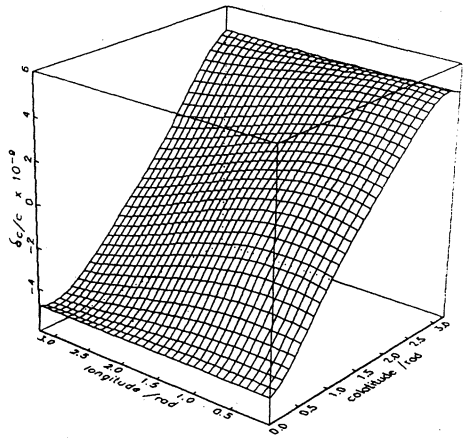


Fig. 5: Magnitude ($\delta c/c$) of the adjusted anisotropy as a function of its direction in a geocentric non-rotating frame.

5. SYSTEMATIC EFFECTS

The three main systematic effects that could affect the data are the satellite clock instability, ephemerides errors and uncertainties in the modelled tropospheric delays.

Pre-launch measurements of the relative frequency stability of the GPS caesium clock showed an instability of $\sigma_y(\tau) \approx 1{,}5 \times 10^{-13}$ (standard Allan deviation) for integration times of ≈ 6 h [9]. On-orbit measurements showed satellite clock instabilities of order 10^{-13} for integration times of 1 day [10], which corresponds to $\sigma_y(6\text{ h}) \approx 2 \times 10^{-13}$ when extrapolated assuming a $\tau^{-1/2}$ dependence. This translates into an accumulated time error over one passage of the satellite of $\delta\tau \approx \tau\sigma_y(\tau) \approx 4{,}3$ ns.

Beutler et al. [11] estimate the uncertainty of the IGS-CODE satellite ephemerides to be 15 cm to 20 cm corresponding to a timing error of $\delta\tau \approx 0{,}7$ ns.

The tropospheric delays were estimated using the standard STANAG model [5]. More accurate values for these delays can be obtained through estimations in a regional or global network (as done for example by the IGS). We have studied such estimates (four values per day and station) for a period of five days in September 1995 and for six of the eight stations used (no IGS estimates were available for the remaining two stations). The differences between the zenithal tropospheric delay obtained from the STANAG model and the 5 day averages of the IGS estimates do not exceed 200 ps for any station with the standard deviation of the averages < 110 ps. For low elevations these differences can increase to maximum 1,1 ns (at elevations of 10°). We conclude that, for our purposes, the STANAG model is sufficient as the dominant limitation of the sensitivity of our experiment is more likely to arise from satellite clock instabilities.

6. DISCUSSION AND CONCLUSION

Timing errors due to the systematic effects discussed in the previous section could give rise to values of $\delta c/c$ of order 10^{-8} ($\delta c/c \approx \delta\tau/(D/c)$) assuming that no cancellation takes place between the systematic effects of different satellite-station pairs. This is an order of magnitude larger than the maximum value observed. The experiment therefore cannot suggest a violation of the second postulate of special relativity.

It is unlikely that, in a global treatment, the systematic errors are correlated with the signature of an anisotropy, as they are expected, in general, to correlate differently with the effect of an anisotropy for each satellite-station pair and, to some extent, for each passage. So the results from a number of randomly chosen subsets of the data can provide an estimate of the reliability of the results presented in section 4. The least squares adjustment was repeated for five different subsets of the complete data, removing observations from 3 to 6 satellites (i.e. 12% to 24% of the total data). The mean of the five adjusted extremum values of $\delta c/c$ is $4{,}6\text{x}10^{-9}$ when considering all spatial directions and $2{,}2\text{x}10^{-9}$ for the component in the equatorial plane, with a standard deviation about the mean of $1{,}3\text{x}10^{-9}$ in both cases. These values should be compared to $\delta c/c$ of $4{,}9\text{x}10^{-9}$ and $1{,}6\text{x}10^{-9}$ obtained from the complete data set (c.f. section 4).

So, assuming no correlation (and resulting cancellation) between the systematic effects in a global treatment and the effect of a possible anisotropy (for the reasons mentioned above), we can set a limit of $\delta c/c < 5\text{x}10^{-9}$ on the spatial variation of c when considering all spatial directions and $\delta c/c < 2\text{x}10^{-9}$ for the component of the anisotropy that lies in the equatorial plane.

Interpretation of the experiment, using the test theory developed by Mansouri and Sexl [6], shows that the value of $\delta c/c$ obtained here is related to the parameter α of the test theory (in special relativity α = -1/2) by the well known relationship (see [7] for details)

$$\frac{\delta c}{c} = (1+2\alpha)\frac{v}{c}, \qquad (2)$$

where v is the velocity of the Earth with respect to the "universal frame" defined in the theory. Additionally one obtains an explicit expression for the double difference Y (c.f. Fig. 1) that is independent of α to first order in v/c [7], which confirms the validity of the experimental principle in this test theory. Taking v as the velocity of the Earth with respect to the "mean rest frame of the universe" ($v \approx 300$ km/s) in the direction of the dipole anisotropy of the cosmic microwave background (declination = -6,1°, right ascension = 11,2 h) [13, 14] we obtain a limit of $|\alpha+1/2| < 1\text{x}10^{-6}$ ($\delta c/c < 2\text{x}10^{-9}$ in this direction). This, to our knowledge, is the smallest limit for the parameter α published up to date.

In principle, any space mission flying a highly stable atomic clock that is equipped with a time transfer system could be used for a similar test of special relativity. An example is the ESA/RSA ExTRAS mission [15] that is currently "on hold". This mission is of interest as the space-clock will be a hydrogen maser expected to be significantly more stable in the short term than the GPS clocks, and because the optical two-way time transfer system planned should diminish the effects of the time transfer (in particular tropospheric) and ephemerides uncertainties [15].

Finally we would like to emphasise that an experiment like the one described in this article can be carried out at minimal cost by virtually anyone as all the IGS data is freely available on the internet via anonymous ftp. This is of particular interest in view of a recent US presidential decision [16] to switch off SA completely on all GPS satellites within the next ten years.

ACKNOWLEDGEMENTS

We thank the IGS stations and centres for making their data and their products freely available, Claudine Thomas for helpful discussions, and D.A. Blackburn for revision of the manuscript.

REFERENCES

[1] Krisher T.P. et al., Phys. Rev. D (rapid communication) **42**, 731, (1990).
[2] Riis E. et al., Phys. Rev. Lett. **60**, 81, (1988).
[3] Vessot R.F.C. and Levine M.W., General Relativity & Gravitation **10**, 181, (1979).
[4] Turner K.C. and Hill H.A., Phys. Rev. **134**, 1B, B252, (1964).
[5] *NATO Standardization Agreement* (STANAG) 4294, Arinc Research Corporation, 2551 Riva Rd, Annapolis, MD, USA, Publ. 3659-01-01-4296, (1990).
[6] Mansouri R. and Sexl R.U., General Relativity and Gravitation **8**, 7, 497, (1977); General Relativity and Gravitation **8**, 7, 515, (1977); General Relativity and Gravitation **8**, 10, 809, (1977).
[7] P. Wolf, Ph.D. thesis, Queen Mary and Westfield College, London, (in press).
[8] Zumberge J.F., Liu R., Neilan R.E. ed., *IGS Annual report 1994*, JPL publication 95-18 9/95, Jet Propulsion Laboratory, Pasadena, CA, USA, (1995).
[9] Wisnia J.A., in Proceedings 24th PTTI, 199, NASA conference publication 3218, (1992).
[10] McCaskill T.B. et al., in Proceedings 23rd PTTI, 307, NASA conference publication 3159, (1991).
[11] Beutler et al., Manuscripta Geodetica **19**, 367, (1994).
[12] Tourenc P., Melliti T., Bosredon J., General Relativity and Gravitation **28**, 9, 1071, (1996).
[13] Fixsen D.J., Cheng E.S., Wilkinson D.T., Phys. Rev. Lett. **50**, 620, (1983).
[14] Lubin P.M., Epstein G.L., Smoot G.F., Phys. Rev. Lett. **50**, 616, (1983).
[15] Wolf P., Phys. Rev. A **51**, 5016, (1995).
[16] GPS world, May 1996, 50, (1996).

Limits on the anisotropy of c

Direct measurements:		
Krisher T.P. et al. (1990)	$\delta c/c < 3{,}5 \times 10^{-7}$	component in equatorial plane
GPS Test (this experiment)	$\delta c/c < 5 \times 10^{-9}$	all spatial directions
GPS Test (this experiment)	$\delta c/c < 2 \times 10^{-9}$	component in equatorial plane
Indirect measurements:		
Riis E. et al. (1987)	$\delta c/c < 3 \times 10^{-9}$	component in equatorial plane
Vessot R.F.C. et al. (1979)	$\delta c/c < 3 \times 10^{-9}$	component in one particular direction
Turner K.C. & Hill H.A. (1964)	$\delta c/c < 3 \times 10^{-8}$	all spatial directions

Tab. 1: Tests of the second postulate of special relativity showing the limits they set on anisotropy of c and their respective spatial sensitivities.

where Δ_0 is a constant characterising the initial phase difference of the two clocks. Einstein's second postulate requires that, for a series of measurements in different directions (e.g. during a complete passage of the satellite), T-D/c should remain constant, after correction for the relative rate of the two clocks due to the gravitational redshift, second order Doppler shift and the intrinsic (proper) frequency difference of the clocks.

In the theoretical framework generally used for the interpretation of experiments that test the second postulate of special relativity, the speed of light is anisotropic along a particular spatial direction (in an inertial frame) i.e. its component along this preferred axis is $c+\delta c$ in one direction and $c-\delta c$ in the opposite direction. The experiments then determine whether the special relativistic postulate $\delta c = 0$ is confirmed within the uncertainty of the experiment and set an upper limit on the anisotropy of the speed of light, that is, on the parameter $\delta c/c$. A more sophisticated theoretical approach to all tests of special relativity (including those testing the second postulate) was developed by Mansouri and Sexl [6]. A detailed interpretation of our experiment in this framework is presented elsewhere [7]; only the results are given here (see Sec. 6).

The effect of a possible anisotropy of magnitude $\delta c/c$ on an individual link would be $(\delta c/c)(D/c)\cos\alpha$, where α is the angle between the direction of the anisotropy and of the transmitted signal, resulting in a measurable variation of T as a function of direction. However, such an anisotropy might also affect the determination of the satellite ephemerides, and therefore the value of D, leaving the difference T-D/c unchanged. A meaningful test of special relativity using the above principle therefore requires a method of satellite orbit determination which is insensitive to a non-zero value of $(\delta c/c)$.

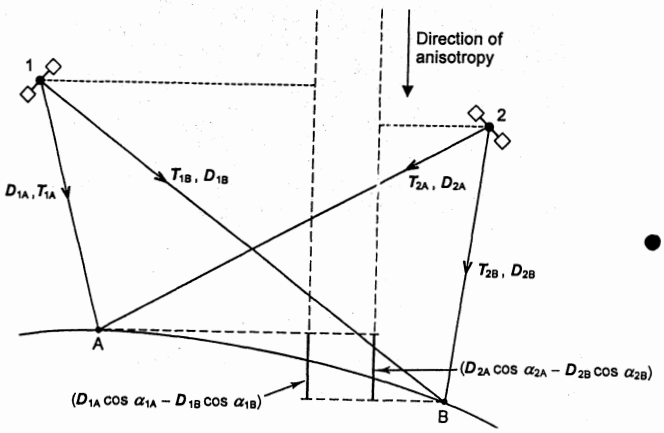


Fig. 1: Double difference (Y) for pairs of stations (A and B) and satellites (1 and 2), $Y=(T_{1A}-T_{1B})-(T_{2A}-T_{2B})$. The effect of an anisotropy on Y given by $[(D_{1A}\cos a_{1A}-D_{1B}\cos a_{1B})-(D_{2A}\cos a_{2A}-D_{2B}\cos a_{2B})]\delta c/c^2$, where α is the angle between the directions of the anisotropy and of the transmitted signal, vanishes.

This is the case for the GPS ephemerides obtained by the IGS-CODE processing centre [8]. The method used adjusts a post-Keplerian, non-relativistic orbit model to doubly differenced GPS timing data (see Fig. 1). The effects of anisotropy on the individual links cancel (to first order) when the double differences are formed (see Fig. 1). The IGS-CODE method is used to simultaneously adjust a number of parameters, including the satellite ephemerides and the ground station coordinates, thereby providing values of D which are unaffected by anisotropy of c.

Additionally one has to ensure that corrections applied to the raw timing data used for orbit determination and the measurement of T do not presuppose the second postulate. In fact two corrections are routinely applied to GPS timing data which are of relativistic origin and therefore do imply the isotropy of c [5]: the correction for the gravitational redshift and the second order Doppler shift of the rate of the satellite clock with respect to coordinate time, and the correction for the so called Sagnac effect which is due to the rotation of the Earth during signal transmission. Both of these are small corrections of order c^{-2} hence the effect of an error in these corrections due to an anisotropy would be negligible with respect to the first order effect on T (for a more detailed theoretical analysis see e.g. [12]).

Observation of GPS satellites in varying spatial directions thus provides a meaningful test of the second postulate of special relativity via relation (1) with D obtained using IGS-CODE ephemerides and station coordinates.

3. EXPERIMENTAL PROCEDURE

The IGS is a global network of ground stations that continuously observe the GPS satellites for civil, geodetic purposes [8]. From the raw data the IGS processing centres calculate (among other parameters) precise satellite ephemerides and ground station coordinates. These, together with the raw observations, are freely available through the internet via anonymous ftp [8]. We use data from eight ground stations for our experiment: Brussels (Belgium), Algonquin (Canada), Yellowknife (Canada), Fairbanks (Alaska, USA), Kokee Park (Hawaii, USA), Fortaleza (Brazil), Santiago (Chile) and Hobart (Australia). The motivation for this choice of ground stations is to ensure global coverage whilst providing maximum ground clock stability (for averaging times $\approx$ 6 h (one passage)) by using only stations which are equipped with hydrogen-maser clocks. The GPS receivers used are all AOA Rogue or Turbo-Rogue geodetic receivers providing raw phase measurements of the two GPS carrier frequencies at a sample interval of 30 seconds. The data sets cover six days (1994 September 18 to 23) and contain observations of all 25 GPS satellites available at the time. During this period (coinciding with the military intervention in Haiti) all GPS signals

were free of the intentional degradation (Selective Availability, SA) which is imposed by the US military. In general, this affects all but two satellites, making them unusable for the experiment described here. Of the 25 satellites used 19 are equipped with caesium clocks, and 6 with rubidium clocks.

From the raw data the differences *T-D/c* are formed, taking into account corrections for the variable part of the gravitational redshift and second order Doppler shift, the Sagnac effect, the ionospheric delay (using an ionosphere-free combination of the two frequencies), and the tropospheric delay (using the STANAG tropospheric model [5]). The amount of data is reduced by averaging over nine 30 s points, in order to save computer space and to make the data more manageable. Fig. 2 shows a typical example of a resulting data set for a station-satellite pair.

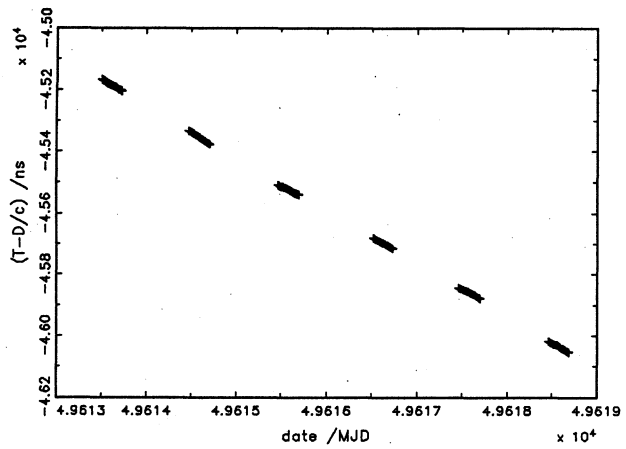


Fig. 2: Measurements of (*T-D/c*) for 6 passages of PRN04 over Algonquin(CAN).

For a test of the isotropy of *c*, one is interested in the variation of the difference *T-D/c* during individual passages of the satellite over the ground station i.e. variations over time scales of roughly 6 hours. Therefore we first filter the data, excluding all long-term variations (greater ≈ 5 days) effectively subtracting the relative rate of the two clocks. Then an arbitrary offset per passage is adjusted as only the *variation* of *T-D/c* during the passage is of interest.

As is well known, measurements of the GPS carrier phase are subject to an unknown phase ambiguity error of an integer number of cycles. This does not present a problem for our purposes as long as the induced error remains constant during each passage, which is the case if the receiver stays locked onto the satellite over the complete passage. Therefore all passages that are incomplete (data gaps indicating a possible loss of the satellite) were excluded.

The 0,5 sidereal day period of the GPS satellites implies that a station sees each passage of a particular GPS satellite at the same time of day (in sidereal days) and in the same directions (in a geocentric non-rotating frame). So, to see the data more clearly, all passages can be projected onto the same day, by shifting them individually by an integer number of sidereal days. Figure 3 shows the residuals of a typical data set after filtering and adjustment of an offset per passage, and with all passages shifted onto the same day. The standard deviation of these residuals for the complete data set (all stations and satellites) is 2,2 ns.

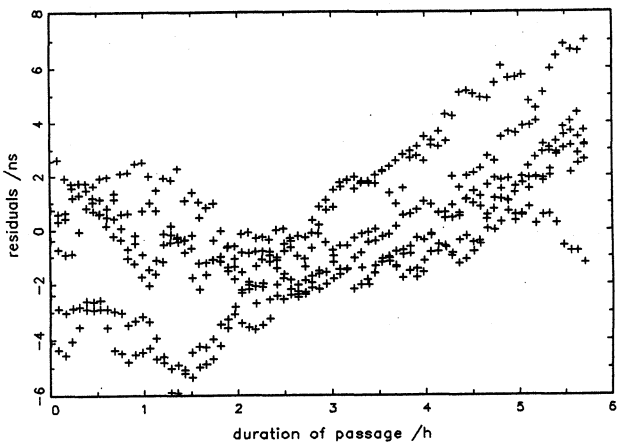


Fig. 3: Residuals of (*T-D/c*) after filtering and adjustment of an offset per passage. The graph shows 6 passages of PRN22 over Brussels(B) shifted onto the same day.

4. RESULTS

Figure 4 shows the spatial directions of signal transmission for the individual links in a non-rotating geocentric frame. There are no links at colatitudes below ≈20° and above ≈163° which is due to the 55° inclination of the satellite orbits and implies that the experiment was least sensitive in the N-S direction.

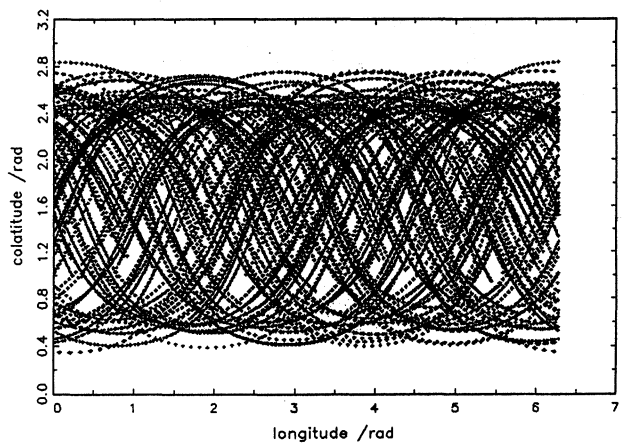


Fig. 4: Directions of signal transmission, in a geocentric non-rotating frame, for all satellite-station links.

The effect of an anisotropy of *c* on the transmission time *T* for a particular link is given by $(D/c)\ (\delta c/c)\ \cos\alpha$ where α is the angle between the direction of signal transmission and the direction of the anisotropy. This model was fitted to the data using the least squares method, adjusting the magnitude of the

anisotropy ($\delta c/c$) and an offset per satellite-station pair. The adjustment was performed for a range of anisotropy directions, spanning colatitudes and longitudes from 0 to π in a grid of 0,1 rad x 0,1 rad. It is sufficient to cover half of all possible spatial directions, as opposing directions correspond to the same anisotropy with the opposite sign. Directions are given in the non-rotating geocentric frame that is coincident with the ITRF Earth fixed frame at MJD 49755 (March 7, 1995) 0h 00 (UTC).

Figure 5 shows the adjusted magnitudes of the anisotropies as a function of their direction. The extremum value of ($\delta c/c$) is $4{,}9\text{x}10^{-9}$ at a colatitude of 2,9 rad and a longitude of 0,5 rad. In the equatorial plane (colatitude of 1,6 rad) the extremum value of ($\delta c/c$) is $1{,}6\text{x}10^{-9}$ at a longitude of 0,6 rad. The formal statistical standard uncertainties of these values are below $3\text{x}10^{-10}$.

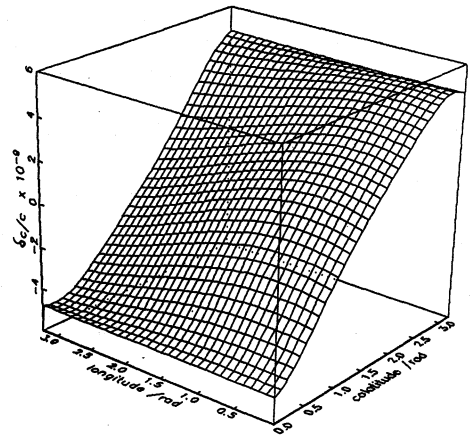


Fig. 5: Magnitude ($\delta c/c$) of the adjusted anisotropy as a function of its direction in a geocentric non-rotating frame.

5. SYSTEMATIC EFFECTS

The three main systematic effects that could affect the data are the satellite clock instability, ephemerides errors and uncertainties in the modelled tropospheric delays.

Pre-launch measurements of the relative frequency stability of the GPS caesium clock showed an instability of $\sigma_y(\tau) \approx 1{,}5 \times 10^{-13}$ (standard Allan deviation) for integration times of $\approx$ 6 h [9]. On-orbit measurements showed satellite clock instabilities of order 10^{-13} for integration times of 1 day [10], which corresponds to $\sigma_y(6\text{ h}) \approx 2 \times 10^{-13}$ when extrapolated assuming a $\tau^{-1/2}$ dependence. This translates into an accumulated time error over one passage of the satellite of $\delta\tau \approx \tau\sigma_y(\tau) \approx 4{,}3$ ns.

Beutler et al. [11] estimate the uncertainty of the IGS-CODE satellite ephemerides to be 15 cm to 20 cm corresponding to a timing error of $\delta\tau \approx 0{,}7$ ns.

The tropospheric delays were estimated using the standard STANAG model [5]. More accurate values for these delays can be obtained through estimations in a regional or global network (as done for example by the IGS). We have studied such estimates (four values per day and station) for a period of five days in September 1995 and for six of the eight stations used (no IGS estimates were available for the remaining two stations). The differences between the zenithal tropospheric delay obtained from the STANAG model and the 5 day averages of the IGS estimates do not exceed 200 ps for any station with the standard deviation of the averages < 110 ps. For low elevations these differences can increase to maximum 1,1 ns (at elevations of 10°). We conclude that, for our purposes, the STANAG model is sufficient as the dominant limitation of the sensitivity of our experiment is more likely to arise from satellite clock instabilities.

6. DISCUSSION AND CONCLUSION

Timing errors due to the systematic effects discussed in the previous section could give rise to values of $\delta c/c$ of order 10^{-8} ($\delta c/c \approx \delta\tau/(D/c)$) assuming that no cancellation takes place between the systematic effects of different satellite-station pairs. This is an order of magnitude larger than the maximum value observed. The experiment therefore cannot suggest a violation of the second postulate of special relativity.

It is unlikely that, in a global treatment, the systematic errors are correlated with the signature of an anisotropy, as they are expected, in general, to correlate differently with the effect of an anisotropy for each satellite-station pair and, to some extent, for each passage. So the results from a number of randomly chosen subsets of the data can provide an estimate of the reliability of the results presented in section 4. The least squares adjustment was repeated for five different subsets of the complete data, removing observations from 3 to 6 satellites (i.e. 12% to 24% of the total data). The mean of the five adjusted extremum values of $\delta c/c$ is $4{,}6\text{x}10^{-9}$ when considering all spatial directions and $2{,}2\text{x}10^{-9}$ for the component in the equatorial plane, with a standard deviation about the mean of $1{,}3\text{x}10^{-9}$ in both cases. These values should be compared to $\delta c/c$ of $4{,}9\text{x}10^{-9}$ and $1{,}6\text{x}10^{-9}$ obtained from the complete data set (c.f. section 4).

So, assuming no correlation (and resulting cancellation) between the systematic effects in a global treatment and the effect of a possible anisotropy (for the reasons mentioned above), we can set a limit of $\delta c/c < 5\text{x}10^{-9}$ on the spatial variation of c when considering all spatial directions and $\delta c/c < 2\text{x}10^{-9}$ for the component of the anisotropy that lies in the equatorial plane.

Interpretation of the experiment, using the test theory developed by Mansouri and Sexl [6], shows that the value of $\delta c/c$ obtained here is related to the parameter α of the test theory (in special relativity α = -1/2) by the well known relationship (see [7] for details)

A Caesium Flying Clock Experiment between NPL and USNO

John A Davis, James McA Steele*

National Physical Laboratory, Teddington, UK.
*Consultant to NPL.

ABSTRACT

1996 marked the 25th anniversary of a round the world flying clock experiment performed by Hafele and Keating of USNO, to test Einstein's theory of relativity. To mark this anniversary NPL in collaboration with BBC Horizon undertook a new flying clock experiment, between NPL and USNO using a portable HP5071A caesium standard. Flying caesium clock experiments are now rare, and hence this new experiment was an ideal opportunity to test the suitability of the HP5071A Cs clock as a flying standard. A detailed description is given of the predicted relativistic changes, and of the recorded clock changes, both on the trip to USNO and after the return to NPL and the resulting closure. The time scales of NPL and USNO are routinely compared using both GPS common-view time transfer and Two-Way Time Transfer. The flying clock experiment however provided a third independent method of comparing the two time scales. The results of these intercomparisons are presented.

Keywords: Time transfer, Relativity, Caesium clock

1. INTRODUCTION

In the first two weeks of October 1971 Hafele and Keating conducted their celebrated experiment (Refs. 1, 2, 3) in which four Hewlett-Packard Model 5061A caesium clocks were taken first East to West and then West to East round the world, making use of commercial airline flights. The objective was to provide convincing experimental evidence which would confirm the predictions of the General Theory of Relativity (GTR) in relation to time transfer in the vicinity of the (rotating) Earth and set to rest the controversies which had arisen on the subject. In this they were successful, the predictions being in good accord with experiment, as the results for mean of the four clocks, shown in Table 1, confirm.

Table 1
Results of Hafele - Keating experiment

	Relativistic clock gain (ns)	
	Eastward	Westward
Experimental results ± 1σ	-59 ± 10	273 ± 7
Prediction ± error estimate	40 ± 23	275 ± 21

In early 1996 the British Broadcasting Corporation (BBC) approached NPL with a proposal for a flying clock experiment, albeit on a markedly reduced scale as compared to a round-the-world exercise, which would figure in a 'Horizon' programme devoted to the wider and more speculative aspects of time travel. This programme would be broadcast close to the 25th anniversary of the Hafele-Keating experiment and would feature both of the original participants. A preliminary analysis was conducted based on a flight along a great circle path between the airports of Heathrow in the UK and Dulles in the USA. It indicated that a time transfer across the north Atlantic with a *single* clock from the latest generation of high-stability caesium standards (H-P Model 5071A with a high-performance option) should enable meaningful results to be achieved. In addition to being a demonstration of the principles involved in relativistic time transfer for the purposes of the Horizon programme it could also have scientific advantages in relating the time scales of the NPL and the US Naval Observatory (USNO), Washington DC. These are already connected on a routine basis by satellite time transfer using both GPS in 'common-view' and also by Two-Way Satellite Time and Frequency Transfer (TWSTFT) via the Intelsat communications satellite at 307°E. Both of these techniques have some limitations in accuracy at the level of a few nanoseconds and the possibility of an additional *ad hoc* time transfer between the two establishments could be useful in resolving any residual biases in the existing , radio-based, time transfers.

2. INITIAL ASSESSMENT OF THE EXPERIMENT

2.1 Relativistic corrections

The relations arising from the application of the General Theory of Relativity (GTR) to time transfer on the rotating Earth are well known and are codified in Recommendation ITU-R 1010 of the International Telecommunications Union on 'Relativistic Effects in a Coordinate Time System in the Vicinity of the Earth' for the exchange of a time reference by means of both portable clocks and electromagnetic signals (Ref. 4). With some very minor changes in nomenclature as compared to the ITU text the basic equation giving the coordinate time accumulated during a clock transport in a rotating reference frame, which is the appropriate choice for this experiment, can be written as,

$$dt = \int_{path} ds \left[1 - \Delta U(\mathbf{r})/c^2 + V^2/2c^2 \right] + 2\omega A_E / c^2 \qquad (1)$$

where dt is the elapsed coordinate time , $\int ds$ is the elapsed (proper) time on the portable clock, $\Delta U(\mathbf{r})$ is the gravitational potential difference (including the centrifugal potential) between the location of the clock at the terminus of the vector $\mathbf{r}$, with origin at the centre of the earth, and the geoid as viewed from an earth-centred and earth-fixed (ECEF) coordinate system, the convention being that ΔU is negative when the clock is above the geoid. V is the ground speed of the clock (aircraft), c is the speed of light, ω the angular velocity of rotation of the earth while A_E represents the equatorial projection of the area swept out during the transport of the clock along its path, A_E being

taken positive for motion in an easterly direction

The first term in (1) is the general gravitational frequency shift: for the potential differences arising in aircraft transfers ΔU can be approximated by g (θ) where the surface gravity, g , is slightly dependent on the latitude, θ , ie

$$g(\theta) = 9.7803 + 0.0519 \sin^2 \theta \ m/s^2 \qquad (2)$$

For a mean latitude on the flight path between Heathrow and Dulles of 45 degrees and an assumed height of 10 km the gravitational shift becomes,

$$g(\theta)/c^2 = 10.91. 10^{-13} \qquad (3)$$

The second term in (1) is the second-order Doppler shift. For an assumed aircraft speed of 250 m/s it amounts to ,

$$V^2/2\,c^2 = 3.48 .10^{-13} \qquad (4)$$

The third term is the Sagnac effect (Ref. 5) required to ensure consistent and transitive synchronization of clocks on the rotating Earth. It is independent of the speed of the transfer and depends only on the form and direction of the flight path. In the absence of prior detailed knowledge of the flight plan a great circle path was assumed for the purposes of assessment. The magnitude of this correction is,

$$2\,A_E/c^2 = 1.6227. 10^{-6} A_E \ ns/km^2 \qquad (5)$$

Before assembling the several corrections it will be convenient to rewrite Equation 1 in a form where the results will be directly comparable with those in Table 1, representing the relativistic gain of the flying clock with respect to coordinate time, ie

$$ds - dt = \text{clock gain} = \int_{path} [g(\theta)\, h/c^2 - V^2/2c^2] - 2\omega A_E/c^2 \qquad (6)$$

Substituting the appropriate values we obtain the results shown in Table 2

Table 2
Computed relativistic clock gain (ns).

Journey	Height	Speed	Sagnac	Total
E→W	+25.5	-8.1	+19.2	+36.6
W→E	+25.5	-8.1	-19.2	-1.8

The table entries are based on an estimated flight time of 6.5 hours at an altitude of 10 km and speed of 250 m/s on a great circle path from Heathrow Airport (51°28' N, 0°27' W) to Dulles Airport (38°55' N, 77° 30' W)

These results demonstrate, of course, the expected East-West asymmetry with the reversal of the Sagnac effect in the two directions of travel. The magnitude of the resulting time shift is appreciable in terms of existing standards of time-keeping and given corresponding stability in the flying clock provides the basis for a viable experiment.

2.2 Portable clock

The Hewlett-Packard 5071A caesium beam standard represents a very considerable advance both in stability and in the reduced effect of environmental changes as compared to the earlier 5061A clocks available to Hafele and Keating, a number of microprocessor-controlled servo loops being employed to ensure a large degree of independence of the frequency for changes in temperature, humidity and ambient magnetic field. The anticipated short-term stability of a 5071A clock *under laboratory conditions* is shown in the form of the Allan deviation in Figure 1 (Ref 6).

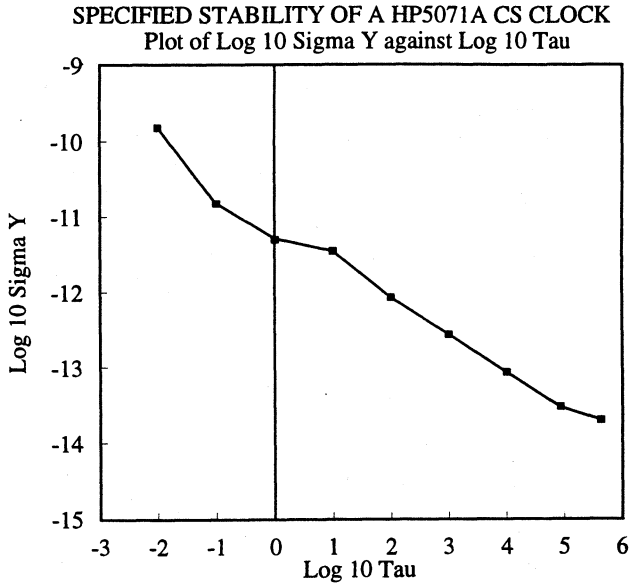


Figure 1 Allan deviation of 5071A-laboratory specification

The proposed experiment envisaged a stay of the flying clock at the USNO for a period of about 24 hours for robust comparison with the Observatory time scale followed by a return to NPL within a total time frame of 48 hours. Over most of this interval the specified clock performance in Figure 1 shows white noise frequency modulation with consequent random walk of phase (time). After an interval of 24 hours this amounts to 2.5 ns while closure with the NPL time scale after 48 hours should be achieved within about 3.5 ns. These figures served to reinforce the opinion that a viable time transfer operation could be mounted, enabling GTR predictions of clock behaviour to be verified at the 10 per cent level.

There remains the question of residual environmental effects. The portable clock will be exposed to fluctuating temperatures and humidity for brief periods during ground transport at each end of the path. In addition, for most of the duration of the flight the cabin pressure will be in the region of 700 mb. The relevant listed coefficients for the effects on frequency and on time are given in Table 3.

Table 3
Effects of temperature , humidity and pressure on the performance of a HP5071A Caesium clock.

Parameter	Frequency change parts in 10^{13}	Time excursion in (hours), ns
Temp. 0-55 C	3.0	1.1 (1 hour)
Humidity, 0-80%	3.0	1.1 (1 hour)
Pressure to 186 mb	<1	<2 (6 hours)

The ranges of parameter in the table are extreme and it would seem that perturbations in time of greater than 1 ns were not likely to arise from the combined effects of these variables. Some confirmation of this view is provided by measurements at the Tokyo Astronomical Observatory on the earlier 5061A caesium standard. These included a pressure coefficient of frequency of $+1.92.10^{-14}$ /100 mb and on the basis of this result the total effect due to pressure alone would amount to a loss in time of 1.2 ns over the flight duration.

2 Experiment of 15-17 June 1996

The results obtained in the preliminary simulation appeared to justify the intended experiment and at NPL. A 5071A caesium clock with the high-stability option was selected and adjusted to approximately zero rate relative to UTC (NPL) some days prior to the experiment. A plot of the performance of this clock, Serial No 784, in terms of a scale controlled by a hydrogen maser is shown in Figure 2 for an interval of 7 days.

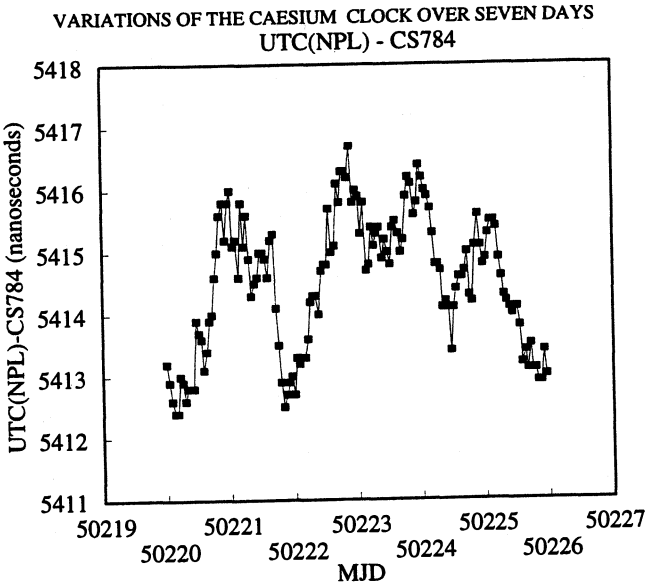


Figure 2 UTC(NPL) - Clock 784

Shortly before departure on 15 June the time indication of the portable clock (PC) was adjusted and set behind UTC(NPL) by 139 ns, ie, at 0800 UTC on 15 June 1996 [UTC(NPL) - PC] = 139 ns.

The clock was carried on scheduled services of British Airways and located together with the power and battery supply on a seat in the upper deck a Boeing 747, being connected to the aircraft power supply for the duration of the flights. The sequence of arrival and departures at various points in the itinerary is shown in Table 4.

Table 4
Timetable for the 1996 clock trip

Dates (1996)	Flight	NPL		H'row		Dulles		USNO
		dep		dep		arr		arr
June 15th	BA217	0800	→	1235	→	2010	→	2200 UTC
		arr		arr		dep		dep
June 16/17th	BA216	0700	←	0551	←	2300	←	1900 UTC

The actual flight times amounted to 7 h 35m westerly and 6h 51m easterly. There were undesirably long delays of several hours between leaving the respective establishments and seeing the clock finally installed in the aircraft, during which time it was operating on batteries and in a substantially unregulated environment. It was also subject to considerable manhandling, especially at Heathrow, where airport steps had be negotiated in the absence of the normal passenger entry port.

In the course of each flight data was provided, usually at intervals of 20-30 minutes, from the flight deck instrumentation on aircraft altitude, speed and position as obtained from the inertial navigator. During the outward journey the flight path was appreciably south of a great circle heading as a consequence of airways congestion at more northerly latitudes. By contrast, on the return journey, a significantly more northerly heading was followed.

On the basis of the navigational information the calculated relativistic time gains (and losses) for the portable clock for each of the contributing effects discussed earlier were calculated in near real time as the aircraft progressed. These values are plotted in Figures 3 and 4 while the overall results are summarized in Table 5. The short ground extensions substantially at sea level from airport to establishment in each country make a negligible contribution to the relativistic budget (< 0.2 ns for the Sagnac corrections).

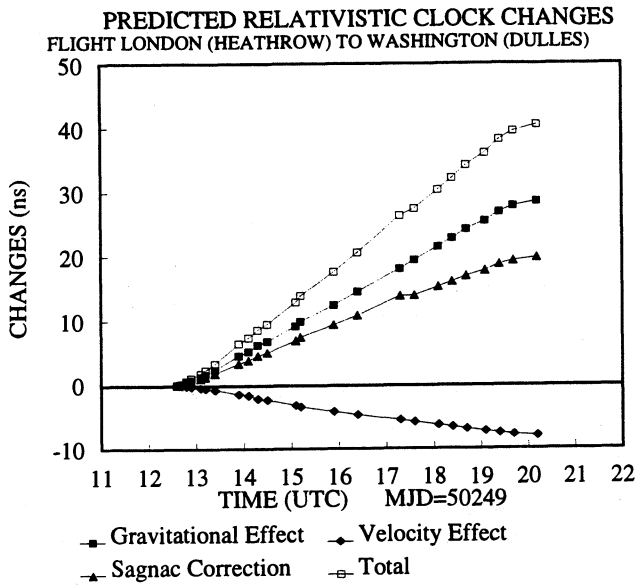


Figure 3 Calculated relativistic clock changes, East to West

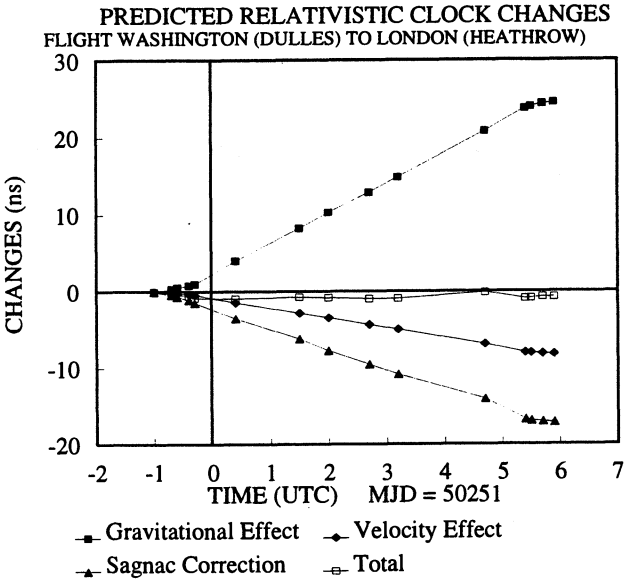


Figure 4 Calculated relativistic clock changes, West to East

Table 5
Computed clock gain

Path	Computed clock gain (ns) due to			Total (ns)
	Height	Speed	Sagnac	
E-W	+28.4	-8.0	+19.8	+40.4
W-E	+24.6	-8.1	-17.1	-0.6

It is not surprising that there are small differences between these results and the predictions appearing in Table 2. The flight time in both directions was greater than the estimated value of 6.5 hours, exceeding it by about an hour in the westward direction where also the operational height was 10.7 km for most of the flight, both factors combining to increase the gravitational shift. The more southerly route on the outward flight tended to increase the Sagnac contribution slightly while the northern track on the return journey had the opposite effect, reducing it by about 10 per cent.

3 NPL - USNO TIME TRANSFER

On arrival at the USNO the portable clock was compared directly with the one Pulse Per Second (1PPS) output from the Master Clock. The results with the GTR corrections applied are given in Table 6 both for the intercomparisons at Washington and Teddington. The exceedingly small closure error of the portable clock on return to NPL of less than 1 ns may be fortuitous but is nevertheless gratifying.

Table 6
Summary of time differences

Date	NPL-PC at NPL	NPL-PC +GTR	USNO-PC	NPL-USNO
June 15 MJD 50249	+139.0	+98.6	+87.5	+11.0
June 16 MJD 50250	-	-	+85.6*	+13.0
June 17 MJD 50251	+100.0**	+99.2**	-	-

*The portable clock was found to have gained 2ns relative to USNO (MC) during its stay.

**On return to NPL the difference between actual and predicted PC times was thus 0.8 ns.

The connection between UTC(NPL) and UTC(USNO), as has been mentioned, is also established by other means. Both institutions contribute to the formation of TAI (International Atomic Time) by the Bueau International des Poids et Mesures (BIPM) and the BIPM monthly Circular T gives, with some delay, the time settings of the contributing clocks relative to UTC. A somewhat more direct and immediate comparison is provided by GPS common-view observations and the table contains these results, derived from the almost 'raw' observations at NPL and USNO but filtered to some extent by a rolling mean of adjacent readings embracing ± 3 days. Finally, NPL and USNO are connected by two-way time transfer which enables a precision of about 1 ns to be obtained three times per week. However, due to a lack of reciprocity in the satellite transponder this can not yet be translated into a corresponding accuracy. These results and those given by the other methods are included in Table 7

Table 7
Time transfers NPL - USNO (ns)

Dates 1996	BIPM Circ T	GPS +filter	TWSTFT	PC
June14th MJD 50248	-	8.5	28.5	-
June 15th MJD 50249	-27	7.6	-	11.0
June 16th MJD 50250	-	6.5	-	13.0
June 17th MJD 50251	-	6.0	29.0	-
June 20th MJD 50254	-26	8.4	-	-

The filtered GPS time transfers and portable clock transfers agreed within experimental error. The TWSTFT results show a discrepancy of 15 ns when compared with the portable clock trip time transfer. This is in part due to an unknown transponder delay asymmetry occurring within the Intelsat satellite. However, this link was calibrated in 1994 using a portable X-band TWSTFT station operating from a satellite free from transponder delay asymmetries. This calibration has been incorporated into the present results.

The differences between the NPL-USNO time transfers determined for circular T and determined via the other methods is surprisingly large. It is proposed to resolve this discrepancy in the near future using the BIPM calibrating GPS receiver.

4) CONCLUSIONS

The flying of a HP5071A caesium clock from NPL to USNO and back provided a useful test of general relativity at the 10% uncertainty level. The use of the HP5071A caesium as a portable transfer standard has been shown to be useful in the synchronization of atomic time scales at primary timing laboratories.

5. ACKNOWLEDGEMENTS

We are indebted to the BBC, the Horizon team, and the producer Judith Bunting for initiating and supporting this enterprise through its many stages. We are grateful also to British Airways for facilitating the time transfer, both in the air and at the terminals and especially to Captain Morris Butler and his colleagues on the flight desk for the provision of navigational information at frequent intervals on both flights. We thank Hewlett Packard for the loan of a HP5089A power supply.

We also thank our colleagues at USNO, in particular James DeYoung and Bill Klepczynski both for their support and assistance with the project and for their hospitality during our visit to USNO.

This work has been supported under the UK National Measurement System programme for time and frequency.

6. REFERENCES

1. Around-the-world Atomic Clocks: Predicted Relativistic Time Gains, J C Hafele and R E Keating, Science, Vol 177, pp 166-168.

2. Around-the-world Atomic Clocks: Observed Relativistic Time Gains, J C Hafele and R E Keating, Science, Vol 177, pp 168-170.

3. Performance and Results of Portable Clocks in Aircrafts, J C Hafele, Proceedings of the 3rd PTTI, 16-18Nov 1971 ,pp 261-284.

4. Relativistic effects in a coordinate time system in the vicinity of the earth. ITU Radiocommunication Study Groups, Document 7/45(Rev.1)-E, 27 Oct 1993.

5. Around the world relativistic sagnac experiment, D W Allan, M A Weiss and N Ashby, Science, vol 228 pp. 69-70 Apr. 1985

6. HP5071A Primary Frequency Standard operating manual, section 6-3.

DORIS PRECISE ORBIT DETERMINATION and LOCALIZATION
Performances in orbit and preliminary studies of DORIS-NG

M. Brunet, A. Auriol, J.P. Berthias, M. Costes, J.L. Issler, P. Sengenes
CNES - Centre National d'Etudes Spatiales
18, avenue Edouard Belin, Toulouse - France

ABSTRACT

The French Space Agency C.N.E.S. (Centre National d'Etudes Spatiales), G.R.G.S. (Groupe de Recherche en Géodésie Spatiale), I.G.N. (Institut Geographique National) have designed and developed DORIS (Doppler Orbitographic and Radiopositioning Integrated by Satellite) able to determine within few centimeters, the position of satellites and ground beacons.

DORIS payload has been integrated on French SPOT satellites, on TOPEX NASA satellite, and will be embarked on ENVISAT ESA and JASON satellites.
Based on one way doppler very accurate measurements from a network of 50 ground beacons, localization performances depends strongly on the stability of on-board and ground USOs.

The paper describes elements of the DORIS system, and gives last performances in orbit computation, localization of ground beacons, and USOs performances.
Preliminary relexions of an DORIS-New Generation (DORIS-NG) are also given.

1. THE DORIS SYSTEM

1.1. The Doris system components (figure 1)

The Doris system comprises a "flight segment" composed of a receiver, an USO and a dual frequency omnidirectionnal antenna, a "ground segment" including the control, command and data processing center, and a network of Orbitography Determination Beacons (ODBs) and Ground Location Beacons (GLBs). GLBs are functionally identical to ODBs, but field packaged.

1.2. The flight segment

The on-board DORIS instrument comprises :
- A fixed omnidirectional dual frequency antenna ;
- A receiver performing the Doppler measurements every ten seconds. Current receivers perform measurements over one beacon at a time. The nominal functioning mode is a programmed mode in which it tracks the beacons signals according to a preplanned work table which is uploaded daily ;
- An Ultra Stable Oscillator (U.S.O.) delivering a very stable frequency reference.

DORIS instrument has been integrated on SPOT2 and SPOT3 French satellite in orbit since february 1990 and october 1993, on TOPEX POSEIDON NASA satellite in orbit since september 1992, and on SPOT4 French satellite which will be launched in february 1998. DORIS instrument will be also integrated on Envisat ESA satellite which will launch in 1999-2000. SPOT, ENVISAT and TOPEX satellites are on circular orbits at an altitude of 830 kms and 1300 kms respectively.

1.3. The ground segment

The raw data are received from satellite ground segment by the DORIS control and processing center located in Toulouse, France and operated by C.L.S. company, a CNES subsidiary.
The DORIS control and processing center performs :
- Acquisition of telemetry raw data
- Ground beacons and on-board instrument monitoring
- On board instrument programmation
- Measurements time-tagging and pre-processing
- Operational orbit determination
- Beacon precise positionning
- Data monitoring and archiving

The CNES Service d'Orbitography DORIS (S.O.D.) performs the precise orbit determination and evaluation.The ground segment includes also a network of 50 Orbitography Determination Beacons (ODBs) (see figure 2) installed in very accurate known place (geodetic points). They are used to calculate continuously the orbit parameters and are the reference system for the location determination. Installation and maintenance of the network is realized by I.G.N.
The location of each beacon is precisely linked to the International Reference Frame (IERS) to get absolute positionning. The frequency and time reference for the DORIS system is provided by two master beacons located in Toulouse and Kourou french Guyana, which are drived by an USO slaved to an atomic frequency standard. The time scale reference is TAI to avoid leap seconds.

ODBs have 4 units :
- A main unit including an USO whose performances are simular to the on-board USO, driving 2 transmitters at 2.03625 GHz and 401.25 MHz with power of 5 watts ;
- An omnidirectional dual frequency antenna ;
- A battery back up unit to provide autonomy while main AC supply is off ;
- A meteorological package providing pressure, temperature and humidity measurements which are transmitted to the up link to allow tropospheric correction to the measurements.

Ground Location Beacons (GLBs) are also installed according the different users needs.
GLBs have same units as ODBs, but are packaged and equiped with autonomus power supply (solar array, ...).

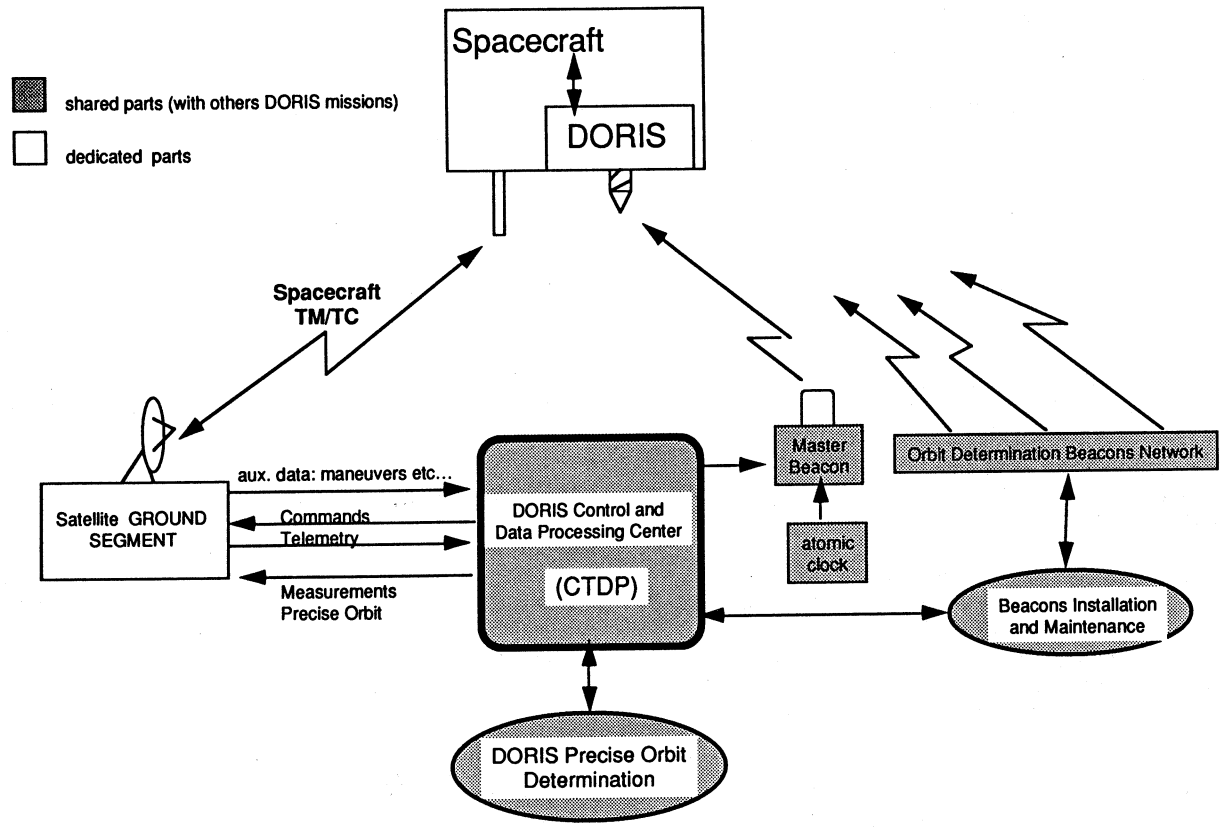


Figure 1 : DORIS system components

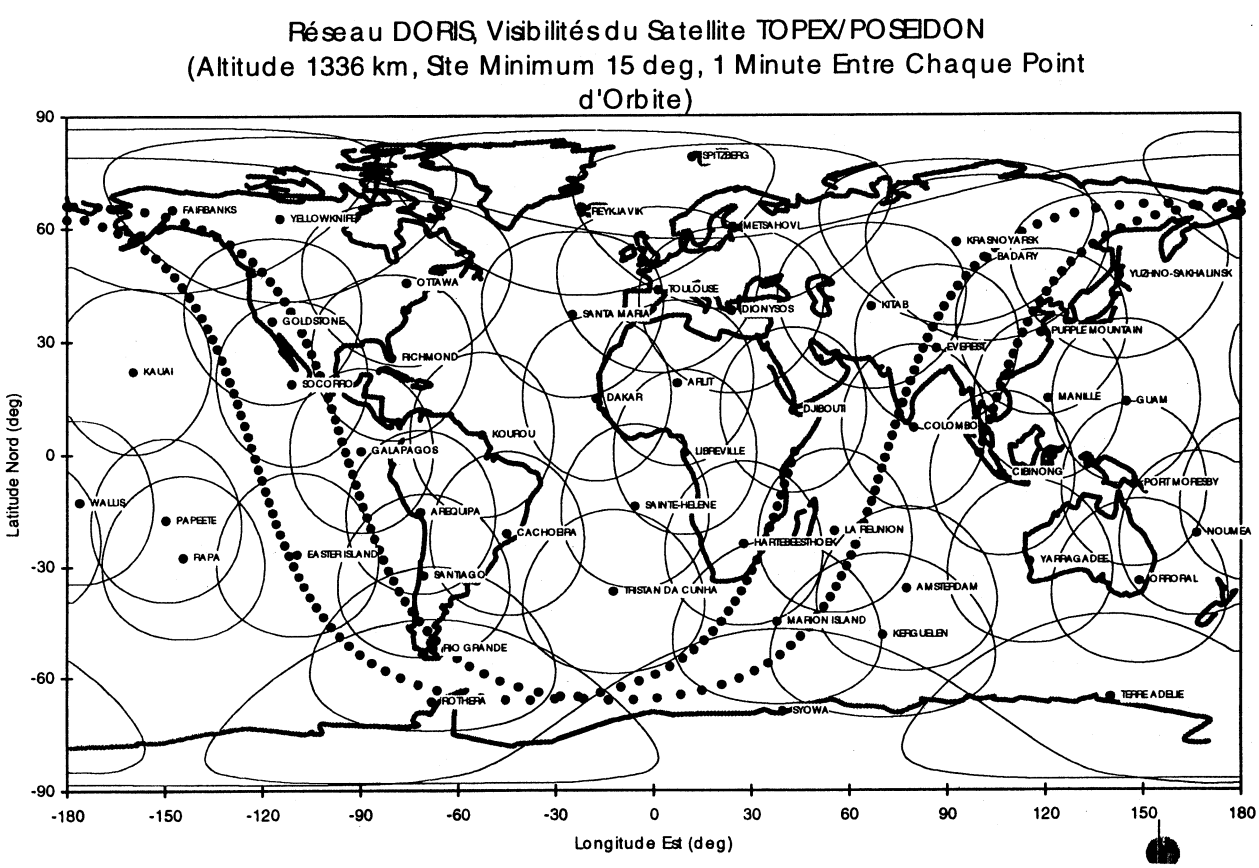


Figure 2 : DORIS network (March 1993 and 15° visibility circles for T/P

2. DORIS PERFORMANCES

2.1. Instrument performances

The main characteristics and performances of the on-board instrument are summarised as follow :

- Total mass of the units composing the instrument (antenna, receiver, USO) : 15 kg.
- Power consumption : 20 W.
- On-board time-tagging accuracy : 2 μs.
- Doppler frequency measurement precision : standard deviation 0.5 mm/s or $1.6.10^{-12}$. This figure includes effects of ground and on-board instruments (USOs, ground transmitter, on-board receiver, antennas) multipaths and atmospheric propagation effects. This standard deviation is composed of :

  1) Short terme noise (rms of 10 seconds measurements) coming from on-board and ground oscillators, thermal noise, quantization and other instrumental errors. This noise is easily shown in the orbit residuals, since it has only high frequencies, and may be isolated by filtering using GUIER. Some very quiet pass show a 0.20 mm/s runs ($6.6.10^{-13}$).

  2) The medium term instrument errors are due mainly, to the natural noise of oscillators (Allan variance of Doris USOs vary from 5.10^{-13} over 10 s to 1.10^{-12} over 10 minutes), and the tropospheric propagation mismodelling during the satellite pass. Vertical tropospheric effects are about 2.4 meters, estimated with an accuracy of 2-3 cm from the three meteorological ground sensors.

2.2. Coverage and availability

DORIS-SPOT2 is in use since February 1990. DORIS-TOPEX since September 1992, and DORIS-SPOT3 since October 1993.
The actual coverage (percentage of time during which the instrument is doing measurements) is 65% for SPOT altitude (830 km) with a minimum elevation of 12° and 80% for TOPEX altitude (1300 km) with a minimum elevation of 15°.
The relability of the system is about 90% in term of Doppler measurement production. Data losses are mainly due to Single Event Upsets which necessitate re-initialization of the on-board instrument programming. Nevertheless, thanks to short data outages, TOPEX precise orbit has been produced 100% on the time since the beginning of the mission. Furthmore, new receivers will use a SEU less sensitive processor.

2.3. Precise orbit determination performances

DORIS data is used routinely to produce high precision orbits for SPOT and TOPEX. To achieve this goal, software developments have been conducted by CNES over the last 15 years. This involved large scale improvement in the quality of models. Such progress were made possible by a collaborative effort with the Space Geodesy Branch of the GSFC and the Center for Space Research of the University of Texas at Austin.

As a result there is now a general conscensus on the fact that the radial component of the T/P orbit is known with an accuracy of about 3 cm rms. This figure is established using various validation techniques of which the use of external data is the most powerful. Satellite laser tracking residuals (fig 3) are at the 3 cm level ms. In addition the comparison of the US and French precise orbits shows that the surface force model is the 1.5 cm level (fig 4) [The US and French orbits use the same gravity field, and this gravity errors do not show up in orbit comparison plots]. Additional comparisons with GPS orbits also show an agreement at the 1 to 2 cm level Reference [7, 8].

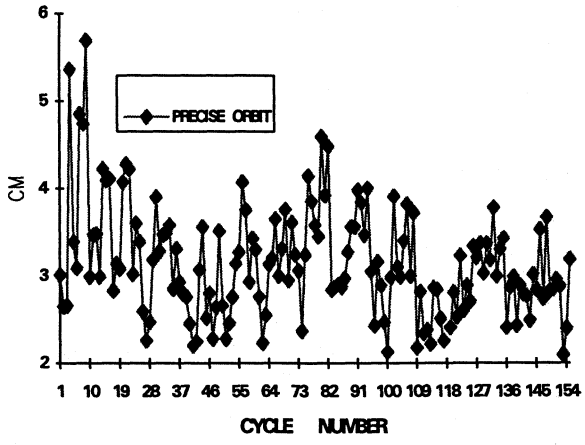


Figure 3 : TP laser residuals versus time (1 cycle = 10 days)

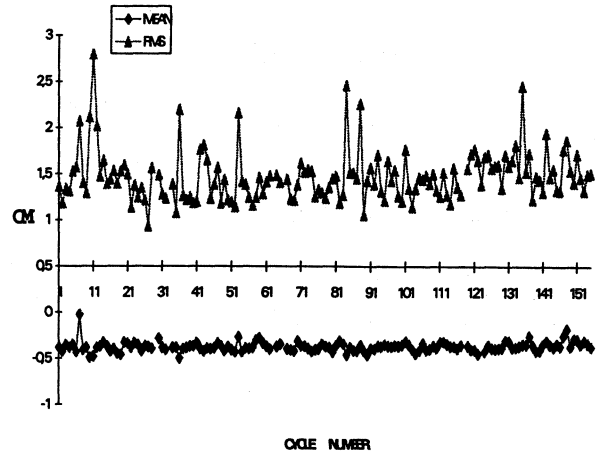


Figure 4 : TP precise orbit radial differences CNES-NASA using DORIS data

Recent progress has been made on the delay at which precise ephemeris products can be obtained. If the high precision processing still takes up to 5 weeks to compute and validate the orbit, faster orbit determination chains now deliver ephemerides with better than 5 cm ms radial error in less than 48 hours. These orbit products receive only minimal validation, but their high speed of delivery makes them of interest for scientists involved in operational processes.

2.4. Real time on-board orbit determination

A real time on-board orbit determination software (DIODE) has been developped by CNES to process the DORIS measurements on-board and complete the spacecraft position with a 10 sec. time step.

This software will fly on SPOT4 and on all the new DORIS instruments (ENVISAT, JASON, ...). The accuracy achieved with the latest version of the software is better than 1 meter 3D (1 sigma). This real time product completes the set of DORIS orbits as shown in figure 5. This DORIS products span all possibilities from real-time to 5 weeks delayed, with according levels of accuracy.

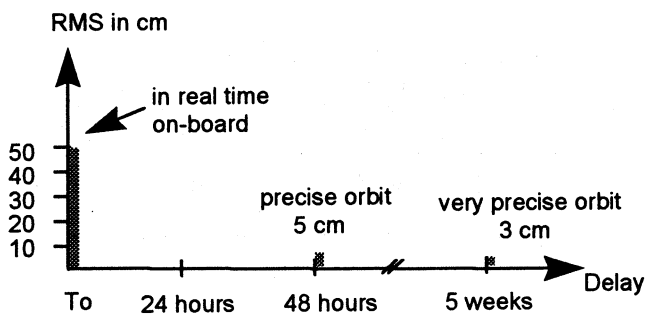


Figure 5 : DORIS orbit accuracy in cm versus delay

2.5. Precise Positioning Ground Beacons performances

DORIS is particulary well suited for global kinematic studies of large scale tectonic motions. 19 DORIS beacons collocated with reference from other tracking systems (laser, GPS, VLBI) are installed on 9 major tectonic plates. Each beacons is located using DORIS data since 1990.

Figure 6 from [4] shows the interplate velocities of the DORIS stations on the African plate relatively to the others plates. 40 to 50 mm per year are observed for these stations.

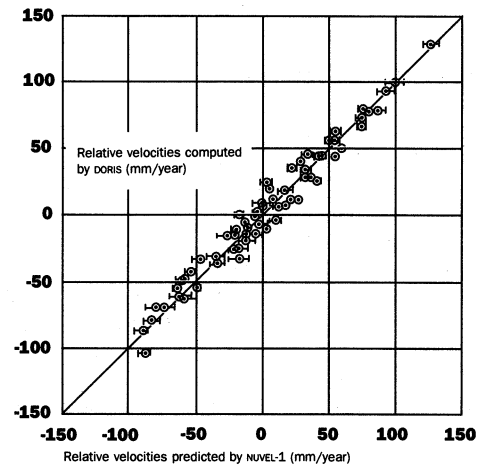


Figure 6 : Interplate velocities of the stations on the African plate relative to the stations on the Eurasian, Pacific, North American, South American, Indian and Australian plates, computed with DORIS and compared to NUVEL-1. L. Soudarin and A. Cazenave from GRGS have estimated the relative rates in millimeters per year of several tectonics plates.

2.6. Polar motion determination with DORIS data

The DORIS system has been used also for polar motion determination by I.E.R.S. on regular basis. Figure 7 shows the polar motion from 1990 to 1996, clusters of points representing the DORIS data.

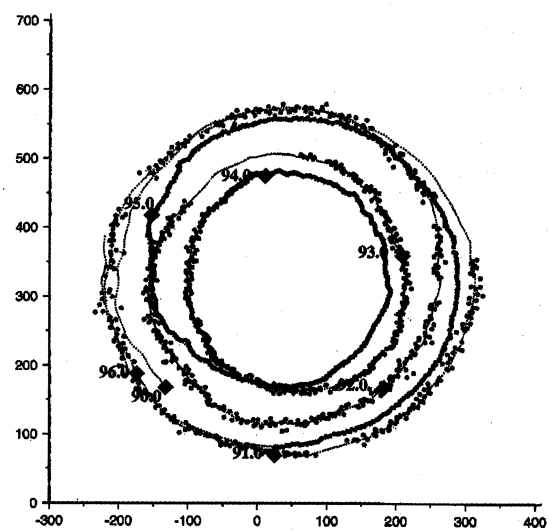


Figure 7 : Polar motion from 1990 to 1996 by IERS using DORIS data

3. DORIS USOs SPECIFICATIONS AND PERFORMANCES

3.1. DORIS USO Specifications

The main characteristics required for the on-board (OB) and ground USOs are defined in the following Table 1.

Parameters	OB. USO	ODB. USO	GLB. USO
Frequency	10 MHz	5 MHz	5 MHz
DC. Supply variation	$\pm 1.10^{-11}$ (1%)	$\pm 5.10^{-11}$ (1%)	$\pm 5.10^{-11}$ (1%)
Load variation	$\pm 1.10^{-10}$ (10%)	$\pm 2.10^{-11}$ (10%)	$\pm 2.10^{-11}$ (10%)
Pressure variation	$\pm 1.10^{-8}$/bar	$\pm 1.10^{-12}$/mb	$\pm 1.10^{-12}$/mb
Acceleration variation	$\pm 1.10^{-9}$/g	10^{-9}/g	10^{-9}/g
Vibrations (sine, random)	$\pm 5.10^{-9}$	-	-
Magnetic variation	$\pm 5.10^{-11}$/gauss	-	-
Temperature variation	$\pm 2.10^{-10}$ (air, vacuum 0 to 40°c)	$\pm 2.10^{-10}$ (0 to 60°c)	$\pm 1.10^{-10}$ (-20 to 60°c)
Short term stability τ = 10 s (Picinbono variance)	5.10^{-13}	5.10^{-13}	5.10^{-13}
Doris mid term stability			
with θ = 0.1°C/mn	0 to 40°C	0 to 60°C	-20 to 60°C
R.M.S. SLOPE	$\pm 4.10^{-13}$/mm	$\pm 4.10^{-13}$/mm	$\pm 4.10^{-13}$/mm
Residual (1σ)	1.10^{-12}	1.10^{-12}	1.10^{-12}
Long term aging after 1 month			
1 day	$\pm 1.10^{-10}$	$\pm 3.10^{-11}$	$\pm 3.10^{-11}$
1 month	$\pm 1.10^{-9}$	$\pm 5.10^{-10}$	$\pm 5.10^{-10}$
Phase Noise dBc/Hz			
10 Hz	- 120	- 130	- 130
1 KHz	- 140	- 145	- 145
DC. Power			
air, 0°C	6 W	6 W	6 W
Vacuum, 0°C	3,5W	-	-
Power supply	20 VDC	24 VDC	24 VDC
Random vibration	18 grms	-	-
Sine vibration	20 g - 10 to 100 Hz	2 mm 5 - 25 Hz	2 mm 5 - 25 Hz
Radiations	TOPEX	-	-

Table 1 : DORIS USO SPECIFICATIONS

The very important characteristics of one way Doppler technique are :

1) The short term stability, expressed in Allan variance : the initial specification is $\sigma_y \leq 5.10^{-13}$ $\tau = 10$ s, and $\sigma_y \leq 3.10^{-13}$ for the ODBs 2e generation.

2) The medium term stability over the duration of the pass (15 minutes), which is defined for DORIS by slopes S (min^{-1}) and residuals R at 1 sigma, of 90 samples of USO frequency measurements made over 10 seconds every 10 seconds. S and R are calculated by the least square method, and must be :

	USOs specifications	
	OB and ODG 1e generation	ODB 2e generation
Slopes S	$\pm 4.10^{-13}$/min	$\pm 1{,}5.10^{-13}$/min
Residuals R	1.10^{-12}	3.10^{-13}

Each on board and ground USOs must be compliant to theses specifications, and need very low short term stability and very low sensitivity to all environmental conditions : temperature, magnetic field, radiation in orbit ; temperature pressure, humidity on the ground.

3.2. DORIS USO procurement

On board flight model USOs were procured from FEI (USA) for DORIS-SPOT2, then from THOMSON-CEPE (France) for DORIS-SPOT3, SPOT4, Topex, Envisat. CEPE use one precise thermal control gigogne structure, and BVA-4 (also called QAS) space qualified swept SC cut crystal resonator.

ODB and GLB USOs (100 units including spare models) were provided from Oscilloquartz (Switzerland). Models used are 8600.03 type, using BVA AT 5 MHz crystals, and 2 ovens and Dewar techniques.

The 50 ODBs network must be replaced by new packaging and will be equiped by better performances USOs (from CEPE and Oscilloquartz).

3.3. DORIS USOs performances

All OB and ODB-GLB USOs are tested in detail, first at the suppliers, then at the CNES Time and Frequency lab. and during all equipment and satellite test in near real life conditions.

Table 2 gives typical performances of OB and ODB USOs

Parameter	OB USO (CEPE)	GLB USO (OSA)
Short term $\sigma y(\tau)$ $\tau = 10s$	3.10^{-13}	3.10^{-13}
Doris medium term with gradient		
Slopes	$\pm 5.10^{-13}$/min	$\pm 5.10^{-13}$/min
Residual	1.10^{-12}	1.10^{-12}
Aging	1.10^{-10}/day after 40 days	2.10^{-11}/day after 30 days
Phase Noise dBc/Hz		
10 Hz	- 142	- 142
1 KHz	- 155	- 155
Temperature Sensitivity	1.10^{-10}pp - 10 to +50°C in vacuum	5.10^{-11}pp - 20 to +55°C in air
Magnetic Sensitivity	0.2 to 2.10^{-11}/gauss	NA
DC power at 20°C		
in air	6.5 W	2.4 W
in vacuum	2.0 W	NA

Table 2 : Typical DORIS USO performances

3.4. DORIS USO performances via orbit determination

CNES-SOD processes on-board data using the "ZOOM" software ref [4]. Several parameters are ajusted :
- orbital elements
- dynamic parameters
- frequency differences between OB and ODBs, for each pass.

The absolute frequency differences are determined when each satellite passes are over Master beacons of Toulouse and Kourou, because the frequency of these master beacon is given by a cesium atomic clock and known within 1.10^{-13} or less.

SOD calculates absolute frequency of OB and each ODB and GLB USOs. The accuracy of this determination is estimated to 4.10^{-12} (1 sigma).

The following table 3 give the drift in orbit of DORIS SPOT2, Topex and SPOT3 USOs.

	SPOT2	TOPEX	SPOT3
after 1 month	$+1.1.10^{-11}$/day	$+3.3.10^{-11}$/day	$+8.10^{-11}$/day
after 1 year	$+8.10^{-12}$/day	$-4.5.10^{-11}$/day	$-2.1.10^{-11}$/day
after 2 years	$+6.10^{-12}$/day	-3.10^{-12}/day	$-1.1.10^{-11}$/day
after 3 years	$+610^{-12}$/day	$-2.8.10^{-12}$/day	$-1.0.10^{-11}$/day
after 4 years	$+6.10^{-12}$/day	$-2.6.10^{-12}$/day	"
after 6 years	$+6.10^{-12}$/day		"

Table 3 : In orbit DORIS USO drift

4 . FROM DORIS FIRST GENERATION TO DORIS NEW GENERATION

Since the first flight on the SPOT2 satellite in February 1990 of a DORIS on-board equipment, several improvements have been implemented on the later instruments. The DORIS second generation on-board instrument which will be accommodated on SPOT4 and ENVISAT satellites, will present as main improvements the hereafter capabilities :

- A new software, called DIODE, is integrated to the DORIS instrument software to allow the on-board processing of Doppler measurements and achieves a real time estimation of the satellite trajectory. This software, developed in ADA by CNES, will perform orbit determination with an accuracy better than 1 m rms along each direction of the orbital reference frame.
- The second generation DORIS instrument which will be on-board the ENVISAT satellite, will be able to provide simultaneous Doppler measurements for two ground beacons : improvements in orbit determination and/or ground beacon localization are expected.

4.1 « Limits » of the current DORIS system

In spite of its great performances and the improvements made on the second generation instruments, November 1994 DORIS workshop pointed out some drawbacks or limitations of the current DORIS system :

- Limitation on the maximum number of ground beacons that can be tracked at the same time (one for the first generation equipment, two for the second generation), which is a handicap in the relative positioning of colocated beacons.
- Conflicts between the functions of orbit determination (that requires to track the Orbitography Determination Beacons of the permanent DORIS network) and localization of Ground Location Beacons (« customer » beacons).
- Necessity to schedule the emitting periods of each beacon in order to be sure that the on-board DORIS equipment will track the wanted beacon.
- Narrow bandwidth signals sensitivity to radioelectric interferences (inter mobiles radio links).
- On-board equipment time reference scale synchronisation with international time reference scales (TAI, for instance) offers an accuracy of a few microseconds when new space applications require an on-board time synchronisation with an accuracy of a few nanoseconds.
- Navigation of geostationary satellites : beacons transmitting a narow band signal at the same frequency cannot been separed by the Doppler division multiple acces (D.D.M.A.), as it can be in LEO.

In order to allow DORIS to be competitive with some other systems which can be used to achieve the same functions on-board a satellite (GPS, for instance), it was noted that a strong reduction of mass, size, power consumption and cost of the on-board equipment was required, without increasing the cost of ground beacons or the cost of system operation.

4.2 DORIS New Generation

In order to bring some significant improvements to the current DORIS system on the points discussed in the previous paragraph, CNES decided to start the study of a new system called « DORIS New Generation ». One of the main changes from the current system comes from the use of spread spectrum techniques in the DORIS NG beacons signals which offers a better protection against jamming. The two DORIS NG carrier frequencies - band L and band S instead of 400 MHz / 2 GHz for the current system - are modulated (two states phase modulation) by the sum of a pseudo-random noise code bits and of the beacon message chips. As each ground beacon transmits its own PRN code, several beacons can be tracked at the same time for each type of orbit (LEO, GEO, GTO) - DORIS NG on-board receiver are currently specified to be able to get measurements simultaneously on fourteen beacons -, so the beacons emission scheduling is no longer required.

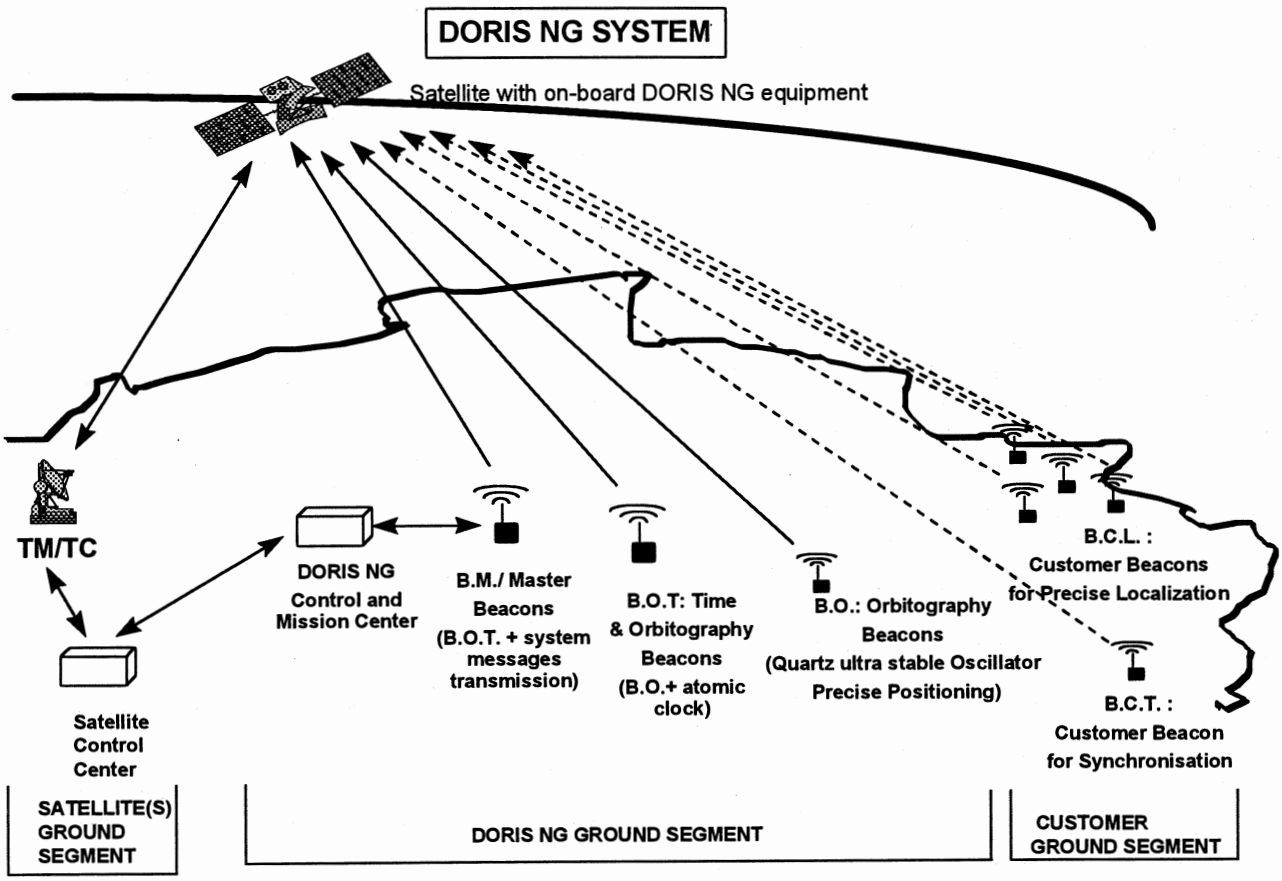
The use of PRN codes provides the DORIS NG on-board equipment the ability to realize pseudo-range measurements with an expected precision of a few meters : this opens the way to the synchronisation of on-board time reference with an accuracy of a few nanoseconds. As it is possible to get pseudo-range measurements on several beacons which are in view of the satellite, DORIS NG can be used to perform the orbit determination of geostationnary satellites.

The development of the on-board equipment is based on the use of digital signal processing techniques in order to reduce the mass, the power consumption and the cost of the on-board receiver and to protect the DORIS system against the risk that the technologies it uses become obsolete.

A brief description of the DORIS NG system is given on the hereafter figure. Two main sets of ground beacons are considered :

- On one hand, the beacons which form the Orbitography and Time permanent network. This network includes the Orbitography Beacons (BO), the Time and Orbitography Beacons (BOT) and the Master Beacons (BM). The positions of all these beacons are known with a very high accuracy, all these beacons use a quartz ultra stable oscillator to allow the realization of Doppler measurement with a very high accuracy : the short term error on Doppler measurement is expected to be of 0,1 mm/sec (1σ Picinbono). Doppler measurements performed on the permanent network beacons are used for the orbit determination of Low Earth Orbit satellites. The DORIS NG system time reference is broadcasted by the Master Beacon in Toulouse ; the shift between this time reference and the TAI is monitored by the DORIS NG Control Center and the value of this shift is transmitted to the DORIS NG on-board equipment via the Master Beacon message. The frequency of BOT USOs are controlled by the frequency delivered by an atomic clock in order to obtain a good long term stability of the BOT time reference. The shift between each BOT time reference and the DORIS NG system time reference is determined by the DORIS NG Control Center from the pseudo range measurements performed on these beacons, then broadcasted via the Master Beacon message.
- On the other hand, the « customer » beacon which are to be located or synchronized with respect to the DORIS NG time reference by the system.

The development of the DORIS NG system is currently under study and the first in-flight evaluation should be performed in early 2000.



Description of DORIS NG system

Acknowledgements

The authors are grateful to many collegues, especially A. Piuzzi (Orbit Metroloy Department) and J.F. Dutrey, J.B. Laporte (Time and Frequency Department) for all results, and valuable discussions.

REFERENCES :

[1] M. Dorrer, B. Laborde, P. Deschamps
Doris : System assessment results with Doris on SPOT2
IAF 90-336-Dresden - Germany, October 1990

[2] F. Nouel
Precise orbit determination of SPOT platform with DORIS AAS-AIAA conference paper 91-409, 1991

[3] L. Soudarin, A. Cazenave
Large scale tectonic plate notions measured with the Doris space geodesy system
Geophysical Research letters, vol 22, n°4, feb 1995

[4] F. Nouel et al
Precise CNES orbits for TOPEX-POSEIDON : is reaching 2 cm still a challenge ?
Journal of Geophysical research TOPEX-POSEIDON special issue, oct 1993

[5] M. Brunet
Doris precise orbit determination and localization system Performances of Ultra Stable Oscillators
6th EFTF, march 1992

[6] M. Brunet et al
Doris precise orbit determination and localization system description and USOs performances
IEEE Frequency Control Symposium 1995

[7] B.D. Tapley et al
Precision orbit determination for TOPEX POSEIDON
Journal of Geophysical research vol 99, december 1994

[8] W.I. Bertiger et al
GPS precise tracking of TOPEX/POSEIDON
Journal of Geophysical research vol 99, december 1994

GPS APPLICATIONS

Chairman: S. Feltham

The Architecture of WAAS Network Time (WNT)

W. J. Klepczynski
G. M. R. Winkler

Innovative Solutions International
8500 Leesburg Pike, Suite 608
Vienna, VA 22182

Abstract

The FAA will soon begin testing the Wide Area Augmentation System (WAAS). The WAAS will improve the US air traffic control (ATC) system by supplementing the GPS navigation signals with transmissions from a geostationary satellite to make it a more reliable and robust system for civil aviation. In order for the system to satisfy its navigation requirements, the signals transmitted by the WAAS will be synchronized to GPS system time. In addition, they will also be traceable to Coordinated Universal Time (UTC) as determined at the U. S. Naval Observatory. The paper will discuss the architecture of WAAS Network Time and its role and relation to the navigation problem.

Introduction

The purpose of the WAAS is to provide a satellite-based navigation capability for all phases of flight within the National Airspace System (NAS) from en route through precision approach. GPS, alone, does not satisfy all the requirements for civil air navigation. In order to meet these requirements, the Wide Area Augmentation System (WAAS) improves the integrity, accuracy, availability and continuity of GPS service by augmenting GPS transmissions with a GPS-like signal from a geostationary communications satellite (GEO) and by using WAAS Reference Stations (WRS) to continually monitor GPS transmissions.

The data obtained by the WRSs are passed on to a WAAS Master Station (WMS) which prepares WAAS navigation messages which are transmitted to the civilian user by the GEO where they are combined with GPS data in a WAAS user navigation receiver to provide improved GPS positions.

The WAAS navigation messages improve the integrity of GPS by providing timely warnings to the users when the system should not be used for navigation due to the inability of the system to provide adequate service. The WAAS navigation messages also contain information on the current accuracy that the system is providing to the user as well as system availability. Because of the redundancy and robustness of the WAAS, it will be able to provide continuous service throughout flight operations.

The use of a GEO to transmit the WAAS navigation messages is one of the main design features of the WAAS. The transmitted signals can be used in several ways. First, if only three GPS satellites are available, the GEO signals could be combined with the GPS signals to determine a position. The GEO signal could be used in a GPS receiver which utilizes some sort of Receiver Autonomous Integrity Monitoring (RAIM) algorithm to isolate a bad GPS satellite. The differential correction data contained within the WAAS navigation message could be used to improve the position determined by a WAAS receiver which is also looking at GPS satellites effected by Selective Availability (SA). However, it must be cautioned that the WAAS signal, which is often described as a GPS-like signal, is only GPS-like because it is on the same L1 frequency as the GPS Standard Positioning Service. The bit rate is not the same and therefore a special receiver has to be used to decode the WAAS navigation message before it can be combined with other GPS signals.

WAAS Overview

The WAAS will be developed in three phases. However, only the initial phase is presently funded. Also, the performance of the WAAS, as it develops, will determine the scope of the final two stages. The objectives of Phase 1 are to provide a signal-in-space for en route through non-precision approach as the primary means for aircraft navigation and to provide a precision approach capability by the end of 1998. The Phase 1 WAAS will use the International Maritime Satellite (INMARSAT) III Program satellite for its GEO satellite capability.

The initial phase will consist of a Functional Verification System (FVS) whose design will be identical to that of the WAAS, only with fewer subsystems. The FVS will be used for early Development Test and Evaluation (DT&E), refinement of contractor site installation procedures, system-level testing, WAAS operational testing, WAAS upgrade impact analysis, and long-term support for the WAAS. There will be 24 WRSs distributed throughout the United States including Alaska and Hawaii. Each WRSs will contain 3 sets of WAAS Reference Equipment (WRE). Each WRE will contain a cesium beam frequency standard, a 12 channel , dual frequency WAAS-GPS receiver, a wide band GPS receiver and a narrow band GPS receiver which will continually track all GPS and GEOs available to them. The data obtained at these 24 WRSs will be transmitted to 2 WAAS Master Stations where the WAAS navigation message will be formed. The message will contain information on satellite orbits, the current state of the ionosphere, timing information and system health. Initially, the WMS will pass the navigation message to two of the 4 planned Geostationary Uplink Stations (GUS) where it will be uploaded to the 2 Phase 1 geostationary communications satellites which will re-transmit it to the user.

The next 2 phases of WAAS development will add or cut additional WRSs, WMSs, GUSs and GEOs as circumstances dictate. The objectives of Phase 2 are to provide a signal-in-space for precision approach as a primary means and a complete functional operations and maintenance capability by mid 2000, including predictions to support the Notice to Airmen System (NOTAMS).

The objectives of Phase 3 are to meet all of the requirements of the WAAS specification by the end of 2000. A System Design Improvement Program will be sustained during Phases 1 and 2 in order to define and implement Phase 2 and 3 contractual options.

The final contract for Phase 1 was awarded to Hughes on 28 October 1996. Since then, progress has been made in several areas. Major systems reviews have been conducted: Architecture Design Review (ADR) in August; 1966, System Design Review (SDR) in October; 1966, and Software Specification Review (SSR) in December, 1966.

The WAAS Process

Each WRS performs the functions of data collection, reasonability checking, data processing, data recording, and data transferring. Each WRS consists of three Wide-area Reference Station Equipment (WREs) collecting independent sets of data including GPS satellite data, GEO satellite data, and local tropospheric data and transmits the data to each WMS in the System. Independence of data sets is ensured by gathering the observable parameters through independent sets of hardware and is necessary to support the verification function performed by the WMS. The data is collected at a rate consistent with its expected level of variation; e.g., slowly changing weather conditions allow this data to be collected less frequently than data from the GPS satellites. Prior to transmitting data to the WMSs, each WRE verifies the reasonability of its collected data. Failed data is marked as having failed the reasonability test and is forwarded to the WMSs. To ensure the availability of the data at each WMS, each WRS transmits data to each WMS through two independent backbone nodes of the Terrestrial Communications System (TCS).

Each WMS performs the functions of correction processing, satellite orbit determination, integrity determination, verification, validation, and WAAS message generation. Once per second, the WMS collects the data received from all WRSs and processes it to support the functions listed above. This processing is performed on all available WRS data and results in the transmission of a formatted 250-bit WAAS message once per second. These WAAS messages are sent to all GUSs. The timing of WMS processing is scheduled to allow broadcast of the resulting WAAS message from the GEO satellite coincident with the desired GPS 1-second Coarse/Acquisition (C/A) code epoch. The WAAS validates the Signal-in-Space by checking the downlinked messages to ensure that they are identical to those transmitted to the GEO satellites and by comparing navigation position solutions from WAAS/GPS with the surveyed WRS locations. Each WMS includes an Operation and Maintenance (O&M) console from which control

over the WAAS can be exercised via a computer human interface. To avoid conflicts, only one WMS within WAAS can be designated as the controlling WMS (active O&M console) at any one time.

Each Geostationary Communications System (GCS) performs the functions of broadcast and ranging. Each WAAS GCS consists of two Signal Generation Subsystems (SGS), two RF Uplink (RFU) subsystems and one Geostationary Earth Orbiting (GEO) satellite. The combination of a SGS and a RFU is a Ground Uplink Station (GUS). The GCS broadcast function starts with the GCS receiving the 250-bit formatted WAAS messages once each second from each WMS in the System. To improve the availability of WAAS messages, each GUS is connected to two WAAS backbone nodes on the TCS. The GUS selects one WMS as its message source and encodes the received message using a 1/2 rate forward error correcting convolution code. The resulting 500-bit message is modulated on a GPS-type signal and uplinked to the GEO satellite. Each GEO satellite is served by two GUSs: one operating as the primary uplink and the other operating as a hot standby. The two GUSs serving a GEO satellite are operationally independent and located at geographically diverse Ground Earth Stations (GES) separated by a minimum of 300 miles. A GES is a facility consisting of one or more GUSs, and provides shelter, power, and operations and maintenance services for the GUS. The GEO satellite "bent-pipe" transponder shifts the frequency of the signal and broadcasts it to the WAAS users. Transition between primary and backup GUS is initiated, when necessary, to maintain the availability of the WAAS signal-in-space.

The GCS ranging function is accomplished by transmitting a signal to the users in the GPS L1 frequency with the following information: (1) a precisely timed Pseudo-Random Noise (PRN) code which is assigned to each GEO satellite, and (2) a precise ephemeris which is contained in the GEO Satellite Navigation Message sent periodically in the broadcast function. This signal structure is achieved in a similar manner to that of GPS, except that the precise timing of the PRN code is carried out on the ground rather than on the GEO satellites. This function will allow the users to apply the GEO satellites as another GPS satellite, thereby increasing overall system availability.

The Role of Time in WAAS

In order for the WAAS signal to supplement the GPS navigation signals, the WAAS transmissions must be synchronized to them, i.e., GPS Time. This is done by establishing WAAS Network Time (WNT) as the reference time for the WAAS. WNT will be steered to GPS Time. Because the accuracy of synchronization will be at the several nanosecond level, a secondary mission for the WAAS has evolved. This is to distribute precise time in a similar fashion to GPS.

There is a third area where time is of concern. This is in the area of recording the time of observation of all data. In the WAAS, this will be GPS Time. It is important to document this since the time of observation will not be UTC.

WAAS Network Time

One of the functions of the WAAS Master Station is to compute real time orbits for each of the GPS satellites which are in view and observed at the WAAS Reference Stations. In addition to computing the orbits, the WMS also computes the offset of WAAS Network time from GPS Time based on the observations of all GPS satellites in view. The WAAS Master Station will compute WAAS Network Time by an averaging algorithm. Initially, WNT will be based on 24 cesium clocks which are associated with the primary WRE system at each WRS.

The cesium clock at the primary GUS will be slaved to WNT. Once per day, the WMS will issue commands to steer the GUS clock in order to reduce any offset from GPS Time. The GUS clock controls the synchronization of the WAAS navigation message from the GEO. It is estimated that the once per day steering may not be sufficient to keep WNT close to GPS Time. Therefore, the hardware steering will be supplemented by correction messages within the WAAS navigation messages. There will be a slow correction message (Message 9) and a fast correction message (Message 2). The slow correction will be updated once an hour. The fast correction will be issued once per second and will compensate for the effects of Selective Availability.

WAAS Time Distribution System

Time distribution is accomplished as a secondary mission of the WAAS. Time distribution is accomplished by providing users with a time offset between WAAS Network Time (WNT) and Universal Coordinated Time (UTC). This time offset is determined by the United States Naval Observatory (USNO) and passed to the WAAS in order to provide time users with a more accurate source of time referenced to UTC.

The core of the Time Distribution System (TDS) will be a unit which is similar to a GPS Time Transfer Unit. However, the TDS will only look at GEOs which are transmitting the WAAS navigation messages. A TDS will be located at the US Naval Observatory. Input to the TDS will come from the USNO Master Clock. The TDS will then determine the difference between the USNO Master Clock and WNT. These differences will be stored and later processed by standard, operational USNO programs. The results of the calculations will be made available to the WMS which will prepare a WAAS message which will denote the difference between WNT and UTC. This message is designated as Message 12.

Concluding Comments

Besides being a navigation system, the WAAS promises to be an excellent Time Distribution System. It will supplement GPS. It will provide timing signals from a geostationary satellite, fixed in the sky. This will allow the design of higher gain antennas and provide a signal which may not be affected by multi-path reflections. The signal will also not be effected y Selective Availability. This may prove to be one of its most attractive features. For it will allow the more rapid determination of the time offset of a local reference clock from UTC than GPS. With GPS, as affected by Selective Availability, one has to average over many satellites and over long periods of time in order to get reasonably accurate time offsets of a local reference clock from GPS.

Acknowledgments

The authors would like to thank their colleagues at the FAA Satellite Navigation Program Office for their comments, suggestions and help, especially Phil Baker and Brian Mahoney.

TIME and FREQUENCY ASPECTS in EURIDIS

Ph. Gouni *, M. Brunet **, H. Secretan **, P. Uhrich ***

* Thomson-CSF/T4S - 105 Avenue du Général Eisenhower - BP1197-31037 Toulouse - France
phone (33) 5 61 19 79 37 - fax (33) 5 61 19 79 57
** CNES - 18, Avenue Edouard Belin - 31401 Toulouse Cedex 4 - France
phone (33) 5 61 27 31 31 - fax (33) 5 61 28 26 13
*** BNM-LPTF - Observatoire de Paris, 61 Avenue de l'observatoire, 75014 Paris - France
phone (33) 1 40 51 22 16 - fax (33) 1 43 25 55 42

1. ABSTRACT

The current capabilities of GPS and GLONASS constellations, although very adequate for some user communities, present some shortfalls mainly concerning availability or integrity of the positioning service. These performances can be improved by augmenting the existing systems by means of signals broadcast with geostationary satellites.
In Europe, EGNOS is being developed to provide GPS and GLONASS regional augmentation to aviation, maritime and land users. EGNOS is a major element of the European Satellite Navigation Programme which is jointly being implemented by the Commission of the European Union, the European Space Agency and Eurocontrol. EGNOS will disseminate accurate time to all users, synchronized to UTC.
EURIDIS is one constituting element of EGNOS sponsored jointly by CNES-DGAC, with the objective to provide a GPS-like ranging signal through INMARSAT 3 AORE. EURIDIS developed by Thomson-CSF, Sextant Avionics and Syseca will implement the generation and the control of a GNSS signal broadcast by the AORE navigation payload. Benefits of a such system can be considered for the time/frequency community in using this R-GEO signal, mainly for applications requesting a precise time transfer. For this purpose, EURIDIS shall use a synchronization process between its three wide spread stations, with an accuracy of 10 to 20 nanoseconds (1 sigma).
The paper describes the time aspects in EURIDIS (architecture, technical solutions, performances budget) and the benefits of such techniques for the time and frequency community.

2. INTRODUCTION

The purpose of EGNOS (European GEO Navigation Overlay Service) is to implement a range of user services complementing the GPS and GLONASS constellations. Based on the concept of the augmentation of these constellations, it consists in providing geostationary overlays broadcasting GPS-like navigation signals containing integrity and differential correction data on GPS and GLONASS satellites, and data on its own GEO service. The association of GPS, GLONASS and the GEO overlays shall constitute the GNSS1 constellation (Global Navigation Satellite System of the first generation).
The data can be received by any users having a receiver with a GEO channel and located within the GEO Broadcast Area (GBA). They address the needs of different categories of users such as Civil Aviation, maritime applications, land applications.
EGNOS will provide three main functions, namely the GEO Ranging signal (R-GEO), the GNSS Integrity Channel (GIC), and the Wide Area Differential (WAD). This paper gives an overview of their respective goals and implementation. EGNOS will make use of the INMARSAT-III AORE (Atlantic Ocean Region East) and IOR (Indian Ocean Region) navigation payloads which are covering the greatest parts of Europe, Africa and the Atlantic Ocean.
The R-GEO function will be first set up using the INMARSAT-III AORE navigation payload and will be based on the EURIDIS system.
EURIDIS will generate and control a R-GEO signal which conforms the requirements proposed by the RTCA and EUROCAE organisations [1]. It is basically a L1 band GPS signal modulated with C/A codes specifically allocated to the GEO satellites. With reference to the GPS message, the data bandwidth is extended to 250 bit per second, so that the whole GNSS integrity and differential data can be broadcast in the near future.
The EURIDIS architecture is presented and particularly the functions which finely control the time and frequency of the R-GEO signal such as the synchronization with the GPS Time and the relative synchronization of the EURIDIS stations. Uncertainty budgets are developed.
Considering that the EURIDIS signal will not be affected by the Selective Availability (SA) - the intentional noise perturbing the GPS C/A signals for protecting the military applications -, synchronization with UTC (USNO) and continuous time transfer between two sites in common view of the GEO satellite can be envisaged with an accuracy of some tenth of nanoseconds. Finally, examples of users applications with a time oriented GNSS receiver are described and the performances assessed.

3. EGNOS AND EURIDIS PROGRAM OVERVIEW

3.1 Missions and European Programmatic of EGNOS

Since 1993, discussion were held between several European organisations to establish a strategic plan for the development of a regional satellite augmentation of GPS and GLONASS which is the first step of implementing the world-wide GNSS1 (see figure 1). As part of the result, the general outlines of EGNOS were adopted. EGNOS [2] is the regional equivalent to the American Wide Area Augmentation System (WAAS) or the Japanese MTSAT Satellite Augmentation System (MSAS). It is part of the European Satellite Navigation Program (ESNP) initiated by the European Tripartite Group consisting of the Commission of the European Union (CEU), EUROCONTROL and the European Space Agency (ESA).

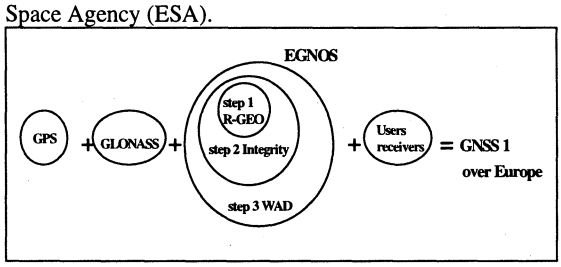


figure 1: the GNSS1 concept

The missions of EGNOS are to offer on a wide zone a navigation service to three main communities of users:

- the Civil Aviation, where EGNOS will be used from en-route navigation of the aircraft till precision approach landing phase (CAT1). The requirements of this application are the more demanding at least in term of coverage, availability and integrity [3], and nearly set up the general design of the system,
- the maritime applications for sea coastal and harbour navigation where need for differential corrections are expressed,
- the land mobile applications where various systems also require the use of differential corrections.

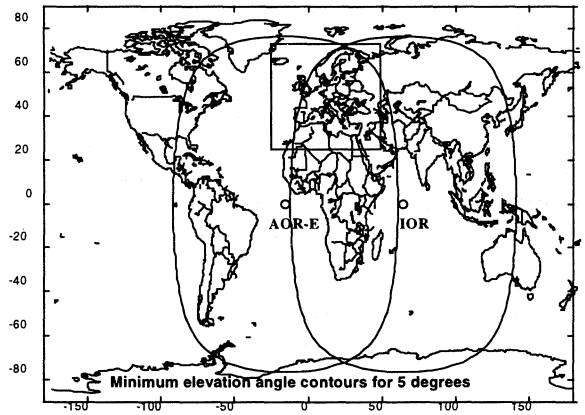


figure 2: AORE and IOR GBAs

For the EGNOS AOC (Advance Operation Capability), above users' zones are covered by the grouping of both INMARSAT III AORE and IOR payloads (see figure 2).

The EGNOS AOC will be developed in three incremental steps.

- Step 1, called ranging or R-GEO, will provide GPS-like signals through the geostationary satellites, that will augment the number of pseudo range measurements available for the GNSS users. It is planned to be operational in fall 1998.
- Step 2, called Integrity, will use the data channel of the Ranging signal to broadcast integrity information on the GPS and GLONASS satellites.
- Step 3, called Wide Area Differential, will provide the users with correction data on GPS pseudo-ranges and ionosphere delays to improve their positioning accuracy.

The industrial team performing these tasks is led by Thomson-CSF (Fr) with the participation of European companies from Spain, UK, Germany, Switzerland and Norway.

3.2 Components of EGNOS or EURIDIS Architectures

The EGNOS architecture is illustrated in figure 3. It consists of the ground segment, the INMARSAT navigation payloads space segment and the user segment (the users' receivers). The basic components of the EURIDIS ground segment are the same but in limited number. The ground segment is composed of :

- the Ranging and Integrity Monitoring Stations (RIMS), forming a wide area tracking network collecting the data from the GNSS1 constellation,
- the Master Control Centres (MCC) receiving the data from the RIMS in order to produce the high level information to be included in the navigation message to be broadcast (e.g GEO ephemeris, slow or fast corrections on the GPS/GLONASS pseudoranges, time parameters, ionospheric correction model...),
- the Navigation Land Earth Stations (NLES) which are the GEO access stations in charge of generating and controlling the GPS-like signals by maintaining a close loop with the GEO. The signal is synchronized in real time at the output of the GEO L1 band antenna with the GPS Time or the EGNOS time,
- and the Wide Area Network for the transport of data between the different sites.

3.3 Time Transfer Applications

Beside the EGNOS services already described above, applications can also be considered with some benefits by the time/frequency community, providing that time oriented GPS/GEO receivers are used. The broadcast of the signal over a wide area gives the capability to synchronize the local user time to the GPS Time in making use of the GEO ephemeris contained in the message, and to relatively synchronize two distant sites which are within the same GBA and even within the association of the two GBAs. In addition, the difference between the UTC USNO and the GPS Times is included in the GEO message, allowing the user to also connect its local time to the UTC USNO time.

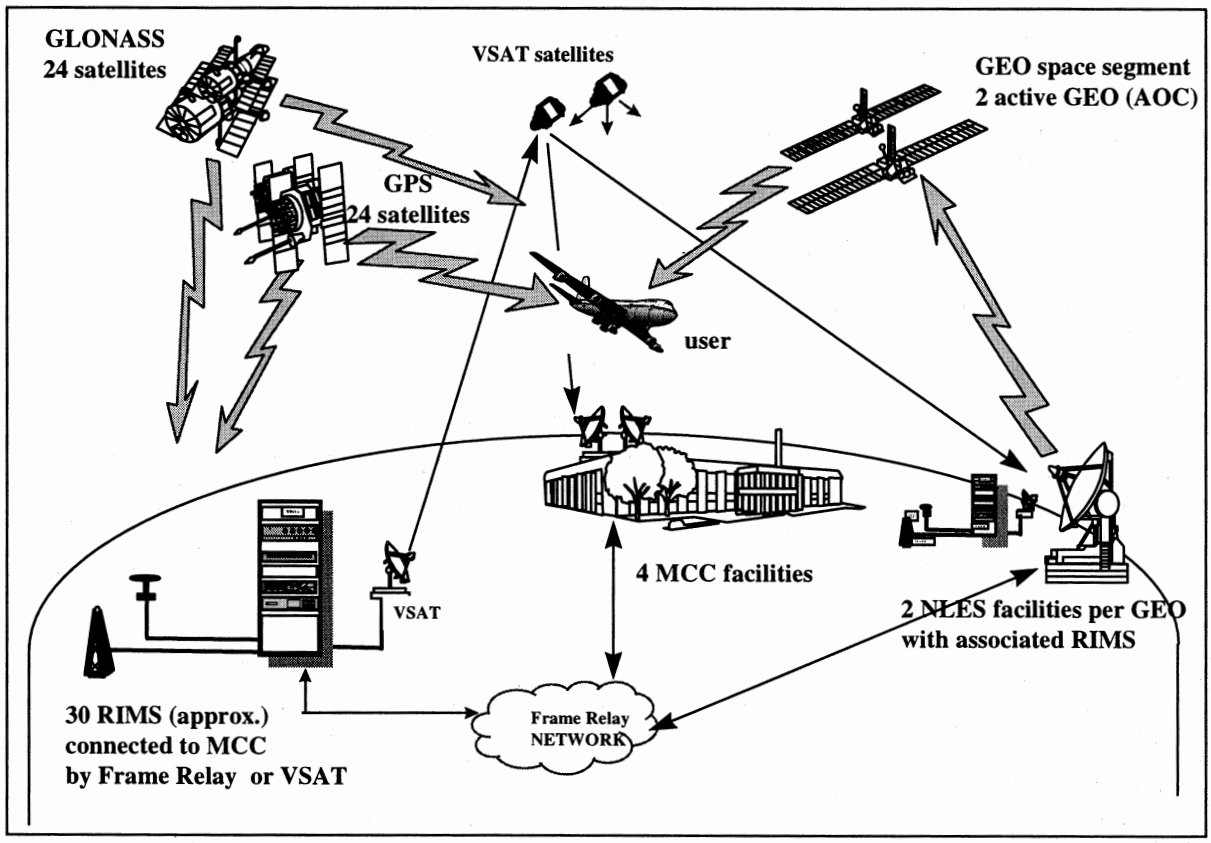


figure 3: components of EGNOS / EURIDIS

4. EURIDIS AND THE R-GEO

The development of the R-GEO EURIDIS has been initiated in the frame of EGNOS by the French space agency (CNES) and by the French Civil Aviation (DGAC) [4]. The project has been running since July 95, and EURIDIS will constitute an In Kind Delivery to the step 1 (R-GEO) of EGNOS. One of the objectives is to operate the AOR-E payload which has been delivered in orbit by INMARSAT in November 96.

The figure 4 shows the implementation of EURIDIS. It is based on three RIMS stations implementing a wide triangular observation basis for easing the trajectography determination of the GEO satellite, on an MCC in Madrid (Sp), and on the use of the existing INMARSAT Aussaguel NLES station (Fr) operated by France Telecom.

5. FREQUENCY/TIME ASPECTS IN EURIDIS

5.1 Synchronization of the R-GEO Signal with the GPS Time

The ability of a user to fix an accurate position with at least four GPS signals is based on the fact that all GPS signals are transmitted in synchronization at the same epoch time (the GPS TOW), or, if not exactly, that data are given in the GPS messages to resynchronize the transmit time of the signals.

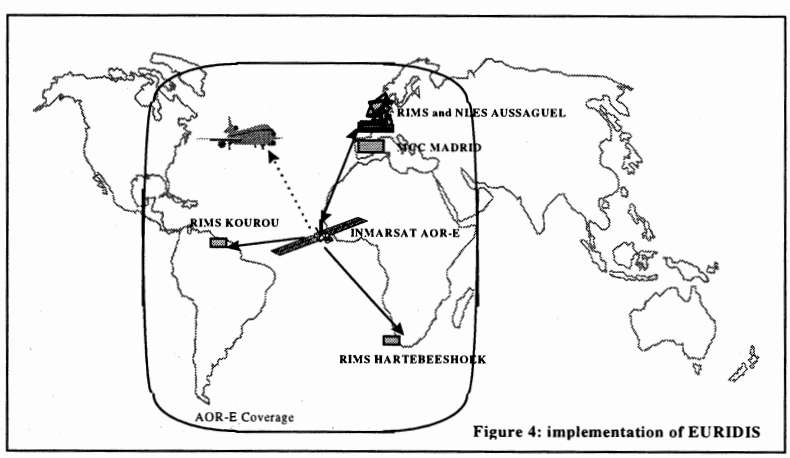


Figure 4: implementation of EURIDIS

As a part of the constellation, EURIDIS shall thus be synchronized with the GPS epoch time, the requirement being that the EURIDIS transmit epoch time at the output of the L1 band antenna of the GEO satellite shall be known by the user with a 28 ns (1σ) accuracy with reference to the GPS epoch time.
EURIDIS performs this synchronization process in actually separating it in two functions:

- (a) achieve the restitution of the GPS Time in the ground segment,
- (b) set up the GEO Time with respect to this GPS restored time.

5.1.1 (a) GPS Time Restitution

Unfortunately, the knowledge of the GPS Time is disturbed by the Selective Availability. Experience [5,6,7] has shown that, with the current SA characteristics, the GPS Time can be restored with an accuracy better than 20 ns (1σ) providing that a sufficient tracking time over a sufficient number of GPS satellites is ensured with sufficient instantaneous accurate measurement.
This function is implemented in EURIDIS by using an all-in-view 15 channels GPS/GEO receiver in a precise known location at the Aussaguel RIMS. This receiver is made by Sextant Avionics (Fr); it has been adapted to receive a caesium frequency reference representing the EURIDIS Central Time (TGNSS) and to deliver each second the pseudoranges of the GPS in view with reference to this Caesium calibration signal (see figure 5). After 2 hours data collection, the MCC computes the difference between this observed "composite" GPS Time (TGPSO) and TGNSS, in using the GPS pseudorange residuals with appropriate corrections (antenna/cables calibrated delays, ionopherics and troposheric corrections, Sagnac effect...). Actually, the MCC gives a linear model of the drift between TGPSO and TGNSS.

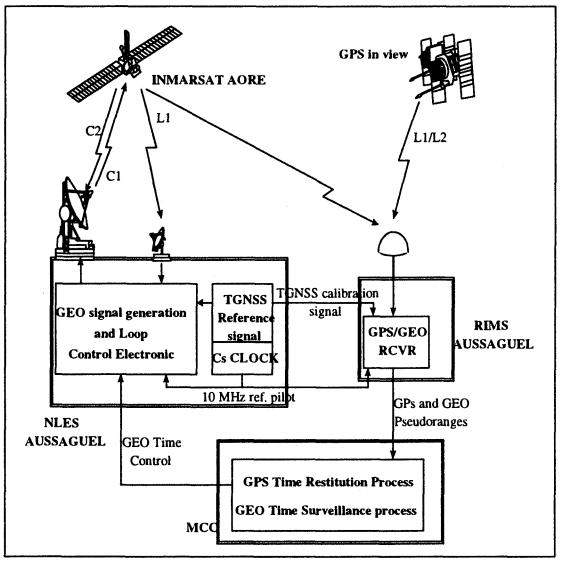


Figure 5: Synchronization process of the on-board GEO Time with the GPS Time

5.1.2 (b) GEO On-Board Time Transfer

The objective is then to transfer the time to the GEO. The requirement is that the additional error due to the transfer shall be less than 20 ns (1σ). Two strategies are possible:

- let the GEO Time be enslaved (equal to) TGNSS and communicate the drift TGPSO/TGNSS in the data of the GEO message,
- or apply this drift to the on board GEO Time and control it in real time so that the GEO Time be the closest possible to TGPSO, in using the close loop of the NLES. In these conditions, no clock correction is needed at user level.

For various reasons, the second solution was chosen for EURIDIS. The MCC sends the drift model to the NLES during the 2 following hours of the model computation, and the GEO Time is monitored by the Aussaguel RIMS, using the pseudorange residuals on the GEO satellite. An additional smoothing process ensures the continuity of the GEO Time transfer. The main sources of error is the calibration errors of the delays in the payload and in the transmit/receive chains at the NLES station.

5.1.3 GEO/GPS Time Synchronization Budget

The total GEO/GPS Time uncertainty budget is presented in table 1. It takes into account all the effects along the sequence described above.

Table 1: Uncertainty budget of the synchronization of the GEO Time versus the GPS Time in EURIDIS.

BUDGET COMPONENTS		uncertainty (1σ)
GPS Time restitution part	SA dither residual (1)	15
	GPS position error (2)	9
	Tropospheric delay error	1
	Ionospheric delay error (3)	3
	antenna coordinates error	< 1
	Receiver noise (2h filtered)	< 1
	GPS multipath (filtered) (4)	1
	Earth station calibration (antenna, cables, LNA...)	2
GEO on-board transfer part	GEO close loop error (5)	5
	prediction error (MCC model)	1
	NLES transmit/receive delays errors (6)	8.1
	payload delay error (6)	5.2
TOTAL RSS		**21 ns (1σ)**

notes:
(1) take as an assumption the SA dither described in ref[1], noise filtered on 2 hours,
(2) no SA ephemeris taken into account,
(3) L1/L2 dual frequency Ashtech receiver also used,
(4) the Sextant receiver has a GPS multipath rejection system,
(5) include electronic noise, ionospheric error and loop lag effect; weighted by the round trip with the satellite,
(6) weighted by the round trip with the satellite.

5.2 Relative Synchronization of the EURIDIS RIMS

The accuracy of the orbitography ordinary necessary to control a GEO satellite is about several 100 m and is not

sufficient to cover a GPS type navigation application. For an accurate positioning or GPS Time synchronization, the GEO radial position shall be known with an accuracy better than 8 m (1σ). This position shall then be delivered in the GEO message under an ephemeris form. As a consequence, the EURIDIS ground segment shall also finely restore the GEO position.

The method consist in doing a "one way ranging" on the three RIMS with the collected GEO pseudoranges. But it is also necessary to synchronize the RIMS within a 10 ns (1σ) accuracy. This synchronization process is called the relative synchronization of the ground segment.

To perform this function, the chosen solution is to maintain the time stability in each RIMS on the medium term (drift < 5 ns over 12 hours), using an ultra precise HP caesium clock, and to measure that drift with respect to TGNSS, the Aussaguel RIMS time, by the GPS common view technique. For system initialisation reasons, the common view with its own GEO payload is not desirable. As a consequence, the relative synchronization is only based on the GPS constellation, even if one drawback is the unknown but limited effect of the SA ephemeris over the common view technique applied to overseas stations [5].

The MCC computes each 12 hours a linear drift model between the RIMS clocks. The raw pseudoranges are collected over 48 heures gliding panels, corrected (same corrections than in the GPS restitution time process) and translated in pseudorange residuals using the known position of the stations and the GPS broadcast ephemeris. The uncertainty budget described in table 2 gives an estimate of 8 ns (1σ).

table 2: Uncertainty budget of the relative synchronization between two RIMS of EURIDIS.

BUDGET COMPONENTS (note 1)	**uncertainty (1σ)**
GPS position error (2)	4.3
tropospheric error	1.4
ionospheric error (3)	4.3
antennas coordinates error	1.4
receivers noise (48h filtered)	1.4
delay error in the receiver	2.8
Earth station calibration (4)	2.8
TOTAL RSS	**7.7 ns (1σ)**

notes:

(1) all components are multiplied by $\sqrt{2}$ to account for both sites to be synchronized,
(2) no SA ephemeris, filtered over 48 h,
(3) dual frequency L1/L2 corrections,
(4) account for antenna, LNA and cables calibration errors.

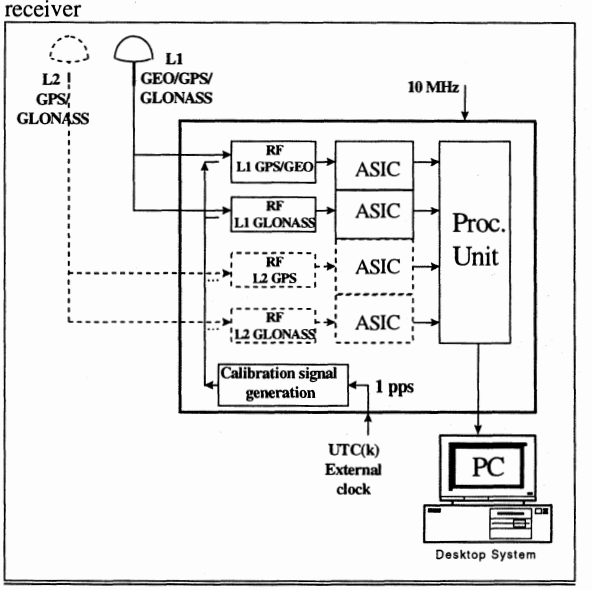
6. USE OF THE R-GEO SIGNAL FOR TIME METROLOGY PURPOSES

For receiving the R-GEO signal, the Time and Frequency community will have to use a time oriented GNSS receiver. The access to GPS Time or UTC (USNO) will be made possible in real time. The well known common-View technique will be applied to the R-GEO signal for continuous time transfer purposes.

6.1 Time Oriented GNSS receiver

The architecture of a future time oriented GNSS receiver is given in figure.6 The signals from GPS, GLONASS and the GEO can be received separately or simultaneously, and processed by multichannel ASICs, which are already currently available. Moreover, a specific GPS-like signal synchronized to the local 1 pps of the external clock UTC(k) of the earth station k, called "pps-C/A-L1", will be sent to the input of the receiver through what is called a "calibration generator". That calibration signal will allow to measure the external 1 pps reference signal towards the time base inside the receiver. At the output of the receiver, one will get UTC(k) - GPS Time via the GEO or via any GPS SV, and UTC(k) - GLONASS Time via any GLONASS SV. The total delay of the calibration generator and of the receiver will also be measured through each channel of the receiver. The process of UTC(k) - GPS Time is performed in the associated PC, with the appropriate correction and filtering software.

Figure 6 Architecture of a time oriented GNSS receiver



note: L2 optional reception if ionospheric corrections are needed (dotted lines)

6.2 Provisional Accuracy of GPS Time Restitution and of Common-View Time Transfer when using the R-GEO Signal

6.2.1 GPS Time Restitution via the GEO

The table 3 gives the provisional uncertainty budget of the GPS Time restitution when using the R-GEO signal and a time oriented GNSS receiver. It characterises the mean evaluation of the time difference UTC(k) - GPS Time over sessions of duration 15 minutes, with a basic sampling of one second. With the data included in the navigation message of any GPS satellite, that measurement gives access to the time difference UTC(k) - UTC(USNO) every 15 minutes.

Table 3: Provisional uncertainty budget for the mean evaluation over 15 min of the GPS Time restitution when using the R-GEO signal, in ns (1 σ).

BUDGET COMPONENTS	**uncertainty (1σ)**	note
GEO ephemeris	25	(1)
GEO Time	28	(1)
Ionosphere (Bent)	15	(2)
Troposphere	1	
Antenna coordinates	< 1	(3)
Antenna and receiver delay	2	(4)
Multipaths	< 1	(5)
Receiver noise	< 1	(4)
Link to the local clock 1 pps	1	
TOTAL RSS	**40 ns (1σ)**	

Notes :

(1) from the specifications of the system design,

(2) that value could be as low as 3 ns with on site measurements on the two Inmarsat carriers L1 and C2. Another option is to use all the GNSS observable to map the ionosphere in the direction of the Inmarsat satellite. In the near future (EGNOS level 3), a ionospheric error model will be broadcast in the GEO message,

(3) the coordinates of the antenna have to be determined in a global geodetic reference frame as ITRF or WGS,

(4) expected or anticipated values (15min filtered noise),

(5) choke-ring antenna.

6.2.2 Common-View Time Transfer via the GEO

The table 4 gives a provisional uncertainty budget of Common-View Time Transfer between two clocks, both connected to a time oriented GNSS receiver, when using the R-GEO signal. The main interest is that the R-GEO signal is continuously available over the area covered by the Inmarsat satellite, with no selective availability or anti-spoofing on the signal. Again, the sessions duration should be of the order of 15 min, or 13/16 min by analogy with the international schedule of GPS Common-Views. It could be foreseen that this would achieve a major change in the Time Transfer in the future, even if one of the drawbacks lies in the potential bias created by the stationary position of the satellite (e.g. multipath).

Table 4: Provisional uncertainty budget of Common-View Time transfer when using the R-GEO signal in ns (1σ)

BUDGET COMPONENTS (note 1)	**uncertainty (1σ)**	note
GEO ephemeris	15	typical (2)
Ionosphere (Bent)	21	(3)
Troposphere	1,4	
Antenna coordinates	< 1	(4)
Antenna and receivers delays	2,8	(5)
Multipaths	< 1	(6)
Receiver noise	< 1	(5)
Link to the local clock 1 pps	1	
TOTAL RSS	**26 ns (1σ)**	

Notes :

(1) all components multiplied by $\sqrt{2}$ to account for both receive chains,

(2) depends on the geometry between the stations and the satellite,

(3) that value could be as low as 3 ns with on site measurements on the two Inmarsat carriers L1 and C2. Another option is to use all the GNSS observable to map the ionosphere in the direction of the Inmarsat satellite. In the near future (EGNOS level 3), a ionospheric error model will be broadcast in the GEO message,

(4) the coordinates of the antenna have to be determined in a global geodetic reference frame as ITRF or WGS,

(5) expected or anticipated values (15min filtered noise),

(6) choke-ring antennas.

7. CONCLUSION

This paper has presented the plans for EGNOS development, and within EGNOS, the characteristics of EURIDIS under a Time/Frequency point of view. EURIDIS will be technically qualified in the second quarter of 98 and it is foreseen that EGNOS level 1 (AOR-E+IOR) be implemented at the end of 98. Due to its stationary position, the R-GEO service should bring a noticeable contribution to the time frequency applications.

8. REFERENCES

[1] MOPS for GPS/WAAS Airborne Equipment, RTCA DO-229, annex A and B, january 16, 96.

[2] EGNOS, the European Regional Augmentation to GPS and GLONASS, S.Loddo, D.Flament, J.Benedicto, P.Michel, ION GPS 96.

[3] Description and Assessment of Integrity Performances for LADGPS and WADGPS Architectures based on Augmented GPS Constellations, P.Gouni, N.Marchal, J.L.Jonquières, T.Carlier, DSNS 94.

[4] Design and operation of the EURIDIS Ranging System, P.Gouni, D.Flament, A.Job, DSNS 96.

[5] Aspects temps du complément européen au Global Positioning System (CE-GPS): résultats expérimentaux des phases continentales et transatlantiques. B.Juompan, P.Uhrich, M.Brunet, Bulletin du BNM n°101, juillet 95.

[6] Time aspects of the European Complement to GPS: summary of the experimental results, P.Uhrich, B.Juompan, R.Tourde, M.Brunet, J.F.Dutrey, EFTF 95.

[7] European Complement to GPS: Main Experimental Results, J.Barbier, M.Brunet and all, ION GPS 94.

THE GPS CARRIER AS A PRACTICAL FREQUENCY REFERENCE

G. Kramer° and W. Klische*

°Physikalisch-Technische Bundesanstalt
D-38023 Braunschweig, Germany
*K+K GPS-Entwicklungen, D-38159 Vechelde

1 ABSTRACT

The paper describes a fully automatic GPS controlled standard frequency and time pulse generator, which uses high resolution phase measurement of the 1.6 GHz GPS carrier as well as the C/A code correlation to determine and tightly „discipline" the frequency of a crystal oscillator. The influence of the so-called „Selective Availability" (SA) mechanism on the frequency precision better than 10^{-11} for all averaging times even without the use of DGPS methods.

Keywords: GPS, frequency standard, carrier phase

2 INTRODUCTION

Like the signals of other navigation systems (e.g. LORAN, OMEGA), the transmissions of the GPS satellites are particularly well suited for time & frequency applications. GPS timing receivers have long been developed and their application is now well established for the comparison of atomic clocks and the synchronisation of national time scales. Common view schedules and centralised data collection and evaluation at the BIPM have made possible the utmost precision.

On the other hand there exists a wide demand for frequency calibrations with only high to medium, but still reliable and traceable, precision. A set of atomic clocks continually monitored and compared to national or international standards would constitute the best possible reference for that. Because of the high cost of caesium atomic clocks this cannot be the solution, and in most cases high stability crystal oscillators monitored by means of such LF transmissions as DCF77 or equivalent standard frequency and time stations have to suffice .

Use of satellite signals for the purpose is made difficult by the fact that the signal sources are in motion with respect to the receiver, giving rise to a Doppler shift (up to $\pm 2.7 \cdot 10^{-6}$ for the GPS satellites), which must be accounted for with high precision for any serious application. The navigation messages, broadcast by each GPS satellite, provide the necessary data, but it still means that a GPS standard frequency receiver must be a full navigation or surveying receiver with special features added for extracting the time and frequency information from the signals.

There now exists a fully automatic timing receiver/standard frequency source, which is based on the hard- and software originally developed for a sail boat navigator. Its main feature is the high resolution carrier phase measurement, which is now being used for frequency measurement instead of Doppler boat speed determination.

The channel characteristics of the GPS carrier are excellent. With coherent detection of the carrier frequency at 1575.42 MHz, receiver noise contributes a mere picosecond in an averaging time of one second (as compared to 10 ns for the C/A code). This means that frequency determination at a precision level of $1 \cdot 10^{-10}$ can be carried out in about 100 ms.

These frequency determinations are then combined in an optimised way with the timing information derived from C/A code **correlation** and message decoding.

3 CARRIER PHASE MEASUREMENT

Because of its spread spectrum characteristics, the evaluation of the whole GPS signal is best described in terms of a cross-correlation process. The complete format of the GPS transmissions is specified in public documents (Ref.3). If a person knows his position and the time, he can calculate, what signal to expect from the satellites. Determining the deviations from these estimates may be called a measurement, where the correlation delay is the interesting quantity.

Generally speaking, the expected antenna signal is somehow formed (synthesised) in the receiver, the incoming signal is multiplied with this expected signal, the time average of the product is the correlation amplitude. The whole correlation process, - of synthesis, multiplication, averaging, and evaluation - is implemented in the receiver by

employing a combination of analogue, digital, and micro computer functions, which I am not going to line out now. If all functions are designed correctly, the performance is not limited by the hardware, but by the properties of the GPS signal itself.

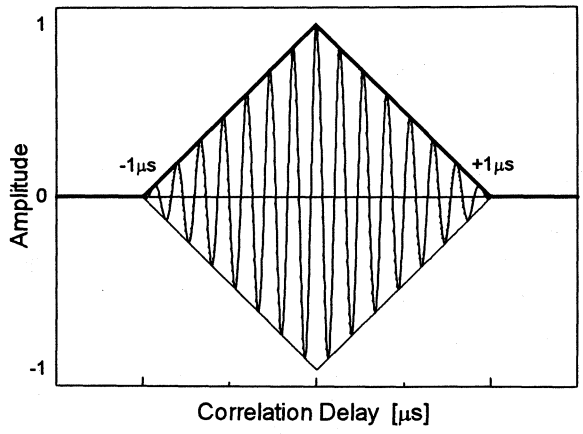


Fig.1 Correlation pulse with carrier

The first figure shows the general shape of the central peak of the cross correlation function. The triangular envelope represents the correlation peak associated with the decoding of the C/A code pseudo random sequence, whose so-called chip length is ≈1 μs, whereas the fast oscillation represents the carrier, of which a number of 1540 periods would fit in one 'chip' length. In order not to completely blot the picture only a fraction of this number was actually drawn. It can be estimated from the numbers though that the sensitivity to delay changes is 10000 (1540·2π) times higher for the carrier decoding than for the C/A code.

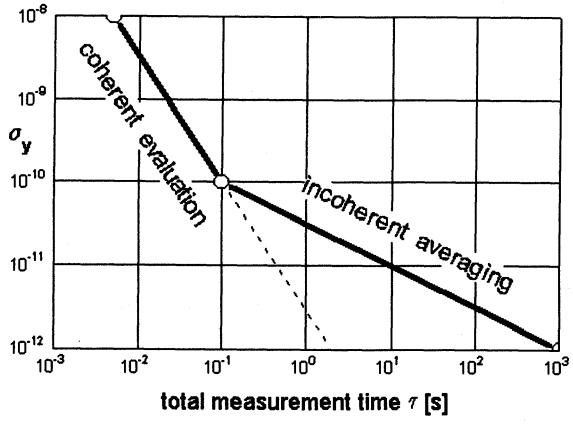


Fig.2 Carrier signal frequency stability

This factor is so immense - it corresponds to a factor of 10^8 in observation time for the same timing noise - that the differences between a multiplexed channel and multi channel system become negligible in this context.

The typical S/N ratio of a GPS receiver with a non directional antenna is 100 to 200 in one second averaging time. Considering the signal slope with respect to the delay time, this means that the corresponding timing noise $\sigma_x(\tau=1s)$ is 10 ns for the code evaluation and 1 ps for the carrier phase. In our receiver the carrier phase measurement is used for frequency determination only. Fig.2 shows the standard deviation of a rate determination $\sigma_y(\tau)$ as a function of the measurement time. For a coherent measurement it varies as $1/\tau^{1.5}$. If we limit the dwell time for each satellite track to 100 ms, we lose the coherency for times longer than 100 ms and gain only with the square root of time by incoherent averaging.

4 SMOOTHING

As no attempt is made to resolve the carrier cycle ambiguity (as opposed to what is done in P-code geodetic receivers), frequency and time information are combined in a process that we call smoothing. Integration of the quasi noise free rate gives the time with an offset, which can then be determined from the C/A code correlation. In the navigator, where a 4-dimensional navigation solution is required, this corresponds to a dead-reckoning integration of a 3-dimensional velocity as well. The high definition 4-dimensional (position + clock) track can then be slowly corrected for an initial offset and possible small systematic offsets which might develop in the course of the integration process. The system works like a mathematical filter(diplexer), which reduces the noise fluctuations like a low pass filter and still preserves the quasi immediate response.

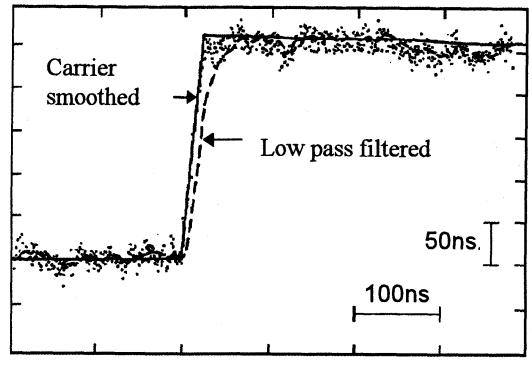


Fig.3 Time response to a clock rate excursion

This property is demonstrated in fig. 3, where we have used the output of a synthesiser as the reference clock of the receiver. The receiver was programmed so as to read out (and output as a data stream) the clock rate y as well as the clock error x as determined from the satellite signal. The

figure shows the response that occurred, when the synthesiser was manually switched to a fractional frequency offset of $y = 1.2 \cdot 10^{-8}$ for a duration of 20 s. The dots represent a pure C/A code evaluation with the scatter associated with the typical signal noise in a 1 s averaging time. The dashed line shows the output of a 10 s recursive low low-pass filter, which gives a delayed but still noisy response, whereas the solid line, the result of the carrier smoothing, looks almost like a synthetic, theoretical curve. The time constant for the smoothing filter is normally set to a few hundred seconds, reducing the contribution of the noise to well below a nanosecond rms.

5 SA DEGRADATION

As we have seen, carrier phase smoothing gives us the possibility of measuring the satellite signals very precisely and almost without noise. Unfortunately, the precision of these signals is deliberately degraded for use by the general public by what is called the **S**elective **A**vailability mechanism. This mechanism mainly shifts the signal clock in a quasi random manner. The SA shifts affect the carrier and the code in the same way. (The phase relation between carrier and C/A code remains stable.) So, if we solve for the receiver clock, this is what we get for a stable reference clock:

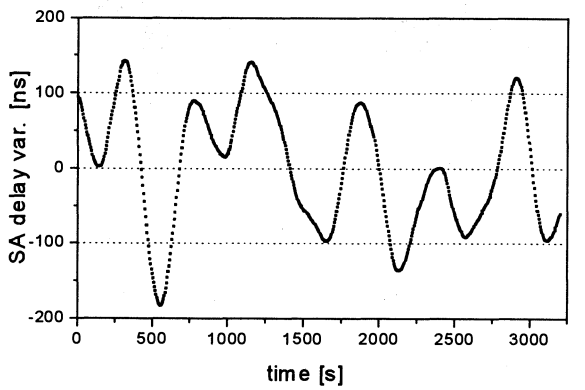


Fig.5

Time variations of 100 ns rms are the normal. The time error is well specified statistically and limits are guaranteed for the system. The so-called URE, the User Range Error is contained in the ephemeris data message, the clock and orbit representation of each satellite. (The value is 32 m, when the SA is switched on).

The spectrum of the SA perturbations, on the other hand, is not specified. So a maximum rate error cannot be guaranteed; we observe peak variations of the clock rate of $2 \cdot 10^{-9}$.

5.1 "SatMix" averaging

For a stand-alone receiver the only way of improving the frequency precision is by averaging. Quite useful precision is already achieved, when receiving and evaluating the signals of all satellites in view. With a number of eight satellites normally in view, this reduces the fluctuations already by a factor of three. Fig.6 gives the measured short term frequency $\sigma_y(\tau)$ and the Allan standard deviation $\sigma_y(2,\tau)$ of the 1 pps output of the receiver, whose (very simple) crystal oscillator is closely locked to what we call the SatMix. 'Closely' refers to a PID servo of less than 10 s time constant, which forces the receiver clock to follow the synthetic, calculated SatMix time scale to within one nanosecond.

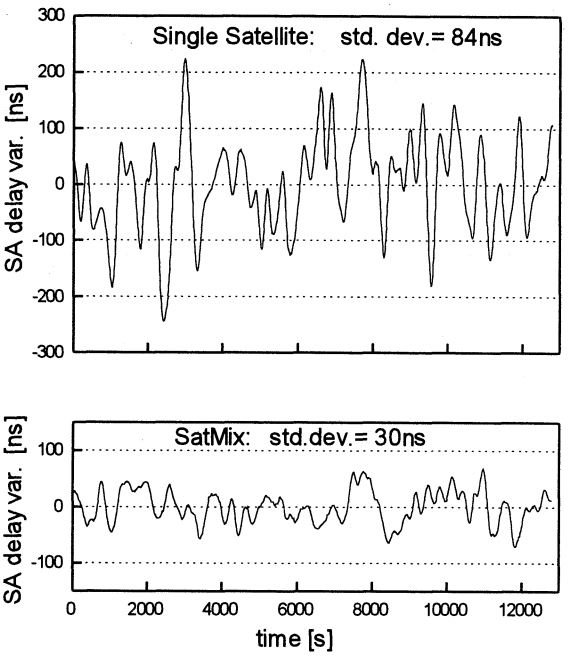


Fig.6 SatMix "All in view"

The SatMix time scale is the best estimate of the GPS (or UTC) time scale, using the carrier phase smoothing and a combination of all satellite signals available. The smoothing process ensures that the SatMix time scale does not jump, that no step occurs, when a satellite drops out of view or a new satellite comes into play (each one with its particular SA timing error of that time). A small step occurs in the rate of the Mix only.

Ideally the SatMix should be the weighted average of all observations available. Setting the weight factors is not trivial, straightforward task, though, but a matter of design goal (taste) and can always be a compromise only. In the past we have used the inverse of the URE (not squared !) as the weight. The main effect now being the greater

emphasis it gives to observations from satellites, on which the SA is temporarily switched off. At present this is one Space Vehicle (SV15), visible for about 5 hours a day. So we think a system should not rely on these observations and should rather consider the fact that the influence of a number of systematic errors (e.g. position error, multipath reception) is minimised or at least kept stable, if observations from all directions are contributing equally.

5.2 Time Averaging

For averaging times up to a few thousand seconds, the intrinsic frequency stability of a good OCXO is better than that of our SatMix and can therefore be used to time average over the SA fluctuations. This is being done by using the OCXO as the master oscillator of the GPS receiver and slowly (but continuously) phase-locking it to the SatMix by means of a D/A converter and a computer servo.

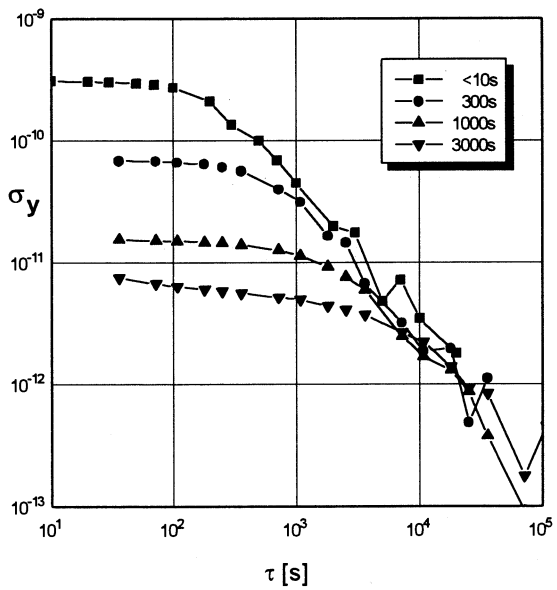


Fig.6 Frequency stability of disciplined Xtal oscillator

The computer servo employs two integrators to bring the average time error to zero even in the presence of a drift of the intrinsic frequency of the oscillator. Its time constants can be easily adapted to the needs.

Fig.6 shows the resulting frequency stability for various settings of the time constant. In a stable environment, 3000 s and more give the best result with this OCXO, which has only simple temperature stabilisation. When the receiver was set up in my office, the variations of ambient temperature were the main perturbations. Particularly the rather sharp temperature rise in the morning was too abrupt for the long time constant. This sharp rise only occurs, when the economy/ecology program of our heating system brings the temperature back into comfort range after having lowered it for the night.

In the figure we show $\sigma(\tau)$, the standard deviation of a frequency measurement of duration τ. As the mean frequency error is zero, this standard deviation is also the rms error to be expected of such a measurement. (The Allan variance would be misleading, suggesting that short measurements were better than long ones.) Another problem with measuring and specifying $\sigma(\tau)$ exist for the long averaging times. (It would be the same for the Allan variance.) Here the fluctuations have the character of 'white phase noise', and taking only two phase readings, at the beginning and at the end of an interval, does not give the best estimate of the frequency, as was mentioned already when talking about the carrier phase evaluation. Again, this is not the place or time for going into the details about statistics and estimators.

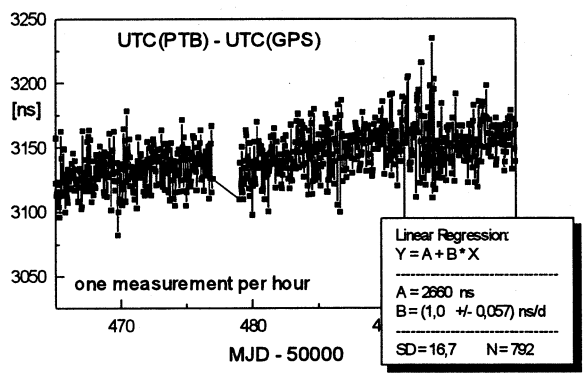


Fig.7 Long term stability (arbitrary time offset)

5.3 Comparing atomic clocks

In the preceding paragraphs I have emphasised the usefulness of the GPS receiver for disciplining a Quartz or Rubidium controlled oscillator, which can then be used as a frequency reference for quickly calibrating the time base of a counter (frequency meter) for example.

If on the other hand you have a stable atomic clock and want to verify its correct rate and time, it is sufficient to continuously record the time difference of the 1 pps outputs of the receiver and the atomic clock. The most simple oscillator in the receiver will do, because the averaging can be carried out in the data. A good OCXO, servo controlled with a long time constant, reduces the

data rate, but is not necessary for the precision of the comparison.

The figure (fig.7) shows data taken and recorded by Dr. Bauch in the PTB time laboratory. It represents the time difference between UTC(PTB) and the SatMix GPS time as received by the new automatic receiver/ generator. With the time constant set to 1000 s most of the time, one measurement per hour is slightly undersampling and results in a standard deviation of these time measurements of 15 ns. The one day averages, which should be particularly stable systematically, will then have a noise contribution of only 3 nanoseconds.

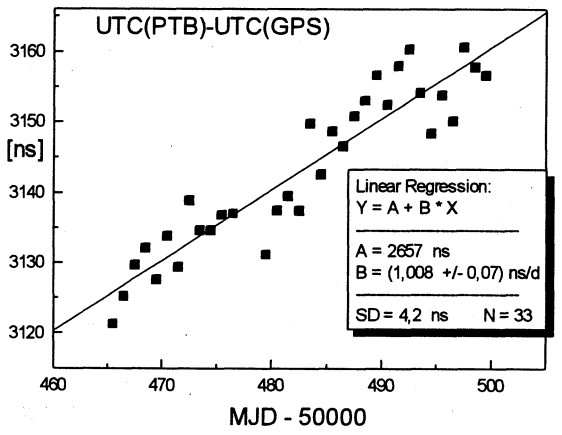


Fig 8 Daily SatMix averages (arbitrary offset)

Fig. 8 shows these daily values recorded for the last month. Only during a week or two has the receiver been really undisturbed, and a linear regression fit to those data confirms the expected small standard deviation of $\leq$ 3 ns for the GPS daily averages. Considering the last month, we find a rate of

UTC(PTB) - UTC(GPS) = (1.01 $\pm$ 0.07) ns/d

6 CONCLUSION

We have presented the principal features of a fully automatic GPS time receiver and some measured performance data. We see three fields of application:

In its simplest form the receiver provides full time information with a precision better that 50 ns.

For quick calibrations on a precision level of better than 10^{-11} (of e.g. the time base of a frequency measuring instrument), the 'disciplined' OCXO can serve as a stable reference with traceable accuracy .

The rates of atomic clocks can be compared at least to a level of $1 \cdot 10^{-15}$ by long term averaging the fluctuations introduced by the SA mechanism. The final limitations of the SatMix as a universal common reference will need further evaluation. Clear advantages are its general availability and simple (for the user) application.

6.1 References

There exists a great amount of literature about GPS subjects. Not all of it is of equal importance. The two most elucidative articles are found in a special issue of NAVIGATION VOL25 No.2 (Summer 1978)

Ref.1 J. J. Spilker: "Signal Structure and Performance Characteristics" p.121

Ref.2 A. J . Van Dierendonck et al. "The GPS Navigation Message", p.145

Ref.3 The official specification document is the STANAG 4292, MAS, NATO unclassified

FREQUENCY COMPARISON USING GPS CARRIER PHASE: SOME EXPERIMENTAL RESULTS

Gérard Petit

Bureau International des Poids et Mesures
Pavillon de Breteuil, 92312 Sèvres CEDEX, France

ABSTRACT

When carried out with C/A code measurements, the capability of GPS for frequency comparison is limited to the level of a few parts in 10^{14} for an integration duration of one day. This performance does not allow to compare, in a sufficently short interval, recently-developed frequency standards for which the stated accuracy is less than 1×10^{-14}. In contrast, carrier phase measurements from multi-channel geodetic GPS receivers have great potential for this purpose because their measurement noise is lower than that of the code measurements by two or three orders of magnitude. Phase and code measurements from TTR-4P receivers in time laboratories and from a geodetic GPS data base maintained by the International GPS Service for Geodynamics (IGS) have been analysed to compare the frequencies of H-masers and other clocks. Preliminary results indicate that, while a fractional frequency stability of several parts in 10^{15} is achieved for averaging durations of several hours, the uncertainty in the frequency comparison is probably dominated by environmental variations in the hardware delays.

Keywords: Frequency comparison, GPS, Carrier phase

1. INTRODUCTION

Quality in time metrology depends on the performance of atomic clocks and on the means used for time and frequency comparison. For the last five years, wide-spread use of the GPS common-view technique has led to major improvements in precision and accuracy, making it possible to compare the best standards in the world at their full level of performance: for integration times of 1 to 3 days, the frequency differences between atomic clocks are consistently measured to about one part in 10^{14}. Recent developments in atomic frequency standards suggest, however, that this performance may no longer be sufficient. The new caesium fountain FO1, built at the BNM-LPTF, Paris, France, shows a short-term white frequency noise characterized by an Allan deviation $\sigma_y(\tau = 1 \text{ s}) = 2 \times 10^{-13}$ and a type B uncertainty of 3×10^{-15} (1 σ) [1]. In addition, it is evaluated over averaging durations of order 10 hours. Frequency standards based on linear ion traps also show excellent short-term stability: their white frequency noise is detected at a level of $\sigma_y(\tau = 1 \text{ s}) = 7 \times 10^{-14}$ [2]. To compare the frequencies of such high-performance standards would call for GPS common-view results to be averaged over 10 days of continuous operation whereas the optimal performance of the standards is reached over much shorter intervals.

Previous studies have shown the potential of carrier-phase and code measurements from geodetic GPS receivers for clock comparisons [3, 4, 5], and have drawn up tentative uncertainty budgets for clock frequency comparisons using phase and code measurements from commercial dual-frequency and multi-channel GPS receivers [6]. This paper describes a test experiment, in which GPS carrier phase measurements are used for frequency comparisons. Examples are given of frequency comparison between clocks located at the BIPM, the NPL, Teddington, United Kingdom, the ORB, Brussels, Belgium, and other stations of the International GPS Service for Geodynamics (IGS). The GPS data are taken from Allen Osborne Associates (AOA) TTR-4P GPS receivers at the BIPM and the NPL, and from AOA Turbo Rogue receivers at the IGS sites. This experimental set-up will be upgraded by incorporating additional GPS receivers from different makers, that will be compared to provide an estimate of hardware instabilities.

2. EXPECTED UNCERTAINTY IN FREQUENCY COMPARISON

A theoretical study has been carried out [6] to estimate the expected uncertainty of the frequency comparison using GPS carrier phase measurements. Figure 1 summarizes the results by showing curves of Modified Allan deviations for averaging durations ranging from 30 s to 1 d using dual-frequency code or carrier-phase measurements.

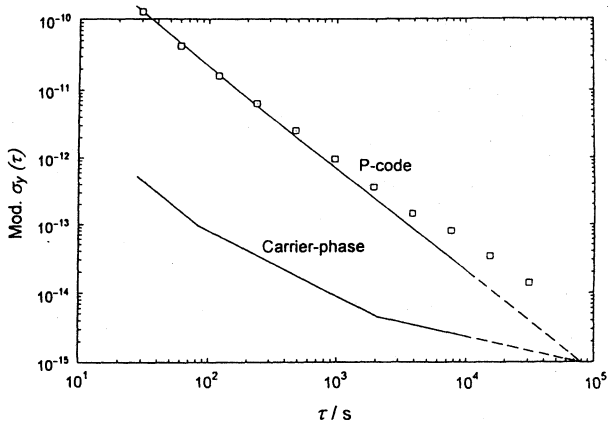


Figure 1: Tentative curves for the Modified Allan deviation characterizing frequency transfer through ionosphere-free code and carrier-phase measurements of all satellites in common view for baselines of up to several thousand kilometres.

The curve shown for code measurements is derived by assuming that white measurement noise averages down according to the $\tau^{-3/2}$ law, thus neglecting possible hardware variations. In this case, an uncertainty of 1×10^{-15} is achieved for a frequency transfer over one day. The curve for carrier-phase measurements has three separate parts. For short averaging durations it follows the $\tau^{-3/2}$ variation caused by white phase noise (measurement, short term troposphere). For the middle term (100 s < τ < 2000 s) the predominant error sources (troposphere, multipath) integrate in a flicker-phase noise represented by a τ^{-1} curve. For longer averaging durations it is necessary to merge data from different satellite paths so the main sources of uncertainty (propagation, station and satellite coordinates) may cause ambiguities in phase restitution. Again neglecting hardware variations due to the environment, phase ambiguities may be the predominant source of uncertainty, giving rise to white frequency noise represented by a $\tau^{-1/2}$ curve reaching a level of 1 part in 10^{15} at one day.

Based on the results of previous studies [7, 8], it is expected that environmental effects on the non-digital part of the hardware may be the dominant source of uncertainty for averaging durations exceeding a few hours. In some cases, the effect of external temperature on the antennas has been evaluated: Expressed as a coefficient linking delays to temperatures, values are found which range from tens to hundreds of picoseconds per degree Celcius. This corresponds to uncertainties in a frequency comparison ranging from 2 to 20 parts in 10^{15} for averaging durations of half a day.

3. TEST EXPERIMENT

An experiment has been set-up to explore the use of GPS carrier phase for frequency comparisons in time laboratories. Currently it involves only a few laboratories and a single type of GPS receiver (AOA TTR-4P), but it will be expanded in the future.

3.1 Experimental set-up

Figure 2 shows the set-up for the present configuration of the experiment. This makes it possible to obtain comparisons between pairs of clocks using both a dual frequency GPS phase technique and the classical single frequency C/A code common view method based on data from GPS receivers of the NBS design. As the uncertainty of the GPS phase technique is likely to be the smaller, the comparison between the two techniques will be meaningful only if it is carried out over many continuous days.

The laboratories involved are the BIPM in Sèvres, France, the NPL in Teddington, United Kingdom, and the ORB in Brussels, Belgium. Some of the equipment involved is in continuous operation and the corresponding data are already being recorded for other purposes, greatly simplifying data collection. This is the case for the NBS type receivers at NPL and ORB, which transmit data to the BIPM for the TAI computation, and for the Turbo-Rogue receiver at ORB, which is part of the IGS network. In addition, GPS data from all other stations of the IGS network may be used as they are freely available on the Internet [9].

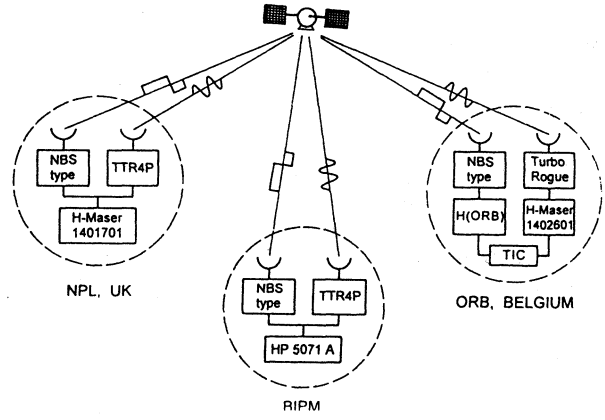


Figure 2: Experimental set-up at the three time laboratories involved. Baselines range from 200 km to 300 km.

3.2 Processing strategy

Given that the network considered in this study has only a few stations, and that the baselines are short (a few hundred kilometres in general), the approach used has been to write a simple software that takes advantage of the results of IGS computations (for station positions and satellite ephemerides) and adjusts parameters only for troposphere and phase ambiguities. IGS results have been taken from the CODE processing centre at the University of Bern. The processing is a multi-step procedure for which the main lines are described below.

The raw dual-frequency phase and pseudo-range measurements are combined to give ionosphere-free carrier-smoothed pseudo-range values for each satellite pass (uninterrupted observation by a station). These values are measurements of $RefSV_i$, the difference between the reference clocks of the station and satellite i. These measurements are combined to form single differences for stations A and B,

$$S_i(AB) = [RefSV_i(A)-RefSV_i(B)] - [T_i(A)-T_i(B)] \quad (1)$$

where T represents values of the time of propagation derived from the best a priori model, computed with IGS station coordinates, IGS satellite ephemerides and a tropospheric model.

Double differences $D_{ij}(AB)$ are formed for all simultaneous quantities $S_i(AB)$ and $S_j(AB)$ and used to solve for the zenith tropospheric delay correction (one value per station, valid for about 6 hours) and phase ambiguities. The computed tropospheric delay corrections and phase ambiguities are then used in (1) to refine the single difference data $S_i(AB)$ for the stations of interest, from which we obtain an estimate of the clock comparison Ref(A)-Ref(B). In this procedure, no attempt is made to ensure the synchronization of the clocks, but the time evolution of Ref(A)-Ref(B) provides the frequency comparison.

In the analysis conducted on this data set, the station coordinates and the antenna offsets for the IGS stations were taken from the ITRF94 reference frame. For the BIPM station, they were obtained from a GPS survey, conducted by the French geographic institute IGN. For the NPL, the coordinates of the phase centre of the antenna have been estimated from the data themselves. A correction for solid Earth tides was applied. Satellite ephemerides were obtained from the IGS CODE processing centre. Overall, it is estimated that the geometric model used provides an uncertainty in the computation of single differences S and double differences D which is below 100 ps. This value, of course, relates only to the rather small regional network considered in this study.

3.3 Preliminary results

Two days of continuous data (MJD 50481 and 50482, February 2 and 3, 1997) were used for the preliminary analysis. The GPS phase receivers at BIPM and ORB operated continuously without reset. The TTR-4P receiver at NPL had a phase reset and so does not directly provide a continuous link over the two day interval. Figure 3 shows results for the link ORB-BIPM and figure 4 shows results for the link between ORB and another IGS station, in Matera (Italy), both of which are equipped with an hydrogen maser.

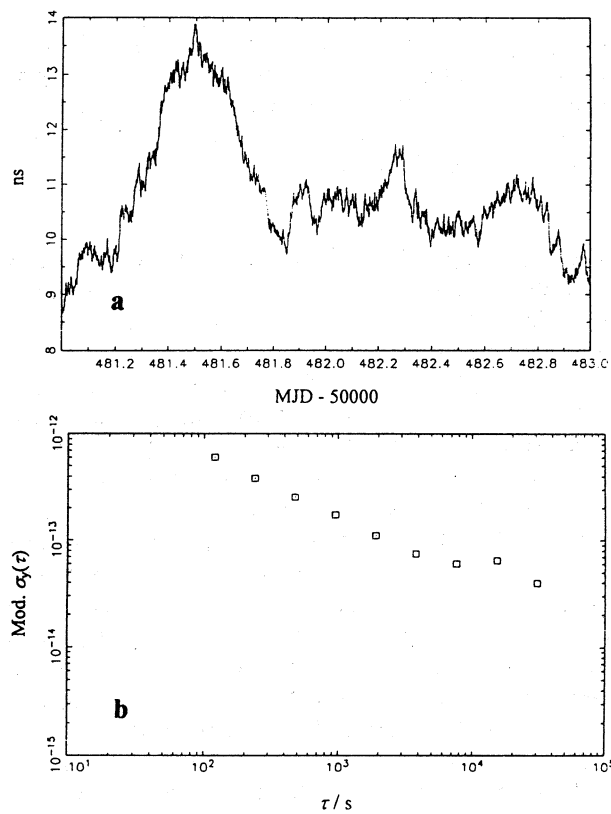


Figure 3: Comparison of a hydrogen maser in operation at the ORB with a HP5071A clock at the BIPM by ionosphere-free carrier-phase measurements of all satellites in common view:

3-a Time differences (constant rate removed)
3-b Modified Allan deviation

The stability implied by Figure 3-b is consistent with the specifications of the BIPM clock and so does not represent the limit of what can be obtained using the GPS phase technique. It is probable, however, that at least part of the change in the time differences data (Figure 3-a) around date 50481,5 is due to some sensitivity to the environment as it is strongly correlated with a simultaneous change in the external temperature at the BIPM. No large temperature change was observed on day 50482 where the time differences are smoother.

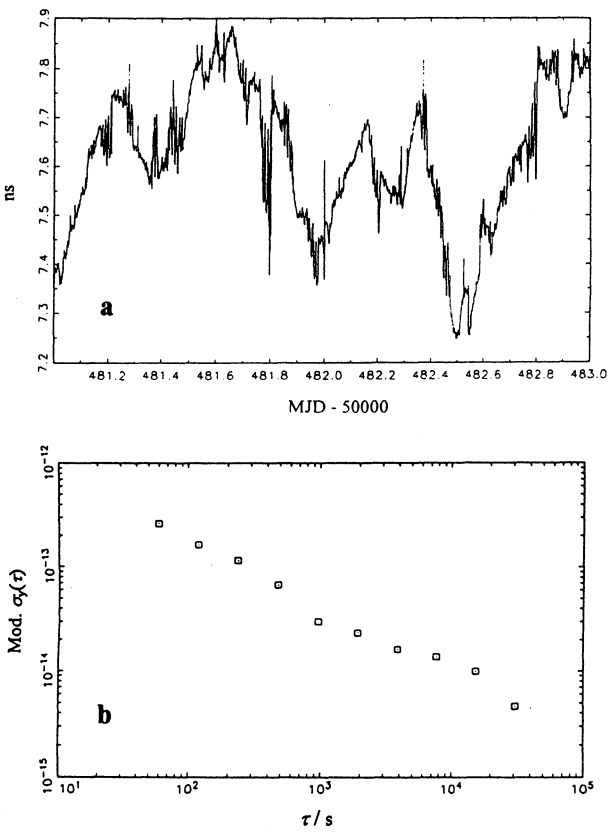


Figure 4: Comparison of two hydrogen masers in operation at the ORB and Matera by ionosphere-free carrier-phase measurements of all satellites in common view:

4-a Time differences (constant rate removed)
4-b Modified Allan deviation

The stability implied by Figure 4-b reaches 5×10^{-15} for an averaging duration of less than 10 hours. As the residual noise cannot be attributed to the masers, it must represent the effect of the environment on this particular type of equipment. The short term behavior of the data in Figure 4-a has its origin in the fact that no filtering was applied and no attempt was made to take care of some small parts of the data set where multipath effects are important. The values obtained could be improved by processing, but that would not change the limit for longer averaging durations.

4. CONCLUSIONS

Previous studies already indicated the potential of GPS carrier phase measurements for frequency comparisons. The achievable uncertainty is estimated to be about 1×10^{-14} for an averaging duration of one hour and about 1×10^{-15} for an averaging duration of one day. This range of durations seems to be optimal for this technique: It is convenient for experimenters who wish to compare clocks in a given experimental configuration and require that the uncertainty in the frequency comparison is comparable with the stabilities of the clocks. The technique can also be used for longer averaging durations (days) but then provides uncertainties similar to those based on using GPS code measurements only.

As previous studies also indicate, the main source of uncertainty, possibly preventing uncertainties from reaching the levels quoted above, is the sensitivity of the hardware to environmental variations. The main part of this sensitivity arises in the antenna when it is subject to large external environmental variations.

Preliminary results from a 2-day subset of the data obtained during a test experiment confirm both initial conclusions: An uncertainty of several parts in 10^{15} for an averaging duration of 10 hours in the frequency comparison of two hydrogen masers with Turbo-Rogue receivers is achieved, but is probably limited by hardware variations. Data from TTR-4P receivers have also be incorporated in the data processing and produce valid results. It seems, however, that the sensitivity of this receiver to the environmental variations is larger, but this needs to be studied in more detail.

Future studies will include systematic comparisons of different GPS phase/code receivers, each optimized for time and frequency applications. In addition to the TTR-4P from AOA, the Z12-T receiver from Ashtech and the R100/30 receiver from 3S-Navigation will be introduced.

ACKNOWLEDGEMENTS

I am indebted to the colleagues who provided help and data: J.A. Davis (NPL), P. Defraigne (ORB) and P. Moussay (BIPM). The role of the IGS in providing data and results is gratefully acknowledged. I thank Claudine Thomas for her support and for helpful discussions, and D.A. Blackburn for revision of the manuscript.

REFERENCES

[1] A. Clairon., S. Ghezali, G. Santarelli, P. Laurent, S.N. Lea, M. Bahoura, E. Simon, S. Weyers and K. Szymaniec, "Preliminary Accuracy Evaluation of a Caesium Fountain Frequency Standard", Proc. Symp. on Frequency Standards and Metrology, J.C. Bergquist Ed., World Scientific, 1995, pp. 49-59.

[2] R.J. Tjoelker, J.D. Prestage and L. Maleki, "Record Frequency Stability with Mercury in a Linear Ion

Trap", Proc. Symp. on Frequency Standards and Metrology, J.C. Bergquist Ed., World Scientific, 1995, pp. 33-38.
[3] C.E. Dunn, S.M. Lichten, D.C. Jefferson and J.S. Border, "Sub-nanosecond Clock Synchronization and Precision Deep Space Tracking", Proc. 23rd PTTI, 1991, pp. 89-99.
[4] R.J. Douglas and J. Popelar, "PTTI Applications at the Limits of GPS", Proc 26th PTTI, 1994, pp. 141-152.
[5] P. Baeriswyl, T. Schildknecht, J. Utzinger and G. Beutler, "Frequency and Time Transfer with Geodetic GPS Receivers: First Results", Proc. 9th EFTF, 1995, pp. 46-51.
[6] G. Petit and C. Thomas, "GPS Frequency Transfer Using Carrier-phase measurements", Proc. 50th Frequency Control Symposium, 1151-1158, 1996.
[7] W. Lewandowski and P. Moussay, "Temperature Protected Antennas for Satellite Time Transfer Receivers", Proc. 11th EFTF, 1997.
[8] F. Overney, L. Prost, U. Feller, T. Schildknecht and G. Beutler, "GPS Time Transfer using Geodesic Receivers: Middle-term Stability and Temperature Dependency of the Signal Delay", Proc. 11th EFTF, 1997.
[9] IGS Annual report 1994, JPL publication 95-18 9/95, Jet Propulsion Laboratory, Pasadena, CA, USA, 1995.

Maser clocks for GPS: A joint experiment

G. Busca
Observatoire de Neuchâtel
Rue de l'Observatoire 58, CH-2000 Neuchâtel

Robert F.C. Vessot
Smithsonian Astrophysical Observatory
Cambridge, Massachusetts 02138, USA

Summary

The Space Hydrogen Maser (SHM) has been identified as a possible future GPS clock for improving the GPS system performances.

Actually, SHM's of the active type and of quite different designs have been already developed or are presently in the course of development, in the frame of non-GPS related NASA and ESA programmes.

Drift removed SHM frequency stability appears capable of challenging the performances also of other space clocks which could become available in future as ion trap, cold Cs atom or improved conventional Cs beam clocks.

The availability of 153 kg of excess payload space on the GPS II F satellites, coupled with the present availability of SHM technology offers the interesting possibility of assessing this technology for GPS applications directly on a GPS satellite. This is the scope of the proposed experiment.

Several configuration of the experiment are presented involving the use of intersatellite links, standard GPS signals and laser time transfer system for evaluating the masers frequency stability and the improved navigation performances. Possible impact of the experiment on other fields of science as time synchronisation and earth observation is also illustrated.

MEASUREMENT TECHNIQUES

Chairman: E. Graf

STANDARD UNCERTAINTY FOR AVERAGE FREQUENCY TRACEABILITY

R.J. Douglas and J.-S. Boulanger
Frequency and Time Standards Group, Institute for National Measurement Standards
National Research Council of Canada, Ottawa, CANADA K1A 0R6
(613) 993-5186, Fax (613) 952-1394, rob.douglas@nrc.ca

A simple method (at the hand calculator level) is presented for using the Allan variance $\sigma_y^2(\tau)$ to estimate the standard uncertainty of average frequency due to the transfer process from one general time interval to another. The method is applicable when the measured $\sigma_y(\tau)$ reveals that the stability can be modelled by a sum of phase noise (white and flicker phase noise), white frequency noise, flicker frequency noise and random walk frequency noise.

Uncertainty, standard uncertainty, average frequency, frequency metrology, traceability.

1. INTRODUCTION

The recommendations of the *Guide to the Expression of Uncertainty in Measurement* [1] have increased demand for the standard uncertainty of all measurands - including SI time interval, or average frequency. Clients delighted by the ease of making excellent frequency comparisons are usually discouraged by the intricacies of non-white noise characterization, which compound as they try obtain a standard uncertainty. A means is needed for time and frequency metrologists to translate their understanding of stability into the standard uncertainty wanted by their clients.

In general the measurand is a (proper) time interval $[t_1, t_2]$ between two events, or equivalently the average frequency of a frequency standard over this interval. The underlying mathematical abstraction is the instantaneous normalized frequency departure, $y(t)$: the frequency departure from nominal, normalized to the nominal frequency. The average frequency of a frequency standard over $[t_1, t_2]$ can be expressed as $\bar{y}_{[t_1,t_2]} = \int_{t_1}^{t_2} y(t)dt$. Frequency calibration involves a comparison with a reference frequency standard, and in many practical cases is done by scaling $\bar{y}[t_1, t_2]$ to be zero for the whole calibration interval $[t_1, t_2]$, although other calibrating fits are also used. It is convenient to define the indefinite integral over $y(t)$, usually referred to as the "phase" of the frequency standard relative to an ideal standard: $x(t) = \int_0^t y(t')dt'$.

Frequency standard stability is characterized by comparing the average frequency over adjoining, equal time intervals. Considering the average frequency over intervals of duration τ as a time series, the Allan or two-sample variance is $\sigma_y^2(\tau) = \left\langle \frac{1}{2} \left(\bar{y}_{[t,t+\tau]} - \bar{y}_{[t+\tau,t+2\tau]} \right)^2 \right\rangle$. Many frequency standards are free from periodic variations, and the residual noise is found to be random and normal but not necessarily white. The *Guide* considers the case of white noise processes which are uncorrelated in time. In addition to white frequency noise, four other non-white power-law spectral distributions are commonly identified, and will be discussed below.

2. TRACEABLE *A* to *B* CALIBRATION

A frequency standard is calibrated for average frequency over one time interval A, and used as a transfer standard to establish an average frequency for a different time interval B (either shorter or longer) at some later time, as illustrated in Figure 1. Occasionally the end-use time interval B is earlier than the calibration, or partially overlaps with the calibration, or is entirely within the calibration interval.

Frequency or phase comparisons are made across A between the transfer oscillator and a SI frequency standard, fitted and used to predict the SI frequency across the interval B. The fit may include frequency drift terms or periodic variations. The fitting process may be any linear combination of the frequency differences measured across subintervals of A, and be analysed in much the same way that the Allan variance has been analysed, in terms of autocorrelation of the phase $x(t)$ or may include any linear combination of the phase comparisons across the subintervals of A by using methods discussed in [2].

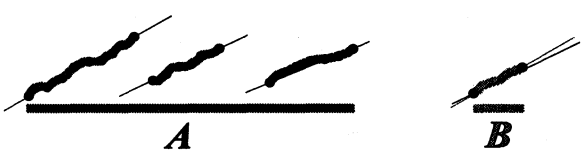


Figure 1: General Frequency Calibration over A for use over interval B. At B, the predicted (solid line) and actual average frequency (dotted line) have an rms deviation which is the standard uncertainty of the frequency transfer process.

3. STANDARD UNCERTAINTY at B

The standard uncertainty needed for quantitatively documenting traceability in frequency transport from interval A to B is to be thought of as an ensemble average of the entire process: with each member of the ensemble having a separate embodiment of both the SI frequency reference and the transfer standard. The ensemble average of the frequency references at B is the appropriate reference frequency that should be used to judge the quality of the calibrated frequency derived from the transfer standard at B. This choice of reference allows correct analysis of cases where the transfer standard is more stable than the reference frequency standard, where the overall standard uncertainty of the transfer oscillator at B can be less than the standard uncertainty of the frequency reference at B, although not less than the standard uncertainty of the frequency reference across A. Where the reference standard is well-behaved, one can separate the standard uncertainty due to the reference standard across A, and the standard uncertainty due to the instability of the transfer standard relative to an ideal transfer standard. This is a significant simplification, since procedures for characterizing transfer oscillators relative to an ideal reference are well established, using the Allan variance or the modified Allan variance [?, 4].

4. GENERAL A-B STRUCTURE FACTORS

Most calibration processes will use an analytic function with parameters which are linear combinations of the phase measurements made over the calibration interval. Its standard uncertainty will depend on the stability of the transfer oscillator through the autocorrelation function of $x(t)$, $\mathcal{R}(\tau) =< x(t)x(t-\tau) >$. If the phase measurements are made at times $\{t_i\}$ scattered across intervals A and B, then the square of the standard uncertainty will depend on a sum over $\mathcal{R}(t_i - t_j)$, with a weight determined from the details of the fitting process used. We refer to this sum over differences between the lattice of sampling times as the A-B structure factor for the calibration process. For a pure type of power law noise, it is often convenient to normalize it to the Allan variance for some specific time.

5. END-POINT CALIBRATION

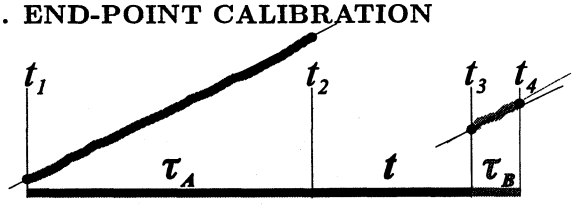


Figure 2: End-Point Frequency Calibration over A for use over interval B.

The average frequency over the calibrating interval, illustrated in Figure 2, is the simplest and often is also the most accurate frequency estimator of the reference standard relative to an ideal standard : $[t_1, t_2]$, is $(x(t_2) - x(t_1))/(t_2 - t_1)$. This is the frequency estimate which will be used over the second time interval $[t_3, t_4]$, although the average frequency over this interval, in terms of $x(t)$, is $(x(t_4) - x(t_3))/(t_4 - t_3)$. The variance in the average-frequency transfer process is just the mean-square difference between these two frequencies, or

$$\left\langle \left\{ \left[\frac{x(t_4) - x(t_3)}{t_4 - t_3} \right] - \left[\frac{x(t_2) - x(t_1)}{t_2 - t_1} \right] \right\}^2 \right\rangle. \quad (1)$$

The variance can also be expressed in terms of the autocorrelation function of $x(t)$, $\mathcal{R}(\tau) = \langle x(t)x(t-\tau) \rangle$. As shown in Figure 2, it is convenient to denote the duration of the calibration interval as $\tau_A = (t_2 - t_1)$, the duration of the end-use interval as $\tau_B = (t_4 - t_3)$ and the duration of the "dead time" between the intervals as $t = (t_3 - t_2)$. Intervals τ_A and τ_B which have any overlap would have negative "dead time" t. Collecting the product terms in $x(t)$ within the average and changing to autocorrelation function notation gives the variance as

$$2\left[\frac{\mathcal{R}(t) + \mathcal{R}(\tau_A + t + \tau_B) - \mathcal{R}(t + \tau_A) - \mathcal{R}(\tau_B + t)}{\tau_A \tau_B} + \frac{\mathcal{R}(0) - \mathcal{R}(\tau_A)}{\tau_A^2} + \frac{\mathcal{R}(0) - \mathcal{R}(\tau_B)}{\tau_B^2} \right]. \quad (2)$$

We want a strategy which will allow us to estimate $\mathcal{R}(t)$, or more simply *differences* of the form $\mathcal{R}(t) - \mathcal{R}(t')$. $\mathcal{R}(t)$ can be calculated by integrating over the spectral distribution of noise, where experience shows that spectral smoothing of broadband noise is very helpful. Five power law noises (self-similar or fractal-like) are usually used to describe the broadband noise of a frequency standard.

We consider a transfer oscillator with a normalized $y(t)$, which fluctuates with a power spectral distribution $S_y(f) = \sum_{\alpha=-2}^{+2} h_\alpha f^\alpha$. We will assume an upper cutoff frequency $\omega/2\pi$ which models the noise equivalent bandwidth of the measurement systems, and a lower frequency cutoff $\epsilon/2\pi$ which models the finite time over which the power law noise model has been verified. The assumptions also establish stationarity and ergodicity for the autocorrelation functions that we will combine to form uncertainty estimates which will themselves be stationary in the limits as $\omega \to \infty$ and $\epsilon \to 0$. We will show that for many practical purposes the details of the band limits may be ignored, provided that they are the same for the stability characterizing measurements, the calibration measurements and the final-use measurements.

The transfer oscillator's $x(t)$ has an autocorrelation function which can be calculated as an integral linear in $S_y(f)$: $\mathcal{R}(\tau) = \int_0^\infty S_x(f)cos(2\pi f\tau)df = (1/2\pi)\int_0^\infty [S_y(f)/(2\pi f)^2]cos(2\pi f\tau)df$. For all purposes involving only differences of the form $[\mathcal{R}(t_a) - \mathcal{R}(t_b)]$, a less divergent function may be used instead of $\mathcal{R}(\tau)$: $\mathcal{I}(\tau) = \int_0^\infty [S_y(f)/(2\pi f)^2][1 - cos(2\pi f\tau)]df$. For our noise model, $\mathcal{I}(\tau) = \sum_{\alpha=-2}^{+2} |\tau|^{1-\alpha}(2\pi)^{-(\alpha+1)} \int_{\epsilon|\tau|}^{\omega|\tau|} h_\alpha u^{\alpha-2}\{1 - cosu\}du = \sum_{\alpha=-2}^{+2} I_\alpha(\tau)$. Complete analytic forms exist [2, 5].

6. A-B STRUCTURE FACTORS SEPARATED BY POWER-LAW NOISE TYPE

By separately considering each α, we can obtain an easier-to-use form of Eq. 2 by switching to the $I_\alpha(\tau)$ form and normalizing it to the Allan variance at some time τ_0 for each type of noise:

$$2\left[\frac{I_\alpha(t+\tau_A) + I_\alpha(\tau_B+t) - I_\alpha(t) - I_\alpha(\tau_A+t+\tau_B)}{\tau_A\tau_B} + \frac{I_\alpha(\tau_A)}{\tau_A^2} + \frac{I_\alpha(\tau_B)}{\tau_B^2}\right] / \left[(4I_\alpha(\tau_0) - I_\alpha(2\tau_0))/\tau_0^2\right]. \quad (3)$$

The sum of these factors, multiplied by the observed Allan variance, will give the square of the standard uncertainty in the frequency transport from A to B. We take the noise band limits to be the same in the processes for determining the Allan deviation, for interval A, and for interval B, and then take the limits $\epsilon \to 0$ and $\omega \to \infty$.

For each of four classes of noise (one class combining white phase noise ($\alpha = 2$) with flicker phase noise ($\alpha = 1$)) we calculate the limiting ratio of the square of the standard uncertainty in the average frequency $u_y^2(\tau_A, t, \tau_B)$ to the Allan variance $\sigma_y^2(\tau_0)$ for that noise class alone, denoted as $s_{y[\alpha]}^2$.

7. RANDOM WALK FREQUENCY NOISE: $\alpha = -2$

This is the most divergent type of random noise commonly associated with frequency standards. It is the noise type that limits the performance of quartz oscillators, of rubidium vapour frequency standards, and of hydrogen masers at the longest range of time intervals before their deterministic frequency drift becomes evident. Frequency noise of this type would be called "pink" noise, referring to the excess of fluctuations with long periods compared to white frequency noise. The Allan deviation for this type of noise is $s_{y[\alpha=-2]}(\tau/\tau_0)^{1/2}$.

$$\mathcal{F}_{AB[\alpha=-2]} = \frac{1}{[2\tau_0\tau_A\tau_B]}\Big[|t|^3 + |t+\tau_A+\tau_B|^3 - \tau_A^2\tau_B - \tau_A\tau_B^2 - |t+\tau_A|^3 - |t+\tau_B|^3\Big] \quad (4)$$

For extrapolation, where $t > 0$, the expression reduces to $[3t+\tau_A+\tau_B]/\tau_0$. In other words the standard uncertainty is the sum in quadrature of the Allan deviation random walk frequency modulation (FM) line at the first interval τ_A with that at the second interval τ_B with three times that at the dead time t. For extrapolations much larger than τ_A or τ_B the standard uncertainty is $\sqrt{3}$ times the Allan deviation at t.

For $t = 0$ and $\tau_A = \tau_B$, we find the expected ratio: the standard uncertainty is $\sqrt{2}$ times the Allan deviation random walk FM line at τ_A.

For centred interpolation, with τ_B centred on the interval τ_A, i.e. $t = -(\tau_A + \tau_B)/2$, the standard uncertainty becomes $(1 - \tau_B/\tau_A)/2$ times the Allan deviation random-walk FM line at τ_A. For $\tau_B << \tau_A$, the standard uncertainty is one-half the Allan deviation at τ_A. As τ_B approaches τ_A, the standard uncertainty in the average frequency transfer approaches zero, reflecting the increasing correlation between the two intervals.

Thus, in considering the general use of a calibration interval over a time τ_A where the noise is dominated by random walk FM, the Allan deviation at τ_A is revealed as a poor generic estimator of the standard uncertainty: overestimating the standard uncertainty by up to two-fold or more for short intervals within the calibration interval, and underestimating the standard uncertainty by a factor of $\sqrt{3n}$ if the calibration interval is n calibration intervals old.

8. FLICKER FREQUENCY NOISE: $\alpha = -1$

This type of noise is commonly seen as the flat Allan deviation (the "flicker floor") for most types of frequency standards. Like the random walk FM, this type of frequency noise could also be called "pink" noise, referring to the excess of fluctuations with long periods compared to white frequency noise. The Allan deviation for this type of noise is $s_{y[\alpha=-1]}$.

$$\mathcal{F}_{AB[\alpha=-1]} = \frac{1}{[2ln2]}\left[ln\left|\frac{(\tau_A+t+\tau_B)^2}{\tau_A\tau_B}\right| + \frac{2t+\tau_A}{\tau_B}ln\left|\frac{\tau_A+t+\tau_B}{t+\tau_A}\right| + \frac{2t+\tau_B}{\tau_A}ln\left|\frac{\tau_A+t+\tau_B}{t+\tau_B}\right| + \frac{t^2}{\tau_A\tau_B}ln\left|\frac{t(\tau_A+t+\tau_B)}{(t+\tau_A)(t+\tau_B)}\right|\right] \quad (5)$$

For extrapolation of the frequency, as $t \to \infty$, the ratio of standard uncertainty to Allan deviation approaches $\sqrt{ln[t^2/(\tau_A\tau_B)]/(2ln[2])}$.

For $t = 0$ and $\tau_A = \tau_B$, we find the expected ratio: the standard uncertainty is $\sqrt{2}$ times the Allan deviation (flicker FM) line at τ_A.

9. WHITE FREQUENCY NOISE: $\alpha = 0$

Usually, this type of spectral distribution has been the only one considered when standard uncertainties are being estimated. The time separation of the averages does not matter, and the only correlation between two averaging intervals occurs when the intervals overlap at least in part. The Allan deviation is $s_{y[\alpha=0]}(\tau/\tau_0)^{-1/2}$.

$$\mathcal{F}_{AB[\alpha=0]} = \frac{\tau_0}{\tau_A \tau_B}\Big[|t+\tau_A| + |t+\tau_B| + \tau_A + \tau_B - |t+\tau_A+\tau_B| - |t|\Big] \qquad (6)$$

For non-overlapping intervals, where $t > 0$ or $t < -[\tau_A + \tau_B]$, this becomes $[\tau_A^{-1} + \tau_B^{-1}]/\tau_0^{-1}$, the sum of the fractional variances of the calibration and end-use intervals, normalized to the fractional variance at the Allan variance time.

10. PHASE NOISE: $\alpha = 1$ and $\alpha = 2$

These types of noise are common in all measuring systems and phase-lock loops. White voltage noise in time comparators and phase lock loops often results in white phase noise, and flicker (or 1/f) voltage noise in time comparators or phase lock loops will give flicker phase noise. By analogy with white and pink frequency noise described above, they might both loosely be described as "violet noise" in frequency: having more fluctuations for the shorter periods than for longer periods. The Allan deviation curve does not distinguish these two types of noise, and for some purposes it can be very important to do so: for these applications the "modified" Allan deviation is used. The modified Allan deviation is somewhat more demanding than the Allan deviation in that it uses a continuous data stream which is not always easy to ensure in real-world applications where long-term measurements are occasionally interrupted or corrupted. If possible, we would choose not to require its use in estimating standard uncertainties.

The Allan deviation's insensitivity to the type of phase noise led us to investigate the extent to which the distinction was required for estimating standard uncertainties. Where these two noise types do not need to be treated separately, one might hope that the details of the true shape of the noise bandwidth might be ignored for many systems, provided only that they are then same when characterized by the Allan variance, when used in the calibration interval and when employed for final use. For many uses in metrology, where it suffices to estimate the uncertainty with an uncertainty-in-the-uncertainty of 10% to 20%, we were pleased to find that this is the case, provided that no end-point of either interval A or B is any closer than 10 times the period of the high-frequency cutoff (which models the noise equivalent bandwidth for white noise). The Allan deviation for this type of noise is $s_{y[\alpha=1,2]}(\tau_0/\tau)$.

Considering each type of noise separately, the limit $\epsilon \to 0$ presents no difficulty for either noise type, and their formal limits as $\omega \to \infty$ are the same, namely

$$\mathcal{F}_{AB[\alpha=1,2]} = \frac{2}{3}\Big[\tau_A^{-2} + \tau_B^{-2}\Big]/\tau_0^{-2}. \qquad (7)$$

The form is chosen to emphasize that this is the sum of the fractional variances of the calibration and end-use intervals, normalized to the fractional variance at the Allan variance time. The $\omega \to \infty$ limit means that the expression applies only where each of the four end points of the intervals are "many" times $1/\omega$ away from any of the other three end points, and that $\omega\tau_0 >> 1$ as well. When the intervals τ_A and τ_B share an end point, there is a full correlation of the phase noise at that point, and $\mathcal{F}_{AB[\alpha=1,2]}$ is increased or decreased by $2\tau_A\tau_B/(3\tau_0^2)$. The variance is increased by this amount when the common point is inside the other two points (e.g. $\mathcal{F}_{AB[\alpha=1,2]} = 2\tau_0/\tau_B$ for $\tau_A = \tau_B$, and $t = 0$) and decreased by this amount when the common point is an outermost point.

11. STANDARD UNCERTAINTY for MIXED NOISE

For establishing the standard uncertainty of an unknown transfer oscillator for a particular end-point calibration process, it would first be measured with respect to an ideal oscillator (perhaps with the aid of three-cornered hat variance analysis if a single sufficiently good reference oscillator is not available). The noise bandwidth of these measurements must be the same as the noise bandwidth to be used for calibration across interval A and across the end-use interval B, or else an extremely careful characterization of the effective phase noise must be done using the modified Allan deviation to characterize the amounts of white phase and flicker phase noise, and hence the Allan deviation expected in the noise bandwidth to be used in the intervals A and B. If the measurements show a that there are no significant deterministic frequency changes over the anticipated period of a calibration cycle, and the Allan deviation graph appears to be adequately described by a power-law decomposition; then the standard uncertainty of the average frequency over τ_B determined by the end-point calibration over τ_A may be calculated as follows: **(1)** At some convenient value of τ_0 on the Allan deviation graph, the four intercepts of the decomposition are found $s_{y[\alpha=1]}$ for τ^{-1} (combined $\alpha = 1$ and 2), $s_{y[\alpha=0]}$ for $\tau^{-1/2}$, $s_{y[\alpha=-1]}$ for τ^0 and $s_{y[\alpha=-2]}$ for $\tau^{1/2}$. **(2)** The four A-B structure

factors $\mathcal{F}_{AB[\alpha]}$ are computed according to Equations 4-7, using the appropriate values of τ_0, τ_A, t and τ_B. **(3)** The standard deviation of all four noise classes are added in quadrature to give the final standard deviation expected for the process:

$$u_y(\tau_A, t, \tau_B) = \sqrt{\sum_{\alpha=-2}^{\alpha=1} s^2_{y[\alpha]} \mathcal{F}_{AB[\alpha]}} \qquad (8)$$

12. DEGREES OF FREEDOM

The degrees of freedom ν_i is the means recommended by the *Guide* [1] to report the uncertainty-in-u_y for the transfer of average frequency from interval τ_A to τ_B. As with the degrees of freedom for the Allan deviation, the degrees of freedom will depend on the time intervals: τ_A, t and τ_B. If the frequency-transfer process dominates the overall uncertainty the most important range of degrees of freedom is from 1 to 15. If other sources dominate the final uncertainty, the upper limit will be even lower. Beyond 15 degrees of freedom, the uncertainty in u_y is less than 20% and although greater precision may be beautiful it will rarely benefit a metrological client. A statement like "The degrees of freedom associated with the determination of u_y exceeds 15..." should therefore suffice.

If the Allan deviation used for the s_y's in Equation 8 is measured in the usual way (continuous phase comparisons) over an interval which is N times $[\tau_A + t + \tau_B]$, then the degrees of freedom should exceed N, since N independent tests of the calibration cycle would fit in the interval.

One simple guide to the degrees of freedom is available [6, 7] from the degrees of freedom for experimental determination of the Allan deviation $\sigma_y(\tau)$ using overlapping samples. This estimate depends on the noise type, on the time interval τ and on the total time $N\tau$ devoted to measuring $\sigma_y(\tau)$. The degrees of freedom for u_y will be less than or equal to the degrees of freedom for the $\sigma_y([\tau_A + t + \tau_B]/2)$ used in Equation 8, and might be expected to be greater than the degrees of freedom for $\sigma_y([\tau_A + t + \tau_B])$. Within the confines of the noise model, procedures could be developed using established methods [7] to evaluate more rigorously the degrees of freedom.

For frequency transfer, there are also significant disadvantages in relying on the degrees of freedom to express the uncertainty in u_y. This proper use of the degrees of freedom must be distinguished from uncritical use of the degrees of freedom for the refinement of the standard deviation of the mean for frequency transfer: in some cases this will apply, and in others the correlations between repetitions may be important and need to be estimated [5]. In attempting to address the confidence in the quoted uncertainty, the degrees of freedom does not address the issues of conformance to the noise model and frequency calibration model (no frequency drift and no deterministic periodic terms). If the procedures outlined above were applied thoughtlessly to a standard with a large drift or diurnal term, the uncertainty estimate could be hopelessly wrong. If sufficient data is available, a reduced χ^2 approach would probably be better.

13. CONCLUSIONS

We have presented a simple method for estimating standard uncertainty in frequency transfer which is widely applicable in traceability documentation discussed above. The method would also simplify for analysis of time scale work [2], pulsed frequency standards [8], intermittent cesium fountain operation [5], and on-site calibration of GPS disciplined oscillators [9].

References

[1] *Guide to the Expression of Uncertainty in Measurement,* First Edition, International Organization for Standardization, Technical Advisory Group on Metrology, Working Group 3 (1992), pub. in final form in 1993 in the name of BIPM, IEC, IFCC, ISO, IUPAC, IUPAP and OIML.

[2] R.J. Douglas, J.-S. Boulanger and C. Jacques: Proc. 25^{th} Precise Time and Time Interval Meeting, pp 249-266 (1993).

[3] J.A. Barnes, A.R. Chi, L.S. Cutler, D.J. Healey, D.B. Leeson, T.E. McGunigal, J.A. Mullen Jr., W.L. Smith, R.L. Sydnor, R.C. Vessot and G.M.R. Winkler: IEEE Trans. Instrum. Meas. **IM-20** pp 105-120 (1971).

[4] D.W. Allan and J.A. Barnes: Proc. 35^{th} Ann. Freq. Control Symp., pp 470-474 (1981).

[5] D. Morris, R.J. Douglas and J.-S. Boulanger: Jpn. J. App. Phys. **33**, pp 1659-1668 (1994).

[6] D.A. Howe, D.W. Allan and J.A. Barnes: Proc. 35^{th} Ann. Freq. Control Symp., pp A1-A47 (1981).

[7] T. Walter: IEEE Trans. Instrum. Meas. **IM-43**, pp 69-79 (1994).

[8] R.J. Douglas and J.-S. Boulanger: Proc. 1992 IEEE Frequency Control Symp. pp 6-26 (1992).

[9] R.J. Douglas and J.-S. Boulanger: Proc. 28^{th} Precise Time and Time Interval Meeting, (1996).

A COMPARISON OF UP-CONVERTED PM AND AM NOISE IN BIPOLAR JUNCTION TRANSISTOR AMPLIFIER CONFIGURATIONS*

E.S. Ferre-Pikal and F.L. Walls

National Institute of Standards and Technology
325 Broadway, Boulder CO 80303 USA

Abstract - In this paper we report on the sensitivities of phase modulation (PM) noise and amplitude modulation (AM) noise to baseband emitter current noise and collector-base voltage noise of different bipolar junction transistor (BJT) amplifier configurations. Experimental results and predicted sensitivities are compared for a common emitter (CE) amplifier, a common base (CB) amplifier and a common collector (CC) amplifier. Results show that the sensitivities vary with transistor, amplifier configuration and circuit parameters.

Keywords: amplifier noise, AM noise, PM noise.

1. INTRODUCTION

Amplifiers add 1/f phase modulation (PM) noise and amplitude modulation (AM) noise about the amplified carrier signal. This 1/f noise is the result of low frequency (baseband) 1/f noise in the dc currents, dc voltages and amplifier impedances which are up-converted to noise about the carrier signal [1]. In this paper we present experimental results of the PM and AM noise sensitivities to baseband current and voltage noise in different bipolar junction transistor (BJT) amplifier configurations. The experimental results are compared to the sensitivities predicted from theory. Detailed description of the theory and derivations of the equations used are given in [1,2], and will not be repeated here.

The AM and PM noise equations for different BJT amplifier configurations can be derived by applying the definitions of AM and PM noise to the linear gain of the amplifier [1]. When the phase shift of the amplifier is << 1, the resulting AM noise is given by

$$\frac{1}{2}S_a(f) \cong \frac{1}{2}\left\{\left[\frac{\Delta G_o}{G_o}\right]^2 + \left[\delta^2 \Delta\delta^2\right]\right\}\frac{1}{BW} + \frac{kTFG}{2P_o}, \quad (1)$$

where G_o is the midband gain of the amplifier, δ is the phase angle, BW is the measurement bandwidth, k is Boltzmann's constant, T is the temperature in kelvins, G is the gain of the amplifier, F is the noise figure and P_o is the output power. The first term of Eq. (1) is the flicker AM noise added by the amplifier (when operating in the linear region); the second term is the thermal AM noise. The resulting PM noise is given by

$$\frac{1}{2}S_\phi(f) \cong \frac{1}{2}\Delta\delta^2 \frac{1}{BW} + \frac{kTFG}{2P_o}. \quad (2)$$

2. AM AND PM NOISE SENSITIVITIES TO CURRENT AND VOLTAGE NOISE IN CE AMPLIFIER

The AM and PM noise due to baseband current noise in a linear common emitter (CE) amplifier with a resistive load are given by

$$\frac{1}{2}S_a(f) \cong \frac{1}{4}\left(\frac{r_e}{r_e + R_E + r_g/\beta}\right)^2 \left(\frac{\Delta I_E}{I_E}\right)^2, \quad (3)$$

$$\frac{1}{2}S_\phi(f) \cong \frac{1}{4}\frac{\left(\omega C_{bc} r_g r_e R_L\right)^2}{(r_e + R_E + r_g/\beta)^4}\left(\frac{\Delta I_E}{I_E}\right)^2, \quad (4)$$

where r_e is the small signal emitter resistance, R_E is the unbypassed emitter resistance, r_g is the equivalent ac input resistance, R_L is the load resistance, $\omega = 2\pi\nu_o$, ν_o is the carrier frequency, C_{bc} is the collector-base junction capacitance, β is the current gain of the transistor, and ΔI_E represents the fluctuations in the dc emitter current.

Measurements of the AM and PM noise sensitivities to baseband current noise in this configuration were made at a carrier frequency of 5 MHz for two different transistors and for two different values of R_E. For the CE amplifier used $I_E \cong$ 25 mA, $R_L \cong 94\ \Omega$, and $r_g \cong 45\ \Omega$. To make the measurements, current noise was injected at the emitter leg of the CE amplifier. Figure 1 shows a block diagram of the measurement system used. We measured the baseband emitter current noise and the resulting AM and PM noise, and from these measurements obtained the experimental sensitivities. Table 1 shows the experimental results

* Work of U.S. Government. Not subject to copyright.

along with the predictions according to Eqs. (3) and (4). Column A in Table 1 shows the measured AM noise sensitivities to current noise ($\gamma = \Delta I_E/I_E$) when using a 2N2222A transistor and a microwave transistor for $R_E = 0$. Column C shows results from similar measurements when using $R_E = 10\ \Omega$. As expected, the AM sensitivity to current noise is independent of Fourier frequency and of the transistor used and decreases as R_E increases. The theoretical values computed according to Eq. (3) are shown in columns B and D. There is good agreement between theoretical and experimental values.

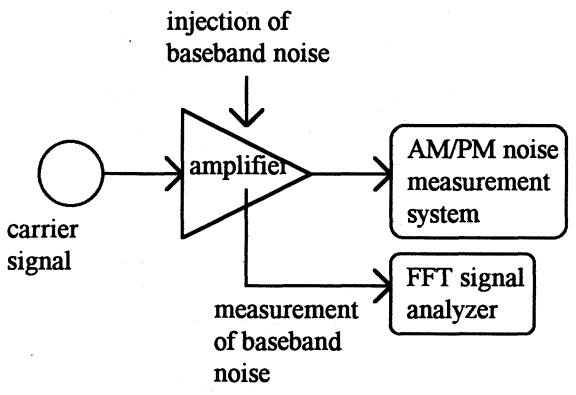


Figure 1. Block diagram of measurement system used.

Columns E and G show the PM noise sensitivities to baseband current noise for $R_E = 0$ and $R_E = 10\ \Omega$ when using a 2N2222A transistor and a microwave transistor. As expected the PM noise sensitivities are smaller when using the microwave transistor (due to the smaller C_{bc}), and decrease as R_E is increased. The theoretical predictions according to Eq. (4) are shown in columns F and H. An extra factor was included in the computation of the PM sensitivity for $R_E = 0$ since the total phase shift for this configuration was close to 1. [Equation (4) was derived assuming the total phase shift was less than 0.2 rad.] For the 2N2222A, the theoretical values are within 1-2 decibels of the experimental values.

Voltage noise was injected at the collector terminal of the amplifier in order to measure the AM and PM noise sensitivities to collector-base voltage noise (ΔV_{CB}). Table 2 shows the predicted and measured values of the sensitivities to ΔV_{CB} in a CE amplifier for $R_E = 0$ and $R_E = 10\ \Omega$ when using a 2N2222A and a microwave transistor. The theoretical sensitivities were computed using

$$\frac{1}{2}S_\phi(f) \cong \frac{1}{4}\left[\omega G_o r_g \frac{nC_{bc}}{V_{bi}+V_{CB}}\right]^2 \Delta V_{CB}^2, \qquad (5)$$

$$\frac{1}{2}S_a(f) \cong \delta^2 \frac{1}{2}S_\phi(f), \qquad (6)$$

where n is a parameter related to the doping profile of the base-collector junction ($n \approx 0.74$ for the 2N2222A and $n \approx 0.55$ for the microwave transistor), $V_{bi} \approx 0.7$ is the built-in potential of the base-collector junction, V_{CB} is the dc collector-base voltage, and G_o is the midband gain given by

$$G_o = \frac{R_L}{r_e + R_E + r_g/\beta}. \qquad (7)$$

Columns A and C show the measured PM sensitivities to collector-base voltage noise for $R_E = 0$ and $R_E = 10\ \Omega$ when using two different transistors: a 2N2222A and a microwave transistor. As expected,

Table 1. Theoretical and experimental sensitivities of AM and PM noise to baseband current noise in a linear CE amplifier at a carrier frequency of 5 MHz for $\gamma = \Delta I_E/I_E \cong 1.9 \times 10^{-5}$, $I_E \cong 25$ mA, and $V_{CB} \cong 9$ V. The output power was approximately 6 dB relative to 1 mW (dBm). The unit dBc/Hz refers to dB below the carrier in a 1 Hz bandwidth.

2N2222A	A	B	C	D	E	F	G	H
Fourier frequency	AM sensitivities (dBc/Hz rel to $\gamma = 1$)				PM sensitivities (dBc/Hz rel to $\gamma = 1$)			
	$R_E = 0$		$R_E = 10\ \Omega$		$R_E = 0$		$R_E = 10\ \Omega$	
	Measured	Predicted	Measured	Predicted	Measured	Predicted	Measured	Predicted
100 Hz	-8.5	-9.3	-22.7	-26.5	-16.3	-16.8	-45.6	-46.9
50 Hz	-8.6	-9.3	-22.6	-26.5	-16.2	-16.8	-45.2	-46.9
20 Hz	-8.6	-9.3	-22.2	-26.5	-15.6	-16.8	-45.8	-46.9
10 Hz	-8.6	-9.3	-22.0	-26.5	-16.3	-16.8	-45.4	-46.9
microwave transistor								
100 Hz	-8.5	-9.3	-22.4	-26.5	-43.2	-34.4	<-48.6	-68.8
50 Hz	-8.5	-9.3	-22.7	-26.5	-42.6	-34.4	limited by	-68.8
20 Hz	-8.5	-9.3	-22.6	-26.5	-42.1	-34.4	system	-68.8
10 Hz	-8.5	-9.3	-22.2	-26.5	-42.4	-34.4	floor	-68.8

Table 2. Theoretical and experimental sensitivities of AM and PM noise to baseband collector-base voltage noise in a linear CE amplifier at a carrier frequency of 5 MHz for $\Delta V_{CB} = 2.8 \times 10^{-4}$ $V_{rms}/\sqrt{Hz}$, $I_E \cong 25$ mA, and $V_{CB} \cong 8$ V. The output power was approximately 6 dBm.

2N2222A	A	B	C	D	E	F	G	H
	PM sensitivities (dBc/Hz rel to ΔV_{CB} = 1 $V_{rms}/\sqrt{Hz}$)				AM sensitivities (dBc/Hz rel to ΔV_{CB} = 1 $V_{rms}/\sqrt{Hz}$)			
Fourier	$R_E = 0$		$R_E = 10\ \Omega$		$R_E = 0$		$R_E = 10\ \Omega$	
frequency	Measured	Predicted	Measured	Predicted	Measured	Predicted	Measured	Predicted
100 Hz	-40.7	-40	-52.3	-52.9	-42.4	-44.1	-64	-68.6
50 Hz	-40.7	-40	-52.4	-52.9	-42.6	-44.1	-63.8	-68.6
20 Hz	-40.6	-40	-52.4	-52.9	-42.8	-44.1	limited by	-68.6
10 Hz	-40.6	-40	-52.4	-52.9	-43.1	-44.1	noise floor	-68.6
microwave transistor								
100 Hz	-73.6	-60.8	<-77	-78	<-63	-86.2	<-69	-116.8
50 Hz	-74.3	-60.8	limited by	-78	limited by	-86.2	limited by	-116.8
20 Hz	-73.2	-60.8	system	-78	system	-86.2	system	-116.8
10 Hz	limited by noise floor	-60.8	noise floor	-78	noise floor	-86.2	noise floor	-116.8

the sensitivities are smaller when the microwave transistor is used, and the sensitivities decrease with increasing R_E. The theoretical predictions are given in columns B and D. For the case of the 2N2222A when $R_E = 0$ an extra factor was added to Eq. (5) since in this case δ (total phase shift) was close to 1 [2]. For the 2N2222A the agreement between predicted and experimental values is very good, while the agreement when using the microwave transistor is not as good. Nevertheless, the sensitivity when using the microwave transistor was smaller than the sensitivity when using the 2N2222A, as expected from theory.

Experimental AM sensitivities to voltage noise are given in columns E and G of Table 2. Except for the case of the 2N2222A and $R_E = 0$, the results were limited by the noise in the measurement system. Nevertheless, a lower sensitivity was observed when using the microwave transistor (when compared to the 2N2222A for $R_E = 0$), as expected from theory.

3. AM AND PM NOISE SENSITIVITIES TO CURRENT AND VOLTAGE NOISE IN CB AMPLIFIER

We also measured AM and PM noise sensitivities to baseband current and voltage noise on a linear common base (CB) amplifier. The CB amplifier used had the following parameters: 2N2222A transistor, $I_E \cong 25$ mA, $R_E \cong 200\ \Omega$, R_S (equivalent source impedance) $\cong 22.2\ \Omega$, and $R_L \cong 387\ \Omega$.

The experimental and predicted sensitivities to current noise for carrier frequencies of 5 MHz and 10 MHz are shown in Table 3. Column A of Table 3 shows the measured AM noise sensitivities to current noise at 5 MHz carrier frequency, while column D shows the measured sensitivities at 10 MHz. As expected, the sensitivities do not change with frequency. The predicted values, computed using

$$\frac{1}{2}S_a(f) \cong \frac{1}{4}\left(\frac{r_e}{r_e + R_s}\right)^2\left(\frac{\Delta I_E}{I_E}\right)^2, \qquad (8)$$

are shown in columns B (5 MHz) and E (10 MHz). The measured and predicted values are within 3 dB. The measured PM noise sensitivities to current noise are shown in columns C (5 MHz) and F (10 MHz). These values do not change with frequency as expected from theory and are probably the result of AM to PM conversion in the mixer of the PM noise measurement system (which is usually between -15 and -25 dB) [3].

PM and AM noise due to ΔV_{CB} in a CB amplifier are given by

$$\frac{1}{2}S_\phi(f) \cong \frac{1}{4}\left(n\omega R_L \frac{C_{bc}}{(V_{bi} + V_{CB})}\right)^2 \Delta V_{CB}^2, \qquad (9)$$

$$\frac{1}{2}S_a(f) \cong \delta^2\left(\frac{1}{2}S_\phi(f)\right) \cong (\omega C_{bc} R_L)^2\left(\frac{1}{2}S_\phi(f)\right). \qquad (10)$$

The predicted and measured AM and PM noise sensitivities to voltage noise in the CB amplifier tested for carrier frequencies of 5 and 10 MHz are given in Table 4. Columns A (ν_o = 5 MHz) and E (ν_o = 10 MHz) show the measured PM sensitivities, and columns B (ν_o = 5 MHz) and F (ν_o = 10 MHz) show the predicted values. The measurements show an increase of 5 dB in the sensitivity when the carrier frequency was doubled. This is expected from theory which predicts that the PM noise (due to ΔV_{CB}) is proportional to ν_o^2, and thus it should increase 6 dB

Table 3. Sensitivities of AM and PM noise to baseband current noise in a linear CB amplifier for $I_E \cong 26$ mA and $V_{CB} \cong 7$ V. For these measurements the output power was adjusted to 9 dBm and $\gamma \cong 2.6 \times 10^{-4}$.

2N2222A	A	B	C	D	E	F
	5 MHz			10 MHz		
	AM sensitivities		PM sensitivities	AM sensitivities		PM sensitivities
Fourier	(dBc/Hz rel to $\gamma = 1$)			(dBc/Hz rel to $\gamma = 1$)		
frequency	Measured	Predicted	Measured	Measured	Predicted	Measured
100 Hz	-28.9	-32.5	-49.8	-30.0	-32.5	-50.3
50 Hz	-29.0	-32.5	-49.8	-29.8	-32.5	-50.0
20 Hz	-29.0	-32.5	-49.9	-29.6	-32.5	-49.9
10 Hz	-29.1	-32.5	-50.0	-29.7	-32.5	-49.2

Table 4. Sensitivities of AM and PM noise to ΔV_{CB} in a linear CB amplifier for $I_E \cong 26$ mA and $V_{CB} \cong 5.5$ V. For these measurements the output power was adjusted to approximately 4 dBm and $\Delta V_{CB} \cong 5 \times 10^{-4}$ $V_{rms}/\sqrt{Hz}$.

2N2222A	A	B	C	D	E	F	G	H
	5 MHz				10 MHz			
	PM sensitivities		AM sensitivities		PM sensitivities		AM sensitivities	
Fourier	(dBc/Hz rel to $\Delta V_{CB} = 1 V_{rms}/\sqrt{Hz}$)				(dBc/Hz rel to $\Delta V_{CB} = 1 V_{rms}/\sqrt{Hz}$)			
frequency	Measured	Predicted	Measured	Predicted	Measured	Predicted	Measured	Predicted
100 Hz	-50.6	-47.2	-59.6	-65.5	-45.6	-41.2	-50.0	-53.5
50 Hz	-50.5	-47.2	-59.8	-65.5	-45.6	-41.2	-50.0	-53.5
20 Hz	-50.6	-47.2	-59.8	-65.6	-45.7	-41.2	-50.0	-53.5
10 Hz	-50.5	-47.2	-59.4	-65.5	-45.8	-41.2	-50.1	-53.5

when the frequency is doubled. The predicted PM sensitivities, computed using Eq. (9) are within 4 dB of the measured values. The AM noise sensitivities to ΔV_{CB} are shown in columns C (ν_o = 5 MHz) and G (ν_o = 10 MHz). The sensitivity increased by 10 dB when the carrier frequency doubled. This agrees with theory which predicts an increase of 12 dB since $S_a(f) \propto \nu_o^4$. The predicted sensitivities (column D for ν_o = 5 MHz; column H for ν_o = 10 MHz), computed using Eq. (10) are within 4-6 dB of the measured values. At high output powers, the measured AM sensitivities showed some dependence on output power (with the sensitivity increasing as the output power increased).

4. AM AND PM NOISE SENSITIVITIES TO CURRENT AND VOLTAGE NOISE IN CC AMPLIFIER

Measurements of AM and PM noise sensitivities were also made on a linear common collector (CC) amplifier. In the CC amplifier tested the transistor used was a 2N2222A, $R_E \cong 69\ \Omega$, $r_g \cong 41\ \Omega$, $r_e \cong 1.2\ \Omega$, and $I_E \cong 23$ mA. The sensitivities of AM and PM noise to current noise are shown in Table 5. Columns A and D in Table 5 show the measured AM sensitivities to current noise at carrier frequencies of 5 MHz and 10 MHz, respectively. As expected, the measured sensitivities are independent of carrier frequency. The predicted values (columns B and E), computed using

$$\frac{1}{2}S_a(f) \cong \frac{1}{4}\left(\frac{r_e}{r_e + R_E + r_g/\beta}\right)^2 \left(\frac{\Delta I_E}{I_E}\right)^2, \qquad (11)$$

are within 5 dB of the measured sensitivities. This difference is probably due to errors in the values of r_e and R_E used for the computation (small errors in these parameters will cause significant errors in the predicted AM sensitivity). We noticed that at larger output powers [> 2 dB rel to 1 mW (dBm)] the AM sensitivity to current noise increased. This might be due the effect of the ac current i_e in the small signal emitter resistance r_e which was not taken into consideration in the calculations. The measured values of the PM noise sensitivities to current noise are shown in columns C (ν_o = 5 MHz) and F (ν_o = 10 MHz). These values do not show the expected dependence on carrier frequency ($S_\phi(f) \propto \nu_o^2$), and are probably the result of AM to PM conversion in the mixer of the PM noise measurement system.

Table 6 shows the measured and predicted sensitivities of AM and PM noise to ΔV_{CB}. The predicted values were computed using the equations

$$\frac{1}{2}S_\phi(f) \cong \frac{1}{4}\left(\frac{n\omega r_g[R_E + r_e]}{[r_e + R_E + r_g/\beta][V_{bi} + V_{CB}]}\frac{C_{bc}}{}\right)^2 \Delta V_{CB}^2, \qquad (12)$$

$$\frac{1}{2}S_a(f) \cong \delta^2\left(\frac{1}{2}S_\phi(f)\right). \qquad (13)$$

The measured PM noise sensitivities to ΔV_{CB} are shown in column A (ν_o = 5 MHz) and column E

(ν_o = 10 MHz). The predicted values, computed using Eq. (12), are shown in columns B and F. Even though there is a difference of 6 dB between the predicted and measured values, the measured sensitivities scale as ω^2, as predicted from theory. The AM sensitivities in this configuration are so small that the measurements were limited by the noise of the measurement system. Columns B and E show the measured AM sensitivities at carrier frequencies of 5 and 10 MHz respectively. The predicted values, shown in column D (ν_o = 5 MHz) and column H (ν_o = 10 MHz), are smaller than the measured sensitivities.

5. DISCUSSION AND CONCLUSION

We have presented results on measurements of AM and PM noise sensitivities to baseband current and voltage noise in three linear BJT amplifier configurations (CE, CB, and CC). As expected, the sensitivities are larger in configurations with larger gains and phase shifts (high gain CE amplifier). The measured sensitivities are generally in good agreement with the values predicted from theory. We showed experimentally that the AM sensitivity to current noise is independent of carrier frequency and independent of transistor used, while the AM sensitivity to voltage noise is a function of the carrier frequency ($S_a(f) \propto \nu_o^4$). The PM sensitivity to ΔV_{CB} increased with carrier frequency, approximately as ν_o^2. In addition, both the PM and AM sensitivities to ΔV_{CB} were smaller for transistors with smaller C_{bc}. Nevertheless the use of a transistor with small C_{bc} is not necessarily the optimum choice, since the intrinsic flicker noise of the transistor needs to be considered and it might be large for high frequency transistors. Also, the equations discussed here represent the magnitude of the up-conversion from low frequency 1/f noise to noise about the carrier. While these equations can be used to minimize the up-conversion factor, they do not provide insight on the magnitude of the low frequency 1/f noise of the amplifier. Some of the dc parameters of the amplifier and the type of transistor used play a role in the magnitude of the low frequency 1/f noise and should also be taken into consideration [2].

REFERENCES

[1] F.L. Walls, E.S. Ferre-Pikal, and S.R. Jefferts, "Origin of 1/f PM and AM Noise in Bipolar Junction Transistor Amplifiers," to be published in IEEE UFFC Trans., March 1997.

[2] E.S. Ferre-Pikal, "Studies of 1/f PM and AM Noise in BJT Amplifiers and 1/f FM Noise in Crystal Resonators," Ph.D. Thesis, Univ. of Colorado, 1996.

[3] F.L. Walls, A.D. Clements, C.M. Felton, M.A. Lombardi, and M.D. Vanek, "Extending the Range and Accuracy of Phase Noise Measurements," Proc. 42nd Ann. Freq. Control Symp., 1988, pp. 432-441.

Table 5. Sensitivities of AM and PM noise to baseband current noise in a linear CC amplifier for $I_E \cong 23$ mA and $V_{CB} \cong 9$ V. For these measurements the output power was adjusted to approximately 1 dBm and $\gamma \cong 4 \times 10^{-4}$.

2N2222A	A	B	C	D	E	F
	5 MHz			10 MHz		
	AM sensitivities		PM sensitivities	AM sensitivities		PM sensitivities
Fourier	(dBc/Hz rel to γ = 1)			(dBc/Hz rel to γ = 1)		
frequency	Measured	Predicted	Measured	Measured	Predicted	Measured
100 Hz	-36.3	-41.4	-69.9	-36.4	-41.4	-53.5
50 Hz	-36.4	-41.4	-69.7	-36.5	-41.4	-53.6
20 Hz	-36.4	-41.4	-69.5	-36.4	-41.4	-53.8
10 Hz	-36.4	-41.4	-69.6	-36.5	-41.4	-53.8

Table 6. Sensitivities of AM and PM noise to ΔV_{CB} in a linear CC amplifier for $I_E \cong 23$ mA and $V_{CB} \cong 6.8$ V. For these measurements the output power was adjusted to approximately 7 dBm and $\Delta V_{CB} \cong 7.7 \times 10^{-4}$ $V_{rms}/\sqrt{Hz}$.

2N2222A	A	B	C	D	E	F	G	H
	5 MHz				10 MHz			
	PM sensitivities		AM sensitivities		PM sensitivities		AM sensitivities	
Fourier	(dBc/Hz rel to $\Delta V_{CB} = 1 V_{rms}/\sqrt{Hz}$)				(dBc/Hz rel to $\Delta V_{CB} = 1 V_{rms}/\sqrt{Hz}$)			
frequency	Measured	Predicted	Measured	Predicted	Measured	Predicted	Measured	Predicted
100 Hz	-75	-69.9	<-90.4	-108.5	-69.3	-63.9	<-88.6	-96.5
50 Hz	-74.8	-69.9	limited by	-108.5	-69.4	-63.9	limited by	-96.5
20 Hz	-74.9	-69.9	noise floor	-108.5	-69.5	-63.9	noise floor	-96.5
10 Hz	-74.8	-69.9		-108.5	-69.5	-63.9		-96.5

HIGH TIME RESOLUTION PULSAR DATA ACQUISITION SYSTEM

I. Morison, C.A. Jordan,
University of Manchester, N.R.A.L., Jodrell Bank.

1. ABSTRACT

The paper describes a new data acquisition system for use in pulsar timing measurements made by the Lovell 76m Radio Telescope. It is capable of sampling the data in up to 256 channels and synchronously integrating the signal from each into a maximum of 1024 bins across the pulse period. With 256 channels the minimum bin width is 25.6 μsec reducing to 12.8 μsec for 128 channels. The start time of each integration and hence the pulse arrival time measured is directly related to the time frame provided by a hydrogen maser clock. The time from this clock is continuously referenced to the 1pps outputs from three GPS receivers to ultimately refer the measured pulse arrival times to TT(BIPM92). Data from the new system is presented for the pulsars B0531+21, J1911-1114 and J1643-1224.

Keywords: *pulsar timing*
time transfer

2. INTRODUCTION

Over time scales of months to years the millisecond pulsars appear to be more stable than current atomic clocks, and it is thus of great interest to make long term, high precision measurements of their pulse arrival times. This requires a system capable of subdividing the pulse period into sufficient bins to allow the pulse shape to be resolved. One major observational problem is that caused by dispersion in the interstellar medium which causes the lower frequency components to be delayed relative to the higher frequency ones. Its effect is to broaden the observed pulse width as the bandwidth is increased to improve the signal to noise ratio so reducing the accuracy of the measurement. It is possible to digitise the full bandwidth and then carry out a coherent dedispersion process computationally but at present, even using powerful vector processing systems, the time to analyse the data exceeds the observing time by around an order of magnitude. An alternative approach which allows continuous observations is to split the overall bandwidth of the observing system into a large number of narrow sub-bands in which the dispersion is negligible. The signal in each must then be individually sampled and integrated. Following a suitable integration period the signals relating to each sub-band are time shifted in the computer to allow for the dispersion across the band and then averaged together.

3. SYSTEM REALISATION

A block diagram of the system is shown in Figure 1. It is designed for use with a multi-channel filter bank in which the output of each narrow band filter has been square law detected and low pass filtered to give a smoothed output appropriate to the sampling rate of the data acquisition system. These outputs are then sequentially selected and digitised and the resulting samples synchronously integrated into a number of bins covering the pulse period. The system can accept up to 256 sub band inputs and integrate the noise within each into a maximum of 1024 bins.

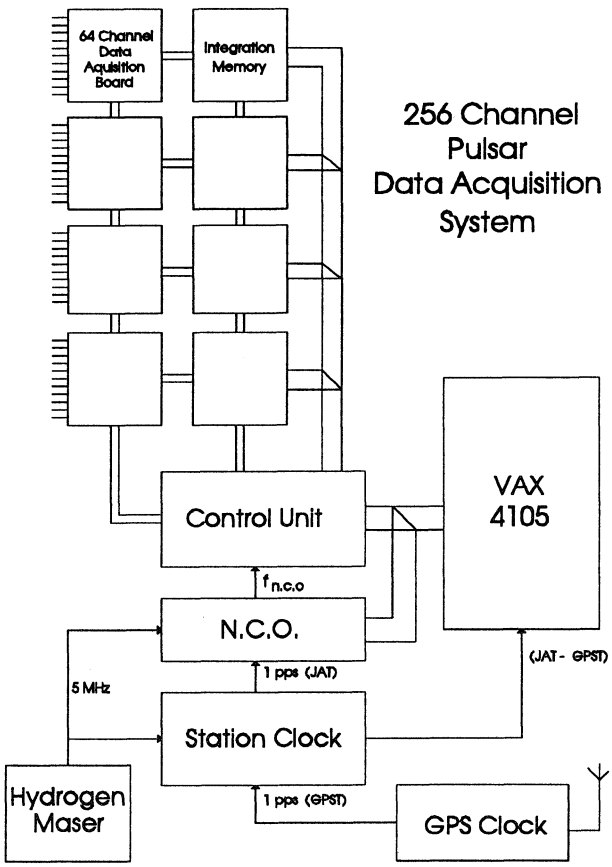


Figure 1: System Block Diagram

A high degree of parallelism is used to achieve the desired time resolution. There are 4 independent banks of 24 bit integrating memories. They provide 1024 bin locations per channel and integrate each sample into the

appropriate channel and bin location in 400nsec. The allowed integration time is dependent on the number of channels and bins in use and is typically of order 10 minutes. Associated with each integration memory bank is a data acquisition board, shown in Figure 2, which samples the data from the 64 channels serviced by it. A single analogue multiplexer and analogue to digital converter would be unable to provide sampled data at an appropriate rate so the data acquisition board uses four 16 channel multiplexers each with its own buffer amplifier and 6 bit 'flash' analogue to digital converter. Thus, on each of the four data acquisition boards, 4 channels are being sampled and digitised

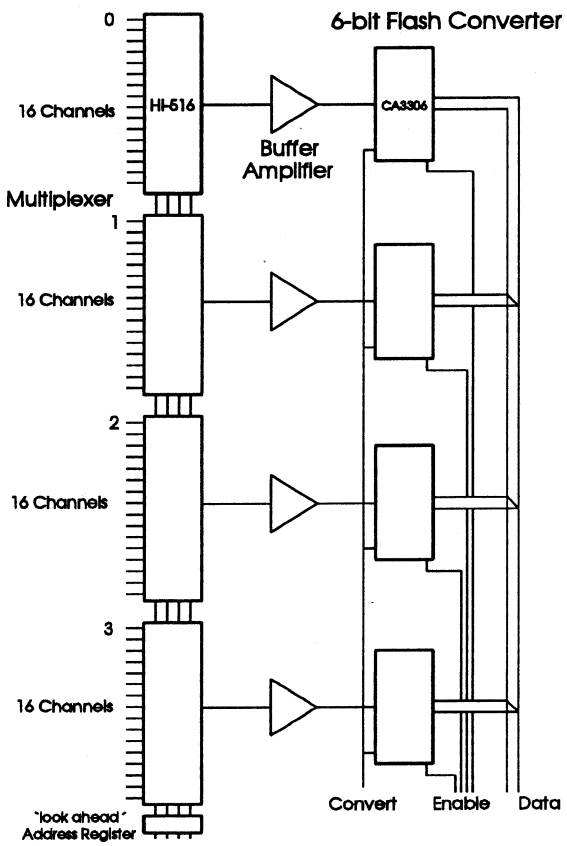


Figure 2: Data Acquisition Board

simultaneously. The system is 'pipelined' so that the acquisition circuitry is operating 4 channels ahead of the associated digital hardware. The four digitised samples are then read out sequentially from each data acquisition board and added into the appropriate locations within the associated block of integrating memory.

4. CONTROL UNIT

This is linked to the VAX control computer through a 16 bit bi-directional parallel interface. It contains the counters which advance the multiplexers and integration memories through the channel and bin locations. Circuitry is included to allow the multiplexers to 'look ahead' as described above. It is important to sample the output of the filter bank detectors at a rate commensurate with their output smoothing time constant otherwise the optimum signal to noise ratio will not be achieved. Thus for all except the very fastest pulsars the output of the detectors must be sampled many times during each bin period. The control circuitry thus allows for the system to 'dwell' on each bin, integrating a specified number of samples before moving onto the next. Each 400 nsec sample and integration period is initiated on the rising edge of the control clock which is generated by a 48 bit Numerically Controlled Oscillator as described below. At the end of each integration, the integrated data is transferred into the control computer; an operation which automatically clears each memory location ready for the next integration.

5. THE NUMERICALLY CONTROLLED OSCILLATOR

The system timing is controlled by the use of a Stanford Telecommunications 48 bit Numerically Controlled Oscillator (NCO). This operates from a 40 MHz clock derived from the 5 MHz output of the station Hydrogen Maser frequency standard. The NCO maintains a record of the phase of the output waveform accurate to 48 bits. At each clock cycle the number which has been downloaded into a 48 bit Δ-Phase register is added to the previous phase. The most significant 13 bits of the phase register control the address of a Sin/Cos lookup table. In this application only the top bit of this table is used to provide a square wave of the desired frequency whose positive edges initiate the integration of each set of four samples into memory. The period of its output frequency is set to be an integral sub multiple of the apparent pulsar period dependent on the the number of channels (NCHAN) and bins (NBIN) and the chosen number, NDWELL, of samples per bin. NDWELL is a free parameter which is chosen to run the system as close to its maximum operating speed as possible and is thus the NCO is set so that its output frequency has a period less than, but close to 400 nsec. If the data collected in successive real-time integrations is to be further integrated, then the pulse must remain in identical bins. This requires that each integration is begun at the same phase of the pulsar rotation with a precision significantly higher than that of the minimum bin period. The NCO control circuitry uses the 1pps pulse from the station clock to control the integration start times. When a RUN command is sent to initiate an integration, a counter system is used to enable the NCO output at a precise time following the next 1 pps pulse. A precision of 200 nsec is achieved by the use of a 24 bit counter clocked at 5 MHz. The NCO frequency is then updated each second precisely on each 1 pps pulse to allow for changes in the apparent pulse period during the integration.

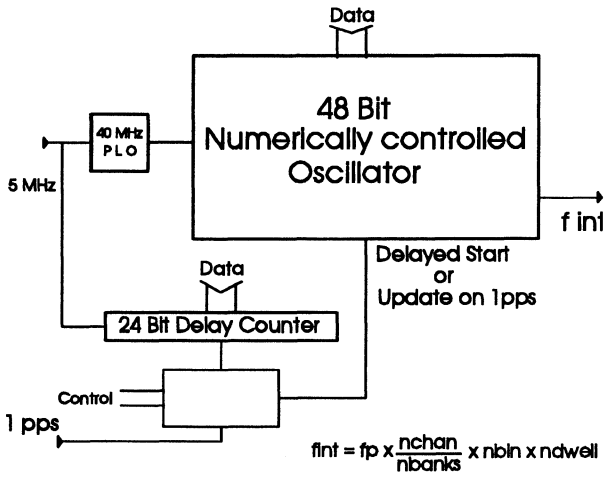


Figure 3: The Numerically Controlled Oscillator

6. THE STATION CLOCK

The 1pps time base used to control the timing of the system is provided by the Station Clock and is normally derived from the 5 MHz output of a Sigma τ Hydrogen Maser, but there is automatic fall back within 100 nsec onto a Rubidium frequency standard should its output fail. This 1 pps waveform defines Jodrell Atomic Time (JAT). The clock continuously monitors the time offset between JAT and the 1 pps waveforms from three GPS clocks which define GPS time (GPST). Two are TRAK SYSTEMS Model 8820 GPS Station Clocks with the third being a Totally Accurate Clock, TAC. All three are operated in the fixed position mode having first used them to find the precise locations of their antennas by long term averaging. This significantly improves the rms jitter on their output 1pps waveforms. Rather than have the difference between JAT and GPST vary both negatively and positively around zero, the Jodrell Clock is initially set so that JAT is 2 μsec fast of GPST. Following accurate adjustment of the maser output the current drift rate between JAT and GPST is very slow. The offset is reset when the difference has changed by 1 μsec. The Station clock continuously monitors the offset between JAT and the 1pps waveforms from the three GPS clocks GPST with a precision of 50 nsec. These difference are transmitted each minute to the VAX 4105 computer system. The header of each pulsar data set includes this information thus enabling the pulse arrival times to be related to GPST. Due to jitter in the GPS 1pps pulses, in part brought about by the use of selective availability, together with the 50 nsec measurement accuracy, the precision of the direct comparison is of order 100 nsec. Measurements of the RMS of the instantaneous measurements for each clock give a value between 45 and 50 nsec. The Jodrell station clock averages these instantaneous measurements over a rolling 3 hour period for each GPS clock to provide measurements of the time offset with a precision close to 10 nsec RMS. This is better than the nominal 50 nsec accuracy of each measurement as the selective availability jitter is effectively providing dither, so by oversampling, a higher precision comparison may be achieved. If it is desired to work to this precision then the published corrections between GPS Time and TT may be incorporated. A graph is presented showing the drift of the Sigma τ Maser with respect to the GPS 1pps waveforms from TRAK 8820 and TAC receivers over a period of 20 days showing both the raw and smoothed data. The raw data presents 1 hr samples of the direct comparison and shows the 50 nsec quantisation of the measurements. The smoothed data are the hourly values of the 3 hr rolling average. The vertical divisions are 100 nsec.

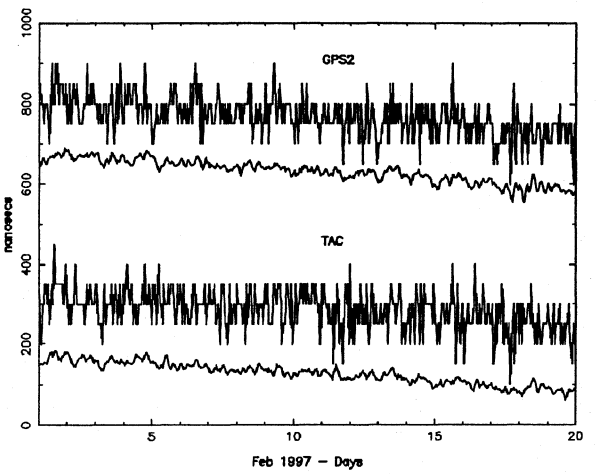


Figure 4: TRAK and TAC clocks referenced against the Sigma τ Hydrogen Maser

7. DATA

All observations presented were made using the 76m Lovell radio-telescope at Jodrell Bank with a bandwidth of 4 MHz centered on 606 MHz in conjunction with a filter-bank which splits the the band into 64 sub-bands each of 128 KHz. Figures 5 and 6 show an observation of the CRAB pulsar, B0531+21 in both hands of circular polarization. 512 bins were specified to cover the pulse period. The first graph plots the integrated data from each of the sub-bands showing the effect of dispersion on the pulse arrival times. The second shows the result of adding the data from each sub-band after having corrected for dispersion. The plot clearly shows the precursor to the main pulse. Figure 7 shows the de-dispersed data for the two hands of polarisation of the 3.626 millisecond pulsar J1911-1114 for which the pulse period was split into 128 bins. Figure 8 shows the de-dispersed data for the 4.621 millisecond pulsar J1643-1224 using the earlier Jodrell Bank Pulsar Synchronous integrator where only 16 bins could

be put across the pulse period. The design of the new PSI which allows 128 bins to be placed across the pulse period, as shown in Figure 9, has thus improved the time resolution by a factor of 16. This is, of course, largely a result of the 16 way parallelism employed in the new device. In further analysis the pulse profile for each pulsar is cross correlated with a template of the average pulse profile to find the bin corresponding to the nominal pulse arrival time. From a knowledge of the start time of the first bin, specified to 200 nsec, and the width per bin the pulse arrival times at the Lovell telescope can then be derived. From this the arrival times at the solar system barycentre are obtained using the JPL DE200 ephemeris. The data collected over a period of time then forms the basis for further analysis such as the measurement of the pulsar spin down rates.

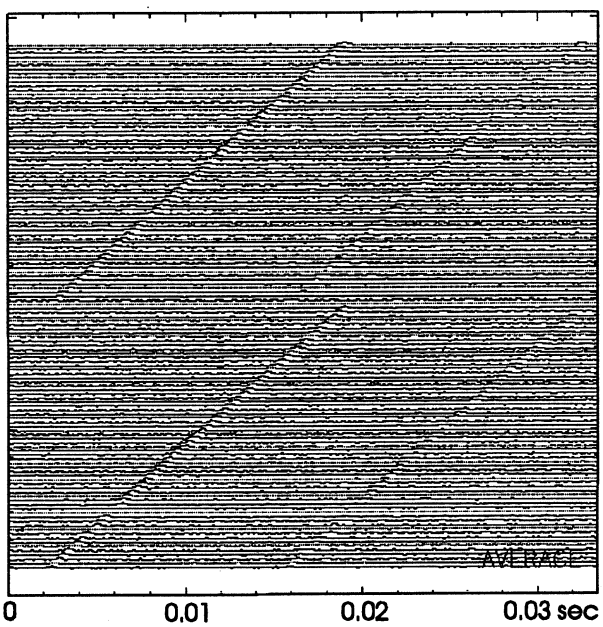


Figure 5: Individual Channel Data for B0531+21

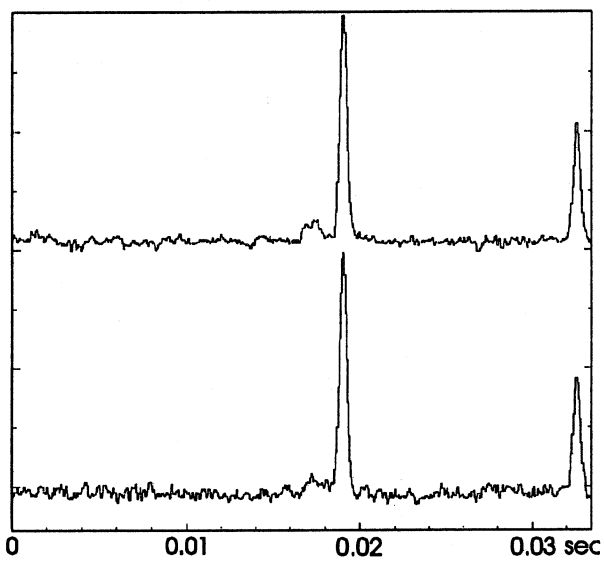


Figure 6: De-dispersed Data for B0531+21

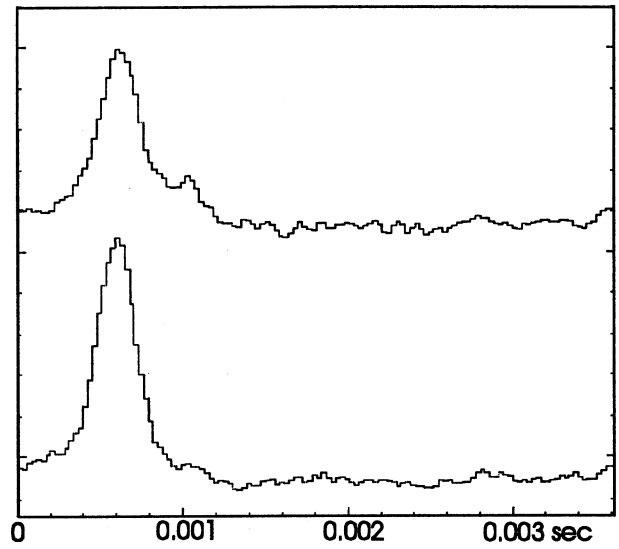


Figure 7 : De-dispersed data for J1911-1114

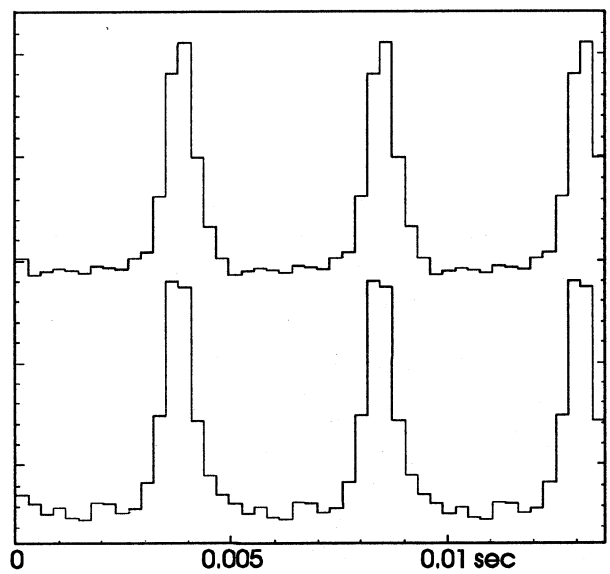


Figure 8: De-dispersed Data for J1643-1224 using 'old' Synchronous Integrator

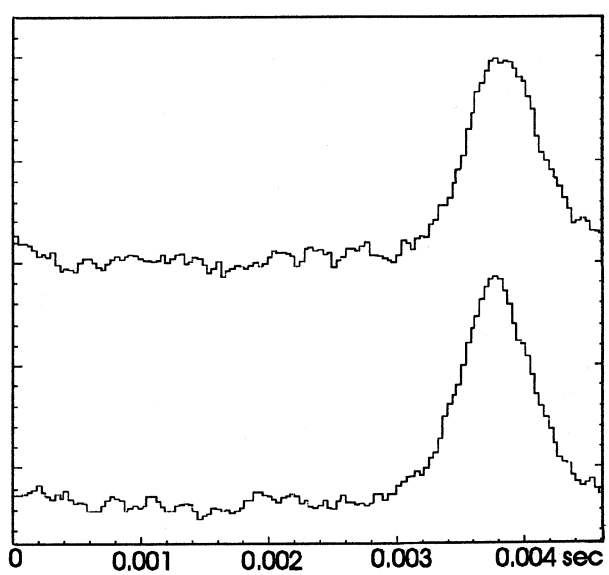


Figure 9: De-dispersed Data for J1643-1224

NUMERICAL SIMULATION OF A PHASE MODULATION NOISE MEASUREMENT SYSTEM OF QUARTZ CRYSTAL RESONATORS

F. STHAL, M. MOUREY

Laboratoire de Chronométrie, Electronique et Piézoélectricité,

Ecole Nationale Supérieure de Mécanique et des Microtechniques

26, Chemin de l'Epitaphe - 25030 BESANÇON CEDEX - FRANCE

1. ABSTRACT

A new numerical model of a phase modulation (PM) noise measurement system of high stability quartz crystal resonator is presented. The measurement device allows to measure the inherent phase stability of quartz crystal resonators in a passive circuit without the noise usually associated with an active oscillator. It is based on the phase bridge method using a resonator pair driven by a low noise source.

The output signal, obtained with a phase detector, is proportional to the phase difference introduced by resonators. The numerical transfer function of each bridge path is given by the model. The output spectral density of phase fluctuations is calculated with these transfer functions and a low-noise source numerical approximation.

The model was applied to 10 MHz BVA quartz resonator. It enables to quantify the source noise rejection versus resonance frequency and unloaded quality factor of quartz crystal resonators.

Keywords: Allan variance, phase noise, phase bridge method, 10 MHz BVA quartz crystal resonator, numerical simulation, power spectral density of phase fluctuations.

2. INTRODUCTION

Quartz crystal controlled oscillators are very important in frequency metrology. They are used in nearly all precision frequency generation and measurement device.

With an oscillating loop it is difficult to distinguish in the output signal the respective contribution of the electronic amplifier and of the quartz crystal resonator to phase fluctuations. So, new techniques are developed to measure the inherent phase noise of quartz resonators.

In the middle of the 1970's, F.L. Walls and al (Refs. 1-2), then J.J. Gagnepain (Ref. 3) have presented new methods, which allow to measure the inherent frequency fluctuations of quartz crystal resonator in a passive circuit without the noise usually associated with an active oscillator. Two crystals which are as identical as possible are driven in π transmission network by a low noise source.

By careful adjustment of crystal tuning and the balancing of the relative Q's, the output signal can be made first order insensitive to both residual amplitude and frequency modulation in the source. These techniques are appropriated to measurements at room temperature and at very low temperature and were used to study a large number of crystals.

To obtain identical resonators is hard and then the noise of the source can't be efficiently rejected. In order to extend the study, a numerical model of a measurement system which is close to the F. L. Walls device has been developed. This simulation shows the measurement system limits and evaluates the acceptable resonators differences so that it is possible to measure their phase noise.

We give a complete transfer function of the measurement system. Different input sources with more or less phase noise can be used. The model was applied to 10 MHz BVA resonators. The source noise rejection is quantified versus resonance frequency and versus unloaded quality factor of quartz crystal resonators. The results are presented in frequency or time domain using the spectral density of phase fluctuations $S_\phi(f)$ or the Allan variance $\sigma^2_y(\tau)$.

3. MEASUREMENT DEVICE

The simulated measurement device is presented on figure 1.

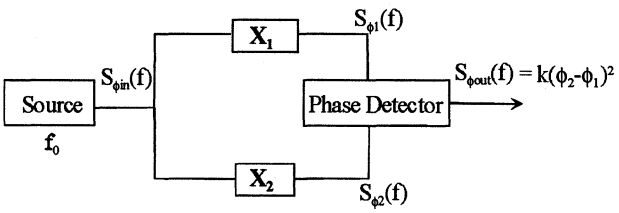


Fig. 1 : Passive measurement device of phase noise of quartz crystal.

A low-noise source of frequency f_0 drives two blocks X_1 and X_2. The blocks X_i have each one a quartz crystal placed in an impedance adaptation network. The inherent quality factor of the resonators are Q_i, the resonance frequency f_{qi}. The resonance frequency of the block can be tuned with a tuning capacitor.

A power spectral density of phase fluctuations $S_{\phi_{out}}(f)$ is obtained from a spectrum analyzer located after the phase detector. $S_{\phi_{out}}(f)$ is proportional to $(\phi_2-\phi_1)^2$ where ϕ_1 and ϕ_2 are the phase introduced respectively by each path X_1 and X_2. By careful adjustment of crystal tuning and the balancing of the relative Q's, the output from the phase detector can be made first order insensitive to the residual amplitude modulation of the source.

4. PROCESS OF MODELING

Model equations of measurement system are presented in the following part.

4.1 Theoretical source equation

The phase noise of the source is given by power spectral density of phase fluctuations $S_{\phi in}(f)$. This noise is computed with an asymptotic curve (Ref. 4).

$$S_{\phi in}(f) = \frac{A}{f^3} + \frac{B}{f} + C$$

Figure 2 shows an example of a low-noise source spectral density $S_{\phi in}(f)$ which is used in the computation.

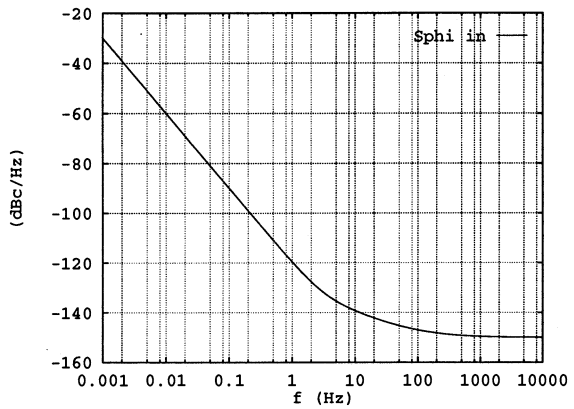


Fig. 2 : Power spectral density of phase fluctuations of the source $S_{\phi in}(f)$ in dBc/Hz.

The function is given versus relative frequency or Fourier frequency, it is normalized with source frequency f_0. To take into account the spectrum in each side of f_0, calculus used single sideband power spectral densities $\mathcal{L}^{+}_{\phi in}(f)$ et $\mathcal{L}^{-}_{\phi in}(f)$. Figure 3 gives an example of these single sideband power spectral densities.

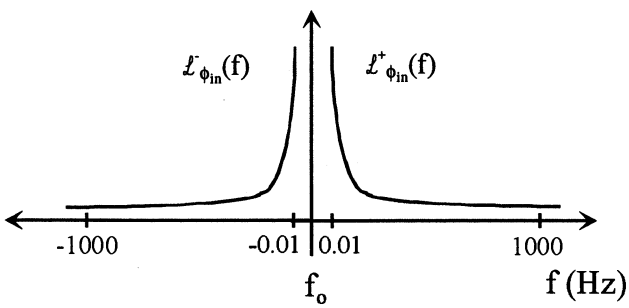


Fig. 3 : Single sideband spectral densities of the source.

4.2 Equivalent circuit of a quartz crystal

Figure 4 presents the usual equivalent circuit of a quartz crystal used in this model.

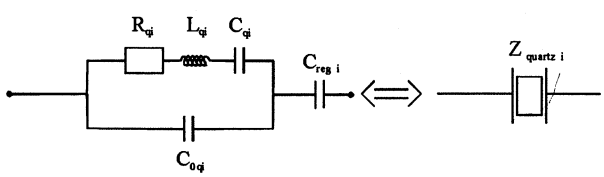


Fig. 4 : Resonator equivalent circuit with a serial tuning capacitor.

R_{qi}, L_{qi}, C_{qi}, are motional parameters. The quartz crystal frequency is given by this parameters. Electrodes of the resonator are represented by the static capacitor C_{0qi}. Serial tuning capacitor C_{regi} allows an adjustment of $Z_{quartzi}$ to f_i frequency. Frequency adjustment is necessary to have quartz crystal frequencies at the same value than the source frequency.

The impedance of the equivalent circuit shown in the figure 4 has the following expression:

$$Z_{quartzi} = \frac{1}{\dfrac{1}{R_{qi} + jL_{qi}\omega + \dfrac{1}{jC_{qi}\omega}} + j\omega C_{0qi}} + \frac{1}{j\omega C_{regi}}$$

4.3 Impedance adaptation

To preserve its resonance quality, quartz crystal is placed in an impedance adaptation network which is composed of R_{1i} and R_{2i} resistors. Resistors are evaluated versus quartz crystal impedance. Figure 5 shows the adaptation network.

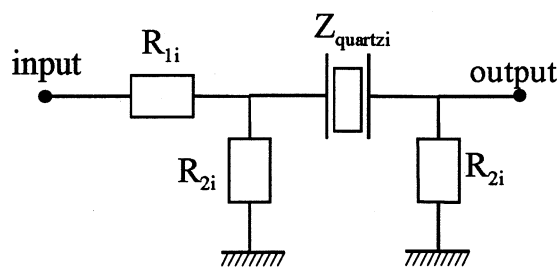


Fig. 5 : Impedance adaptation of quartz resonator.

Transfer function of the figure 5 circuit is:

$$H_i = \frac{R_{2i} Z_i}{(Z_{quartzi} + R_{2i})(R_{1i} + Z_i)}$$

with

$$Z_i = \frac{1}{\dfrac{1}{R_{2i}} + \dfrac{1}{Z_{quartzi} + R_{2i}}}$$

Elements of X_i blocks measurement system are shown in figure 5. Global X_i blocks transfer functions are given for each side of f_0 in relative frequency. The two single sideband transfer functions are expressed in dBc/Hz and normalized versus the transfer function of quartz crystal resonator n°1.

$$X_i^{+}(f) = 20 \cdot \mathrm{Log}\left|\frac{H_i(f + f_0)}{H_1(f_0)}\right|$$

$$X_i^{-}(f) = 20 \cdot \mathrm{Log}\left|\frac{H_i(-f + f_0)}{H_1(f_0)}\right| \quad \text{with } i = 1, 2$$

4.4 Spectral densities expressions

Power spectral density of phase fluctuations of the source $S_{\phi_{in}}(f)$ applied in each path of the device gives the following expressions.

$$Y^+_i(f) = X^+_i(f) + L^+_i(f) \text{ et } Y^-_i(f) = X^-_i(f) + L^-_i(f) \qquad i = 1, 2$$

In output of each path, spectral density $S_{\phi_i}(f)$ is computed from the two previous single sideband power spectral densities.
In dBc/Hz:

$$S_{\Phi i}(f) = 10 \cdot \mathrm{Log}\left[\left| 10^{\frac{Y_i^-(f)}{10}} + 10^{\frac{Y_i^+(f)}{10}} \right| \right] \qquad i = 1, 2$$

Finally, the output spectral density of phase fluctuations $S_{\phi_{out}}(f)$ after the phase detector is given by the phase difference between spectral densities of each path $S_{\phi_1}(f)$ and $S_{\phi_2}(f)$. This phase difference is a linear operation and results are expressed in dBc/Hz.

$$S_{\Phi out}(f) = 10 \cdot \mathrm{Log}\left[\left| 10^{\frac{S_{\Phi 1}(f)}{10}} - 10^{\frac{S_{\Phi 2}(f)}{10}} \right| \right]$$

4.5 Time domain : Allan variance

Short-term stability is usually given in time domain with an Allan variance. Allan variance depends on the relative frequency and the power spectral density of phase fluctuations $S_{\phi_{out}}(f)$ as follow.

$$\sigma_y^2(\tau) = \int_0^\infty S_{\phi out}(f) \frac{2\sin^4 \pi f \tau}{(\pi f_0 \tau)^2} df$$

5. NUMERICAL RESULTS

The model was applied to 10 MHz BVA resonators, SC-cut, third overtone (Ref. 5).

For both resonators, motional resistor R_{qi} is set at 100 Ω. Static capacitor C_{0qi} is 2.4 pF. Q_i-factors, inherent resonator frequencies f_{qi} and f_i of $Z_{quartzi}$ are variable parameters. Motional inductor L_{qi} of quartz crystal is computed with the following approximate equation:

$$L_{qi} = \frac{Q_i \cdot R_{qi}}{2 \cdot \pi \cdot f_{qi}} \qquad i = 1, 2$$

Motional capacitor C_{qi} is computed by a dichotomy method with a 10^{-15} accuracy to obtain the quartz crystal frequency f_{qi}.

Tuning capacitor C_{regi} of the $Z_{quartzi}$ block is also computed to obtain the global resonance frequency f_i. It is drawn versus inherent quartz crystal resonance frequency f_{qi} on the figure 6. The source frequency and the resonance frequency of $Z_{quartzi}$ block are equal to 10 MHz. Q_i-factor is $1.36 \cdot 10^6$.

The 10 MHz BVA resonators used in time base, are made to have their resonance frequency equal to 9999983 ± 10 Hz. The range of the tuning capacitor is fixed by the extreme frequency values.

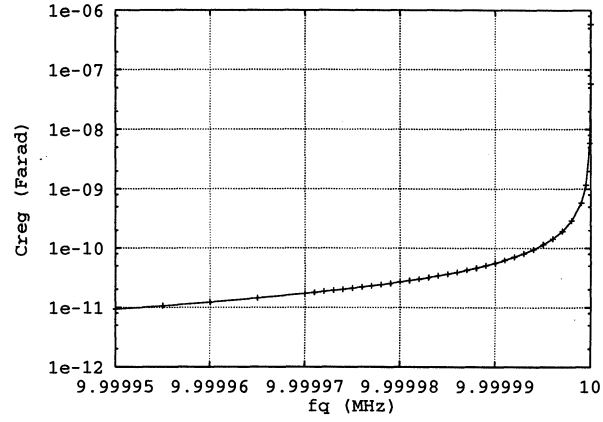


Fig. 6 : Tuning capacitor versus quartz crystal resonance frequency f_{qi}

Maximum and minimum values of the tuning capacitor C_{regi} which bring quartz crystal frequency f_{qi} back to $Z_{quartzi}$ frequency f_i are:

if f_{qmax} = 9999993 Hz, C_{regi} = 81 pF

if f_{qmin} = 9999973 Hz, C_{regi} = 19 pF

The simulation of different quartz crystal is possible with this method. Several cases are studied.

5.1 Identical Q_i and Q_2, $f_1 = f_2 = f_0$

Q-factors Q_1 and Q_2 are equal to $1.36 \cdot 10^6$. Tuning capacitors allow $f_1 = f_2 = f_0$. Figure 7 shows $S_{\phi_{out}}(f)$.

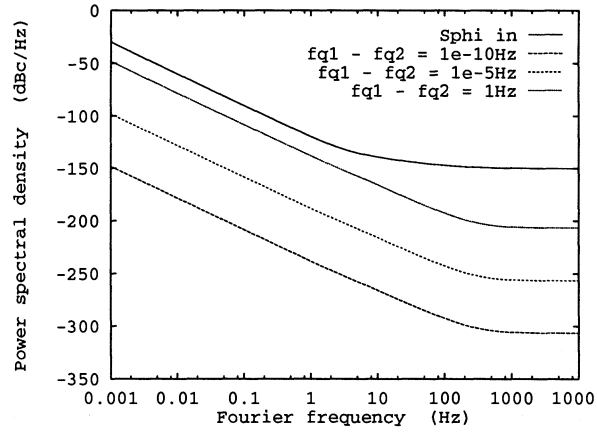


Fig. 7 :Power spectral density $S_{\phi_{out}}(f)$ for different frequency gap of quartz crystal resonators.

The f_{q1} frequency has been chosen very close to 10 MHz to obtain a tuning capacitor equal to 1 Farad. Under these conditions, the resonator n° 1 is almost without frequency correction.
We can note that the source noise is less rejected when the frequency gap of quartz crystals is great. The output's spectral density goes back to the spectral density of the source. The curves show that the resonators must be as identical as possible to obtain a good rejection of the source noise.

Figure 8 shows the case where f_{q1} and f_{q2} are to the extreme values of the bandwidth:

$f_{q1} = f_{qmax} = 9999993$ Hz.
$f_{q2} = f_{qmin} = 9999973$ Hz.

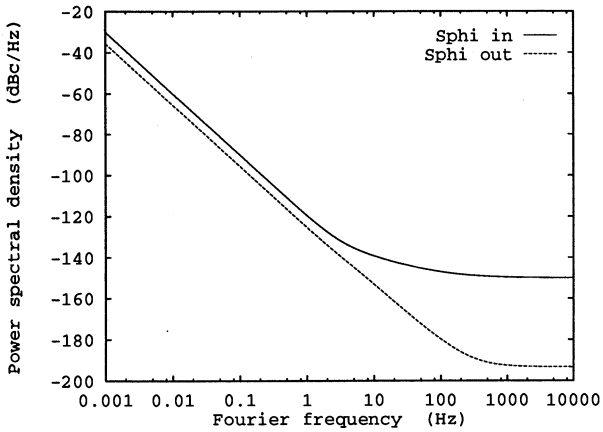


Fig. 8 : Power spectral density $S_{\phi_{out}}(f)$ for $C_{reg1} = 81$ pF et $C_{reg2} = 19$ pF.

Figure 9 shows the output power spectral density of phase fluctuations of the measurement system at 1 Hz versus the resonance frequency gap of quartz crystals. The value of $S_{\phi_{in}}(f)$ is -120 dBc/Hz. The points on the figure are drawn for two cases: the frequency f_{q1} is chosen such as $C_{reg1} = 1$ Farad, then $f_{q1} = 9999983$ Hz, accordingly the tuning capacitor is 32 pF.

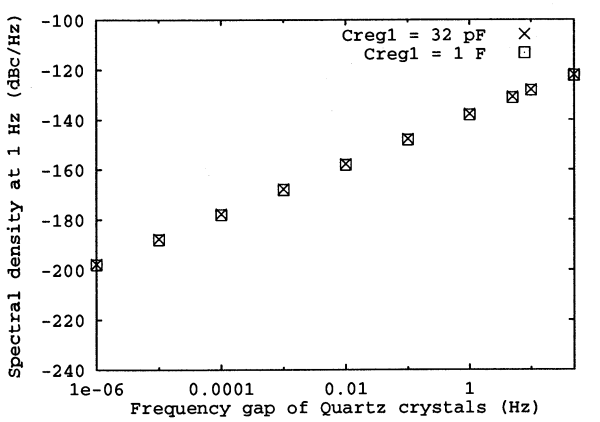


Fig. 9 : Power spectral density $S_{\phi_{out}}(f)$ versus $\Delta f = f_{q1} - f_{q2}$.

The resolution at 1 Hz of the measurement system is given by these points. In the frequency bandwidth, the frequency gap of quartz crystals is therefore an important parameter seeing that the points of each case are superposed. The values of f_{q1} and f_{q2} have a negligible effect on the resolution.

5.2 Identical Q_1 et Q_2 ($1.36 \cdot 10^6$), frequency $f_1 = f_2 \neq f_0$

Figure 10 shows the variations of $S_{\phi_{out}}(f)$ versus f_1 and f_2. Resonance frequencies of quartz crystals are chosen equal to f_{qmax} and f_{qmin} in view to obtain the worse source noise rejection. The values f_1 and f_2 are necessarily greater than f_{qmax} because the tuning capacitor can't decrease the resonance frequency of quartz crystal.
The curves have resonance peaks due to the frequency offset between $f_1=f_2$ and the source frequency f_0. The peaks appear for frequency $f_1=f_2$ below and above the source frequency because the calculus uses each side of the source spectrum.

All curves are 5 dBc/Hz below $S_{\phi_{in}}(f)$ before 1 Hz. The curve of $S_{\phi_{out}}(f)$ where $f_1=f_2=f_0 + 20$ Hz differs more from the curve $f_1=f_2=f_0$. It is about 15 dBc/Hz above the curve $f_1=f_2=f_0$ after 10 Hertz.

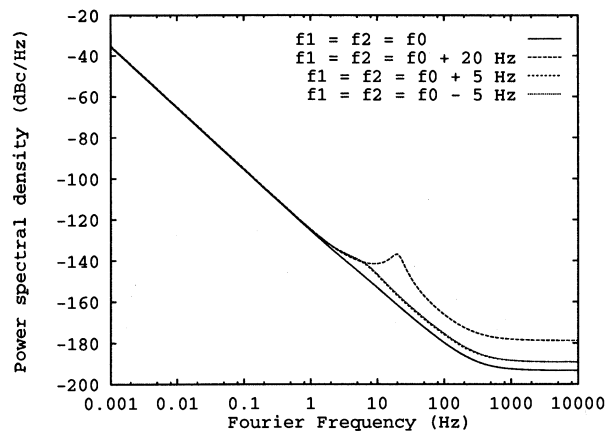


Fig. 10 : $S_{\phi_{out}}(f)$ versus Fourier frequency f for $f_1 = f_2 \neq f_0$.

5.3 Identical Q_1 et Q_2 ($1.36 \cdot 10^6$), $f_1 \neq f_2 \neq f_0$

Figure 11 presents an example of the output spectral density for $f_1 = f_0 - 6$ Hz, and $f_2 = f_0 + 10$ Hz. The quartz crystal frequencies are $f_{q1}=f_{qmax}$ and $f_{q2}=f_{qmin}$, The output spectral density obtained constitutes an unfavorable rejection of the source noise. The values of the output spectral density of phase fluctuations are even above the source noise around 5 Hz.

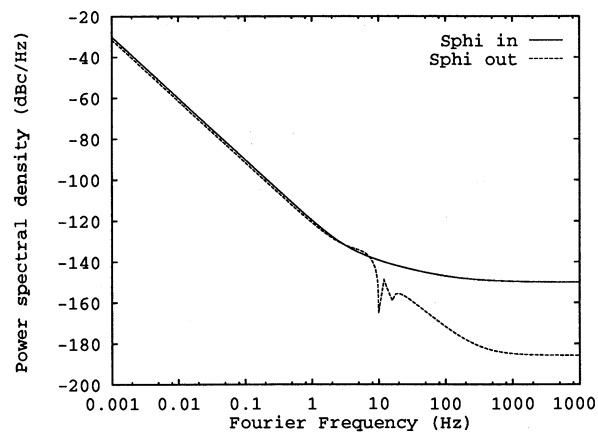


Fig. 11 : $S_{\phi_{out}}(f)$ for $f_1 \neq f_2 \neq f_0$

5.4 Variation of Q-factors of quartz crystals

Figure 12 shows the output spectral densities for identical Q-factors in each path. Curves are obtained for $f_{q1}=f_{qmax}$, $f_{q2}=f_{qmin}$, $f_1=f_2=f_0$. The curves which are obtained for $Q_1=Q_2=0.5 \cdot 10^6$ and $Q_1=Q_2=1.4 \cdot 10^6$ are separated by around 5 dBc/Hz. We can note that when the Q_i-factors increase, the $S_{\phi_{out}}(f)$ is higher, except in the range from 5 Hz to 500 Hz.
We have taken Q_1 equal to $1.4 \cdot 10^6$, then we have observed the variation of the output spectral density of phase fluctuations versus Q_2. The quartz crystal frequencies remain equal to f_{qmax} and f_{qmin}.

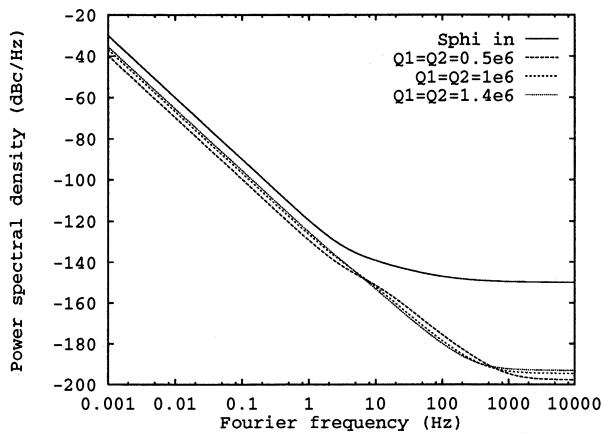


Fig 12 : Output power spectral density of phase fluctuations versus inherent quartz crystal Q-factors.

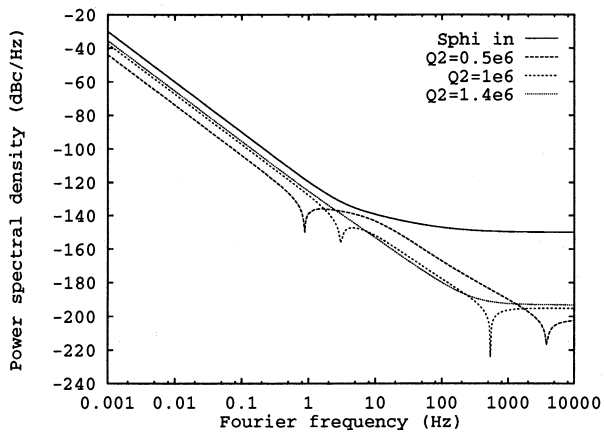


Fig. 13 : Output power spectral density of phase fluctuations versus Q_2.

Figure 13 shows the obtained results. In the frequency range from 10 Hz to 100 Hz, the source phase noise is less rejected when the Q-factors are the more different.
In this area, the values of $S_{\phi_{out}}(f)$ for the corresponding curve of $Q_2 = 0.5 \cdot 10^6$ is about 10 dBc/Hz more than $S_{\phi_{out}}(f)$ with $Q_2 = 1.4 \cdot 10^6$.

5.5 Time domain calculus example

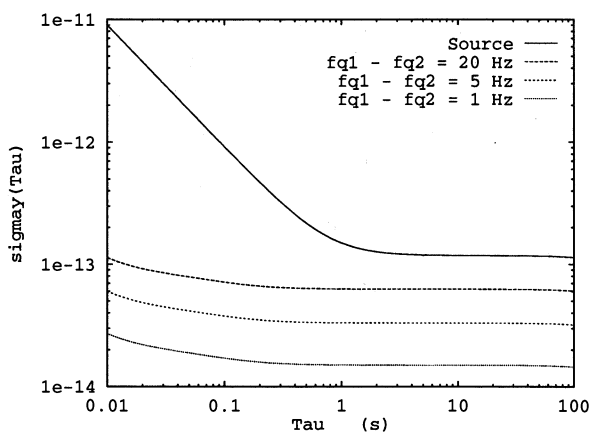


Fig. 14 : $\sigma_y(\tau)$.

Figure 14 shows the standard deviation $\sigma_y(\tau)$ of the source and for the measurement system output for three cases:
- $Q_1 = Q_2 = 1.36 \cdot 10^6$, $f_{q1} = f_{qmax}$ and $f_{q2} = f_{qmin}$.
- $Q_1 = Q_2 = 1.36 \cdot 10^6$, f_{q1}=9999983 Hz and f_{q2}=99999982 Hz.
- $Q_1 = Q_2 = 1.36 \cdot 10^6$, f_{q1}=9999983 Hz and f_{q2}= 9999978 Hz.

$\sigma_y(\tau)$ is under $2 \cdot 10^{-14}$ if the frequency gap of quartz crystals is less than 1 Hz

Spectral densities are evaluated for Fourier frequency ranging from 0.001 to 10000.
Integral calculus of the $\sigma_y(\tau)$ is obtained with integral limits equal to 0.001 and 10000. τ ranges from 0.01 to 100. This restriction on the τ value allows a numerical simulation with a good precision.

6. CONCLUSION

The proposed model allows us to evaluate the output phase noise. The calculation assumes that in order to have a better rejection of source phase noise the quartz crystal resonators must be as identical as possible. The model will be improved if we take into account the noise generated by the phase detector in the computation. In this way, the result will be very closed to the practical case. Nevertheless, the phase detector noise could be eliminated by cross-correlation method (Ref 6).
This paper shows that the measurement system phase noise can be quantified according to the quartz crystal parameters. Thus, we can predict the limits of quartz crystal phase noise measurement.

7. REFERENCES

[1] A.E WAINWRIGHT, F.L. WALLS, W.D. McCAA, "Direct measurement of the inherent frequency stability of quartz crystal resonators", 28th Annual Freq. Control Symp., pp. 177-180, (1974).

[2] F.L. WALLS and A.E. WAINWRIGHT, "Measurement of the short-term stability of quartz crystal resonators and the implications for crystal oscillator design and applications", IEEE transactions on instrumentation and measurement, Vol. IM-24, N 1, Mars, pp. 15-20, (1975).

[3] J.J. GAGNEPAIN, "Fundamental noise studies of quartz crystal resonators", 30th Annual Freq. Control Symp., pp. 84-91, (1976).

[4] F.L. WALLS, P. H. HANDEL, R. BESSON and J.-J GAGNEPAIN, "A new model of 1/F noise in baw quartz resonators", IEEE FREQUENCY CONTROL SYMPOSIUM, pp. 327-333, (1992).

[5] R.J. BESSON, J.J BOY, M. MOUREY, "BVA resonators and oscillators : a review. Relation with space requirements and quartz material characterization", IEEE FREQUENCY CONTROL SYMPOSIUM, pp. 590-599, (1995).

[6] W.F. WALLS, "Cross-Correlation Phase Noise Measurements", IEEE FREQUENCY CONTROL SYMPOSIUM, pp. 275-261, (1992).

PERFORMANCE ASSESSMENT OF A DELAY COMPENSATION PHASE NOISE AND TIME JITTER REDUCTION METHOD

M J Underhill*, M J Blewett*
* School of Electronic Engineering, Information Technology and Mathematics
University of Surrey
Guildford, Surrey GU2 5XH, UK

ABSTRACT

A new phase noise and time jitter reduction circuit was announced at the 1996 EFTF (Ref.1). This paper derives a theoretical model of the processes occurring in this jitter reduction circuit and identifies the dominant source of additional phase noise and time jitter.

The relationship between the lowest sideband frequency that can be suppressed and the time constant of the dc removal circuit is derived. Also the theoretical limit on sideband noise and jitter performance for a given input and output frequency is estimated. Some practical suppression results are given.

Key Words: Phase noise, Time Jitter, Direct Digital Synthesis, Phase Locked Loops, Noise Cancellation

1. INTRODUCTION

The first announcement of a new proprietary technique of removing the time jitter and hence also the phase noise from any frequency source was made at the 1996 EFTF (Refs.1 and 2). The technique is particularly suitable for removing wideband (> 1 - 10kHz) phase noise and jitter sidebands, such as occur in DDS (Direct Digital Synthesis) or mixer type frequency synthesisers. Its unique feature is that unlike the PLL (Phase Locked Loop) it can remove sideband noise up to the carrier frequency spaced from the carrier without significant compromise of speed of switching from one frequency to another.

The main purpose of this paper is to provide a theoretical estimate of the limits to which noise can be suppressed for typical circuit implementations of the new method. The term "jitter suppression circuit" is used in this paper to cover the "Self Adjusting Delay Compensator - SADC" used in Refs. 2 and 3 and also for "Delay Compensated Phase Noise and Time Jitter Reduction Circuit".

2. BASIC MODEL OF THE DELAY COMPENSATOR

Fig. 1 shows the basic delay compensation circuit. The first monostable creates a pulse of fixed duration from the frequency source. Being of fixed duration, the pulse length can be designed to be very stable which is needed if no additional noise or jitter is to be introduced.

The DC removal circuit is shown being implemented as a feedback arrangement. DC has to be removed so that the output of the following integrator circuit does not ramp into positive or negative saturation. However, in so doing suppression of sideband noise down to DC is precluded. The resulting lower cut-off frequency for sideband noise suppression also determines how rapidly noise suppression is recovered after switching from one frequency to another.

The integrator acts on the pulse waveform shown in Fig. 2a to produce the sawtooth waveform as shown in Fig. 2b. The important point shown in this figure is that the timing instants t_1 and t_5 when the integrator output waveform crosses the switching (comparator) threshold level of the second monostable do not vary even for large time jitter displacements of the integrator input pulses. It is this fact that accounts for the ability of the circuit to suppress the time jitter of any frequency source.

The pulse length of the output monostable is chosen to give approximately equal mark-space ratio over the frequency band covered by the input source.

3. FUNCTIONAL MODEL OF JITTER SUPPRESSION

A representative functional model for the circuit of Fig. 1 is shown in Fig. 3. The two signal paths represent the two separate processes occurring simultaneously in the actual circuit. The upper path represents the time jitter information contained in the relevant up or down transitions of the waveforms.

The lower path represents three separate stages of signal processing. The first stage is the conversion of input frequency to a voltage. The mean level of the output from the first monostable is directly proportional to (pulse) frequency provided the pulse length is kept constant. The integrator next operates on this mean voltage to correct it to a voltage corresponding to phase fluctuations of the input signal. A mean level shift in the integrator output waveform inversely causes a time shift proportional to the slope of the integrator waveform because the input switching level of the second compensator/monostable is fixed. Thus the combination of the integrator output sawtooth waveform and the fixed comparator level acts as a "delay modulator" for the relevant waveform transitions appearing at the output of the comparator/monostable.

The time shift fluctuations are then subtracted from the original signal time fluctuations in the summing junction. A major advantage of the circuit is that the integrator slope is automatically the correct value so the fluctuations in the upper path are almost exactly cancelled (in the summing junction) by those in the lower path. The degree of cancellation is only degraded at sideband frequencies at which the integrator does not behave as a perfect integrator.

4. ESTIMATION OF THE CANCELLATION RATIO

The cancellation ratio or suppression is the phase modulation or time jitter at a given sideband (angular) frequency $\omega_m = 2\pi f_m$ on the input in ratio to that on the output at the same sideband frequency. From Fig. 3 the cancellation or suppression ratio can be derived as a Laplace Transform as:

$$\frac{\tau_0(s)}{\tau_i(s)} = \frac{\tau_i(s) - \tau_c(s)}{\tau_i(s)} = 1 - \frac{\tau_c(s)}{\tau_i(s)} =$$

$$1 - G_c(s) \quad \ldots\ldots\ldots\ldots\ldots (1)$$

where $\tau_0(s)$, $\tau_i(s)$ and $\tau_c(s)$ are respectively the Laplace Transforms of the output, input and cancellation time jitter signals. Time jitter can always be converted to and from phase jitter $\phi(s)$ by the equations:-

$$\phi(s) = 2\pi f_o \tau(s) = 2\pi\tau(s)/T_o \quad \ldots\ldots\ldots\ldots\ldots (2)$$

where $f_o = T_o^{-1}$ is the carrier frequency.

In equation 1 $G_C(s)$ is the closed loop transfer function of the lower path:-

$$G_c(s) = G_m(s)G_1(s)k_d/(1 + G_1(s)G_2(s)) \quad \ldots\ldots\ldots\ldots\ldots (3)$$

where

Input Mean Level: $G_m(s) = k_m \; s \quad \ldots\ldots\ldots\ldots\ldots (4)$

Integrator: $G_1(s) = 1/R_1C_1s = k_1/s \quad \ldots\ldots\ldots\ldots\ldots (5)$

DC Removal feedback :

$$G_2(s) = (R_1/R_3)/(1 + R_2C_2s) = k_2(s + b) \quad \ldots\ldots (6)$$

Delay Modulator Gain : $k_d =$

$$(\text{integrator down slope})^{-1} \quad \ldots\ldots\ldots\ldots\ldots (7)$$

The self adjusting action of the circuit corresponds to

$$k_m k_1 k_d = 1 \quad \ldots\ldots\ldots\ldots\ldots (8)$$

which matches the gains of the upper and lower paths. Equations 3 to 8 then give

$$G_c(s) = s(s + b)/(s^2 + bs + k_2) \quad \ldots\ldots\ldots\ldots\ldots (9)$$

The two zeros in the numerator confirm that DC removal action is present and the path gain drops rapidly to zero below the frequency

$$\omega_b = 2\pi f_b = b = (R_2C_2^{-1})$$

It is convenient to adjust the feedback gain k_2 to give coincident closed loop poles at s = b/2 by setting $k_2 = 1/4b$. This occurs if we make $R_2 = 4R_2C_2/C_1$. Then the suppression ratio given by equation 1 becomes

$$S_r = 1 - G_c(s) = 1 - s(s + b)/(s + b/2)^2$$

$$= 1/(1 + 2s/b)^2 \quad \ldots\ldots\ldots\ldots\ldots (10)$$

Putting $s = j\omega = 2\pi f$ then gives for the magnitude of the suppression ratio $|S_r|$

$$|S_r|^2 = 1/(1 + 4\omega^2/b^2) = (b^2/4)/(b^2/4 + \omega^2) \quad ...(11)$$

This shows that below the DC removal cut off frequency

$2\pi f = \omega = b/2$ there is no suppression

because $|S_r| := 1$

Further more detailed analysis shows that the ultimate suppression ratio is limited approximately to the inverse of the loop gain of the integrator at the carrier frequency of operation. In turn it can be shown that the practical maximum loop gain of an integrator is the ratio of the unity gain transition frequency of the dominant lowest frequency transistor in the integrator circuit to be operating carrier frequency. For example if f_t is 250 MHz the maximum suppression for a 10 MHz carrier frequency will be 250/10 = 25 equivalent to 28dB sideband suppression. Thus if more than 28dB suppression is required two or more jitter suppression circuits should be used in cascade.

5. DOMINANT SOURCES OF NOISE

Also shown in Fig. 3 are two equivalent current sources i_{ns} and i_{nR}. Both sources are shown coupled through directional couplers so that the real impedance values existing at the points of equivalent noise insertion do not short out or attenuate the noise currents.

The noise source i_{ns} represents the "shot" noise which will exist on the current charging the integrator capacitor C_1.

This current is always equal to the current supplied from the output stage of the first monostable. No additional current noise suppression feedback is assumed to be present so that the power spectral density of the (white) shot noise in one Hz bandwidth at modulation frequency f_m is given by the standard formula.

$$|i_{ns}(f_m)|^2 = 2q_e I_o \quad ...(12)$$

where I_o is the integrator capacitor charging current during the down slope part of the integrator output waveform and q_e is the charge on the electron where

$$q_e = 1.6 \times 10^{-19} \text{coulomb} \quad ...(13)$$

The equivalent RMS noise voltage $V_{ns}(f_m) \equiv V_n$ at the integrator output for sideband modulation frequency $\omega_m = 2\pi f_m$ is given by:

$$V_n = i_n/\omega_m C_1 = i_n/2\pi f_m C_1 \quad ...(14)$$

But the current I_o into the capacitor C_1 determines the integrator down slope k_d^{-1} so that

$$k_d = C_1/I_0 \quad ...(15)$$

The conversion of V_n to time jitter $\tau_n(f_m)$ can then be determined by the slope value k_d^{-1} and is $\tau_n = k_d V_n$ ...(16a)

$$= C_1 V_n / I_0 \quad ...(16b)$$

$$= \frac{i_n}{\omega_m C_1} \cdot \frac{C_1}{I_0} = \frac{i_n}{2\pi f_m I_o} \quad ...(16c)$$

But from equations 2 and 16c we have

$$\phi_n = 2\pi f_0 \tau_n = 2\pi\tau_n / T_0 \quad ...(17a)$$

$$= (i_n/I_o)(f_0/f_m) \quad ...(17b)$$

$$\left(\frac{i_n}{I_o}\right)^2 = \frac{2q_e}{I_0} \quad ...(18)$$

The phase noise power spectral density $N(f_m)$ relative to carrier power P at sideband frequency f_m is then given by

$$S_\phi(f) = N(f_m)/P = \frac{1}{4}|\phi_n|^2 \quad ...(19a)$$

$$= \frac{q_e}{2I_0}\left(\frac{f_o}{f_m}\right)^2 \quad ...(19b)$$

For a representative case of integrator charging current $I_o = 10\text{mA}$ and output frequency $f_o = 10\text{MHz}$ we have for sideband frequency $f_m = 10\text{kHz}$

$$S_\phi(f) = \frac{1.6 \times 10^{-19}}{2 \times 10^{-2}} \times \left(\frac{10^7}{10^4}\right)^2 = 8 \times 10^{-12}$$

$$\equiv -101\text{dBC} \quad ...(20)$$

This is more than good enough for many applications but the performance of an oscillator good enough for use in high performance communications systems should reach the figure

merit of -120dB at 10kHz, a difference of 9dB.

This indicates that a technique such as feedback for current noise reduction should be applied when better phase noise performance is required. But then the noise from the second equivalent (resistor) noise source i_{nR} can become dominant. This noise contribution is assumed to derive from an equivalent resistor R_e which in Fig. 3 can be taken to be R_1.

We have

$$i_{nR}(f_m) = i_{nR} = (4kT/R_e)^{\frac{1}{2}} =$$

$$(4 \times 4 \times 10^{-21}/R_e)^{\frac{1}{2}} .....(21)$$

But $|i_{nR}|^2$ can replace $2q_eI_o$ in equations

to give

$$S_\phi(f) = \frac{kT}{I_o{}^2 R_e}\left(\frac{f_o}{f_m}\right)^2 .....(22a)$$

$$= \frac{4 \times 10^{-11}}{R_e} .....(22b)$$

$$= \frac{4 \times 10^{-11}}{400} = 10^{-13} \equiv -130\text{dBC}.....(22c)$$

Equation 22c is also for

$I_o = 10\text{mA}$ and with $R_1 = R_e = 400\text{ohms}$

This gives good performance at 10MHz although about 10dB less noise is required if the carrier frequency is moved up to 100MHz.

The conclusion from these two examples is that additional current and resistor noise reduction are desirable but only for the most demanding of applications. There are standard methods for achieving better noise but even for these the noise analysis methods given here remain valid.

6. RESULTS AND CONCLUSIONS

Figs. 4 and 5 illustrate 25dB jitter sideband suppression being achieved at 10 MHz in a particular implementation of the jitter reduction circuit (Ref.2). In this case 2N2222 transistors with $f_T = 250$ MHz were used in the integrator. The 24dB can be compared with a predicted cancellation of 250/10 = 25 ≡ 28dB.

On this basis it can be predicted that a cancellation of about 28dB should be achievable at 300MHz by using the BFP193 with an f_T of 8GHz in place of the 2N2222.

Examination of the design equations show that an increased integrator charging current decreases the phase noise at a rate of 3dB per octave of current I_O increase. The resistor noise contribution is correspondingly decreased at a rate of 6dB per octave of I_O increase. An increase of I_o to 30mA thus would give respectively 9.5dB and 19dB improvements. Power consumption is kept low by decreasing the supply voltage which is facilitated by using a large integrator capacitor to decrease the integrator output voltage swing. Interestingly the noise performance is to a first approximation independent of the integrator capacitor value.

In conclusion the analysis given in this paper has established some of the fundamental limits of the novel phase noise and time jitter reduction technique. The results so far confirm that analysis and further measurements are planned. From this it can be seen that the new technique promises to have wide application to phase noise and time jitter in many electronic systems.

REFERENCES

[1] M J Underhill "Phase noise reduction circuit". UK Patent Application 9603223.

[2] M J Underhill & M J Blewett "Spectral improvement of direct digital frequency synthesisers and other frequency sources", Proc. 10th EFTF, March 1995, pp.452-460 (IEE Conf. Pub. 418).

[3] M J Underhill & M J Blewett, "Phase noise and time jitter reduction circuit". 1996 Conf. Proc. Microwaves and RF, London UK, Oct. 1996, pp.361-366.

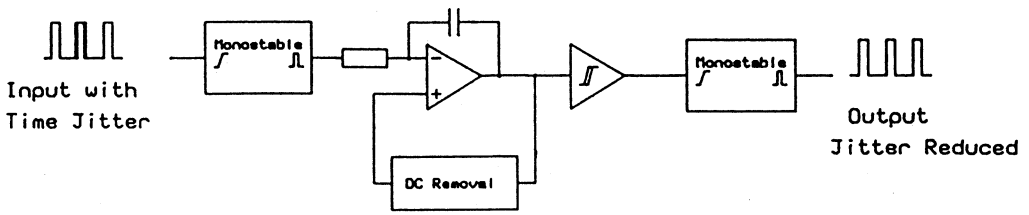


Fig. 1 Phase Noise and Time Jitter Reduction Circuit

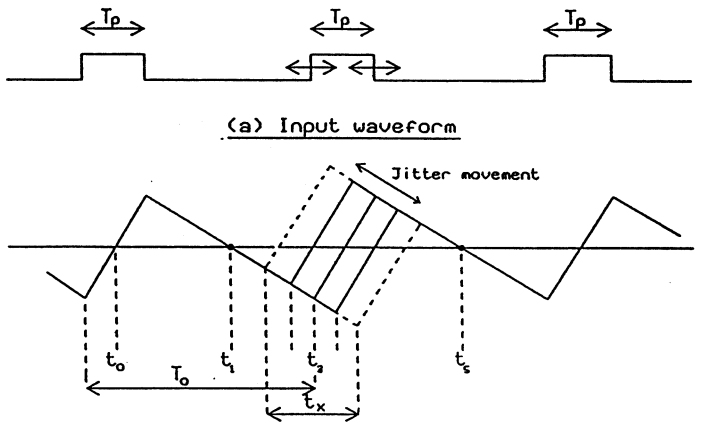


Fig 2 SADC Waveforms

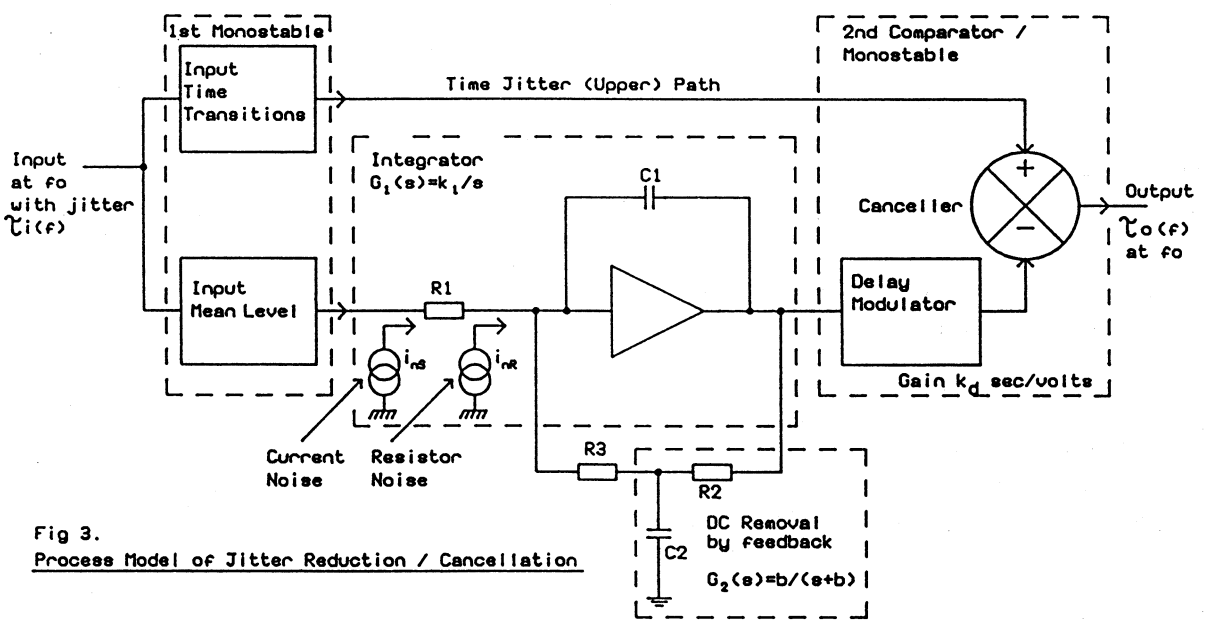


Fig 3.
Process Model of Jitter Reduction / Cancellation

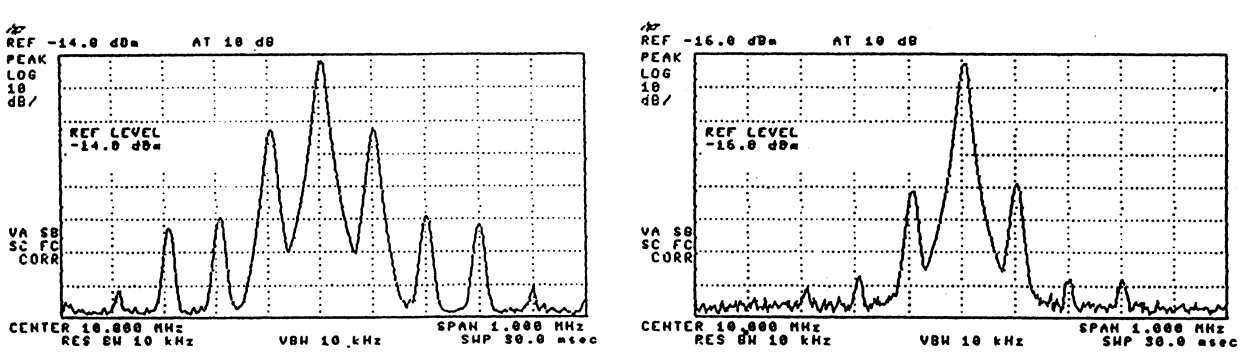


Fig. 4 Spurious Sideband Reduction

POSTERS - PIEZOELECTRICITY

FREQUENCY STABILITY IN A SUPERCONDUCTING TRANSMISSION LINE WITH BUFFER LAYERS

Farhat Abbas, John Gallop and Clifford Hodge

National Physical Laboratory, Teddington, Middlesex, UK TW11 0LW

Abstract

The resonant frequency of a superconducting planar resonator is dependent on the material properties and thicknesses of the superconductors, the dielectric (substrate) and the buffer layers. In this paper, we compute the first and second derivatives with respect to temperature of a wave velocity ratio (with respect to free space) for various combinations of material properties. The dependence of resonant frequency on the dielectric constant and thicknesses of the buffer layers is discussed.

Keywords: Frequency Standards; Superconducting Resonators; Buffer Layers.

1 Introduction

The use of superconducting films in transmission lines has many advantages for signal processing applications such as low dispersion, low loss, and wide bandwidth. In fact, a study of the behaviour of these transmission lines is of interest not only to measure the penetration depth and low-frequency resistance of superconducting thin films, but also to determine the characteristics of thin-film superconducting devices. Passive microwave devices such as filters, resonators and delay lines require high-quality HTS thin-films and substrate materials.

Workers attempting to grow HTS films directly onto high-quality substrate materials have encountered some serious problems, due to large mismatches of both lattice constants and thermal expansion coefficients of the HTS films and some substrate materials. Also, the interdiffusion between the HTS films and substrate materials has been found to severely degrade the superconducting properties (Refs. 1-3). To reduce such problems, it is necessary to investigate suitable buffer layers. The buffer layers should satisfy the minimum five requirements for cryogenic microwave microelectronics applications:

a) prevention of interdiffusion;

b) better thermal expansion coefficient matching;

c) better lattice matching;

d) confinement of the field into the substrate;

e) provide a resonator with a temperature independent frequency.

In this study, an analysis has been completed to explore the possibility of realising a resonator with a temperature independent frequency. The design depends on the material properties and thicknesses of the superconductors, the dielectric (substrate), and

the buffer layers. The first and second derivatives of propagation phase velocity with respect to temperature for a sample in vacuo are computed for various combinations of material properties and geometric factors. The resonator with a temperature independent frequency, as shown in Fig. 1, also depends on the product of dielectric constant of the substrate and the thicknesses of buffer layers, and vice versa.

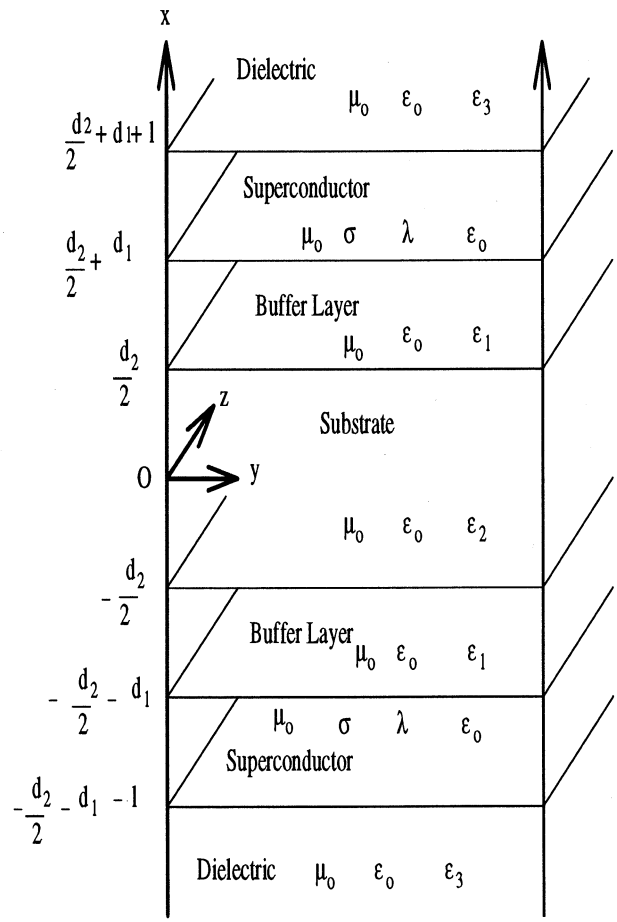


Figure 1: Cross-section of a superconducting transmission line with buffer layers.

2 Field solution

Fig. 1 shows the geometry of the superconducting transmission line with buffer layers which we have considered. The structure consists of a pair of buffer layers (dielectrics, region 1) bridged by a central substrate (dielectric, region 2) and sandwiched by a pair of superconducting thin films. Thicknesses of the thin films, buffer layers and substrate are l, d_1 and d_2, respectively. The dielectrics, in region 3, are considered to be very thick so that the fields in these regions can be assumed to exponentially decay away from the interfaces.

The procedure used to obtain a field solution is the same as in (Ref. 1), and it is summarised below. Consider the propagation of an electromagnetic wave in the z-direction of the transmission line shown in Fig. 1. It is assumed that the dielectric thicknesses (d_1 and d_2) and the penetration depth λ of the high temperature superconductors are very small compared to the width of the line, which in turn is very small compared to the length of the line. From Fig. 1, and the above assumptions, it is clear that the edge effects can be neglected, and there is no y-dependence of the fields and currents.

The two-fluid model is used for the superconductors, in which the total current is the sum of the supercurrent and the normal current. Classical skin effect and London theory are assumed for the normal current and the supercurrent, respectively.

The formulations made in (Ref. 1) can be used, i.e., we are considering a TM wave:

$$\bar{H}_y = \frac{1}{\alpha\mu_o\omega}\left(\alpha^2-\kappa^2\right)\bar{E}_x$$
$$\bar{E}_z = -\frac{\iota}{\alpha}\frac{d\bar{E}_x}{dx}$$

$$\frac{d^2\bar{E}_x}{dx^2}-\kappa^2\bar{E}_x=0 \qquad (1)$$

where, for the dielectrics:

$$\kappa^2 \triangleq K_r{}^2 = \alpha^2-\omega^2\epsilon_r\mu_o, \qquad r=1,2,3 \qquad (2)$$

while for the superconductors:

$$\kappa^2 \triangleq \kappa^2 = \frac{1}{\lambda^2} + \alpha^2 - \omega^2\epsilon_o\mu_o + \iota\omega\mu_o\sigma \qquad (3)$$

Here, α is the propagation constant along the z direction (taking $e^{-\iota\alpha z}$), ω is the angular frequency (assuming $e^{\iota\omega t}$), ϵ_o and μ_o are the permittivity and the permeability of vacuum respectively, ϵ_r is the dielectric constant of the dielectrics, λ and σ are the penetration depth and the conductivity of the superconductors, respectively.

Equation (1) is a second-order differential equation which has two independent solutions of the form $e^{\kappa x}$ and $e^{-\kappa x}$, where κ is taken to be the root of κ^2 with positive real part. In the positive x-direction of the dielectric, region 3, we take only the solution $e^{-K_3 x}$, and in the negative x-direction we take only the solution $e^{K_3 x}$, discarding $e^{K_3 x}$ for positive x-direction, and $e^{-K_3 x}$ for negative x-direction. In the superconductors, the buffer layers (dielectrics, region 1) and in the substrate (dielectric, region 2), both solutions are retained in order to satisfy the boundary conditions. In the superconductors, and in regions 1 and 2 of the dielectrics, we need both solutions in order to satisfy the boundary conditions.

With these solutions in the various media, we have twelve arbitrary constants for the amplitudes of the fields (one each in the dielectrics, region 3, two each in the superconductors, the buffer layers (dielectrics, region 1) and the substrate (dielectric, region 2)). There are twelve boundary conditions that must be satisfied, namely the continuity of the tangential fields $\bar{E}_z$ and $\bar{H}_y$ at the six boundaries shown in Fig. 1. If we ignore any non-linearity in the system, the characteristics of the line are independent of the amplitude of the wave, and eleven of the constants can be determined in terms of the twelfth by using eleven of the twelve boundary conditions. The twelfth boundary condition gives an equation for the propagation constant α, which must be satisfied in order for a solution to exist.

The condition is a transcendental equation for which an exact solution cannot be readily obtained. As discussed in (Ref. 4), the approximations are $K_1 d_1 \ll 1$ and $K_2 d_2 \ll 1$, and physically these approximations mean that higher order modes are ignored. With small d_1 and d_2, higher order modes will not be excited. With these assumptions, the transcendental equation yields:

$$\alpha^2 = \frac{\omega^2\mu_o\epsilon_o\epsilon_1\epsilon_2}{(2d_1\epsilon_2 + d_2\epsilon_1)}\left[2\lambda\coth(\frac{l}{\lambda}) + 2d_1 + d_2\right] \qquad (4)$$

If we make either d_1 or d_2 equal to zero (i.e. removing the buffer layer), equation 4 is exactly the same as that produced in (Ref. 5), and identical results are obtained in (Ref. 1). Also, note the factor of two in equation 4, which indicates that there are two buffer layers and two superconductor layers. This is because we have considered a symmetric case, i.e., the same HTS thin films on each side of the substrate separated by the same buffer layers, as shown in Fig. 1. The wave velocity relative to that in a vacuum can be written as follows from equation 4:

$$V_r = \sqrt{\frac{(2d_1\epsilon_2 + d_2\epsilon_1)}{\epsilon_1\epsilon_2\left[2\lambda\coth(\frac{l}{\lambda}) + 2d_1 + d_2\right]}} \qquad (5)$$

According to equation 5, the wave is dispersionless even though there is a component of the electric field in the direction of propagation, i.e., the group velocity and phase velocity are equal and independent of frequency. The dispersion relation of the wave has been discussed in (Ref. 4).

3 Results

The superconducting transmission line with buffer layers shown in Fig. 1 can be described by the pene-

tration depth λ of the superconductors, the dielectric constants ϵ_r of the r^{th} dielectrics, and the thicknesses d_r and l of the r^{th} dielectrics and the superconductors.

The temperature dependence of the penetration depth λ of a superconductor can be described by any one of several models outlined in (Ref. 5). Any of those models can be used in our analysis. However, we will concentrate on the following approximate result:

$$\lambda = \frac{\lambda_o}{\sqrt{1-(T/T_c)^p}} \qquad (6)$$

In equation (6), if the Gorter and Casimir model is assumed, then p = 4. However, recently (Ref. 6), the spin-polaron theory of high-T_c superconductivity has been explored, in which the charge carriers in a high-T_c superconductor are considered as biholes obeying the Bose-Einstein statistics and localized within a unit cell of the crystal lattice. If the charge carrier system in a high-T_c superconductor is considered as an ideal Bose-Einstein gas, then p = 1.5. In this paper, we will assume p = 4. Also, the temperature dependence of the ϵ_r can be approximated (Ref. 7) as $\epsilon_1 = 9.2 + 2.5 \times 10^{-11}T^4$ (sapphire) and $\epsilon_2 = 113.446 + 0.043T - 0.002T^2 + 7.724 \times 10^{-6}T^3 - 1.072 \times 10^{-8}T^4$ (rutile).

To provide a resonator with a temperature independent frequency, or a transmission line with temperature independent propagation constant (and therefore phase shift) along the line, it is necessary to choose a buffer layer/main dielectric configuration which will cause the propagation velocity (or equivalently the transmission line wavelength) to be as independent of temperature variations as possible at the selected operating temperature.

Expanding expression (5) in a Taylor series about the operating temperature T_0 leads to:

$$\delta V_r(T) = \frac{dV_r}{dT}\delta T + \frac{1}{2}\frac{d^2V_r}{dT^2}\delta T^2 + ... \qquad (7)$$

where the partial derivatives are evaluated at T_0, and where $\delta T = (T - T_0)$. If a certain temperature stability δT can be achieved, then the minimum variation in $V_r(T)$ is attained if as many of the lower order partial derivatives as possible can be made zero, or close to zero. The first order approximation is to produce a turning point in $V_r(T)$ by ensuring $dV_r/dT = 0$ at T_0. But, by judicious choice of geometry factors for a particular combination of dielectrics and superconductors, it is also possible to make d^2V_r/dT^2 zero, and even possibly higher order terms.

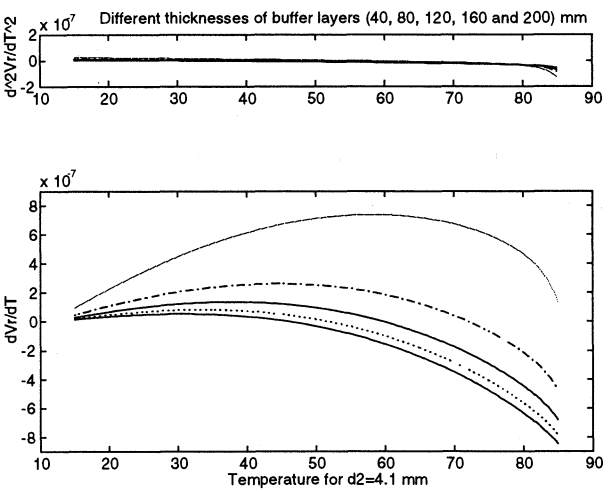


Figure 2: Second and first derivatives of propagation phase velocity as a function of temperature for a variety of buffer layer thicknesses, 40 mm to 200 mm.

Fig. 2 shows the first and second derivatives of propagation phase velocity as a function of temperature for a variety of buffer layer thicknesses, assuming sapphire as the buffer layer and rutile as the main dielectric material. Turning points in $V_r(T)$ can be produced close to any chosen operating temperature in this way.

4 Conclusions

A field solution has been found for a superconducting transmission line with buffer layers, and an expression for the relative velocity of the dominant TM mode has been developed. The derived expression is used to investigate the resonator with a temperature independent frequency which is dependent on the material properties of the superconductors, the dielectric (substrate) and the buffer layers.

Using the Gorter and Casimir model, the variations of the first and second derivatives of V_r (dV_r/dT and d^2V_r/dT^2) with respect to temperature and varying dielectric thicknesses are shown.

An example of YBCO thin films on rutile with sapphire buffer layers has been computed. From this example, it may be concluded that if a turning point can be realized at T = 60 K, and the temperature controlled to better than 0.1 mK, then frequency standards with stabilities of parts in 10^{15} should be attainable.

References

[1] Farhat Abbas and L E Davis, "Propagation coefficient in a superconducting asymmetrical parallel-plate transmission line with a buffer layer", J. Appl. Phys. Vol. 73, No. 9, 1 May, 1993, pp 4494-4499.

[2] X D Wu, A Inam, M S Hegde, B Wilkens, C C Chang, D M Hwang, L Nazar, T Venkatersan, S Miura, S Matsubara, Y Miyasaka and Shohata, "High critical currents in epitaxial $YBa_2Cu_3O_{7-x}$ thin films on silicon with buffer layers" , Appl. Phys. Lett. 54, 754 (1989).

[3] S Y Lee and H H Park, "Effect of $PrBa_2Cu_3O_{7-x}$ Buffer Layer Thickness on the Properties of $YBa_2Cu_3O_{7-x}$ Thin Films Grown on Sapphire by Laser Ablation", J. Superconductivity, Vol. 9, No. 5, 1996, pp 545-549.

[4] Farhat Abbas, L E Davis and J C Gallop, "Ultra-high-Q resonators for low-noise, microwave signal generation using sapphire buffer layers and superconducting thin films", Supercond. Sci. Technol. 7 (1994) 495-501.

[5] Farhat Abbas, L E Davis and J C Gallop, "Field Solution for a Thin-Film Superconducting Parallel-Plate Transmission Line", Physica, C, 215, 132-144, 1993.

[6] N F Mott,"The spin-polaron theory of high-Tc superconductivity", Advances Phys. Vol. 39, No. 1, 1990, pp 55-81.

[7] J C Gallop, C D Langham, L Hao and Farhat Abbas,"Dielectric loaded HTS resonators as frequency standards and low phase noise oscillators", IEEE Trans. Instrumentation and Measurement, Vol. 46, No. 2, April 1997.

ALUMINUM AND ALKALI-RELATED HYDROXYL DEFECTS IN QUARTZ CRYSTALS AND THEIR RADIATION EFFECTS

Harish Bahadur
National Physical Laboratory
Hillside Road, New Delhi-110012. INDIA

Hydrogen is an ubiquitous impurity in quartz and affects nearly every application of it as a device.Hydrogen occurs in the form of hydroxyl defects trapped at numerous unidentified sites in quartz lattice.Most of the OH-related defects in crystalline quartz have the overtone frequencies of their stretching vibrations in the near infrared region. It is now realized that these hydroxyl defects act as a source of hydrogen to compensate the electron excess defects of tetravalent aluminum, substituted for trivalent silicon in quartz lattice, during irradiation of quartz.

This paper presents our investigations on infrared absorption measurements of quartz crystals to characterize the aluminum- and alkali-related hydroxyl defects in natural and cultured quartz crystals. Irradiation was done by using high energy electrons from a Van de Graaff electrostatic accelerator. Quartz samples were irradiated with electron beam of 1.75 MeV and dose of 2 Mrad at 77 K and 300 K. Irradiation at these different temperatures separates the movements of two types of charge compensators at aluminum sites, alkalis and protons in quartz lattice. While the alkalis in quartz move under the radiation field only if the sample temperature is about or above 200 K, the protons move at all temperatures down to 10 K. Using these effects, it has been shown that the technique of near infrared absorption can be utilized to measure the concentration of various impurity-related point defects such as Al-Li^{+} and Al-Na^{+} in natural quartz, Al-Li^{+}, Al-OH^{-}, total aluminum and Al-hole centers in the frequency offsets of quartz resonators in a radiation environment. These studies are primarily important in aerospace frequency control applications such as in satellite-borne frequency standards and in possibly developing better growth procedures in the hydrothermal synthesis of cultured quartz.

PURE MODES IN LITHIUM NIOBATE AND LITHIUM TANTALATE THAT PERMIT ACCURATE DETERMINATION OF FIRST-ORDER MATERIAL PROPERTIES

Carl K. Hruska, Arthur Ballato*) and Peggy Ng

Piezoelectricity Research Laboratory
York University, North York, Ontario M3B 3C7 Canada
*) US Army Research Laboratory, AMSRL-SE, Fort Monmouth, NJ 07703, USA

1. ABSTRACT

This paper lists the first-order (linear) material constants of lithium tantalate and lithium niobate that can be determined using the pure thickness modes of plates generally defined by Kosinski, Lu and Ballato in IEEE Trans. UFFC 1993, pp.258-264. Nine experimental samples of each material of relatively modest orientation accuracy (standard error of 10') can be used to determine four isolated linear material constants with standard errors of the order of 0.01% and two combinations of these constants with errors of about 0.5%.

Keywords: lithium tantalate, lithium niobate, first-order material properties, linear material constants.

2. INTRODUCTION

The propagation of small amplitude plane waves through a piezoelectric medium is governed by the well known equation

$$\begin{vmatrix} \Gamma_{11} - \Gamma & \Gamma_{12} & \Gamma_{13} \\ \Gamma_{21} & \Gamma_{22} - \Gamma & \Gamma_{23} \\ \Gamma_{31} & \Gamma_{32} & \Gamma_{33} - \Gamma \end{vmatrix} = 0 \ , \quad (1)$$

where

$$\Gamma_{ik} = \alpha_j \alpha_l \left(c_{ijkl} + \alpha_m \alpha_n e_{mij} e_{nkl} / \varepsilon \right), \quad (2)$$

$$\varepsilon = \alpha_a \alpha_b \varepsilon_{ab}, \quad (3)$$

and where c_{ijkl}, e_{mij} (and e_{nkl}), and ε_{ab} are, respectively, the first-order, i.e., linear elastic, piezoelectric, and dielectric constants of the piezoelectric material under consideration; α_j, α_l ... are the direction cosines of the plane wave defined in the basic reference frame of the crystal according to the IEEE 1978 Standart (Ref.1). The Einstein index summation is in effect in Eqs.2 and 3.

Experimentally determined eigenvalues Γ have been frequently used to determine the above material constants. Typically, the emphasis has been on finding the minimum number of eigenvalues Γ necessary to determine the maximum number constants. Less attention has been paid to a formal description of their reliability and accuracy.

This paper aims at complementing the past effort by determining an optimum selection of eigenvalues Γ not to produce a set that contains their fewest number, but rather a statistically robust, overdetermined set that will permit the determination of the requisite constants with a compromise between the number of cuts required and the quality of the results. Of necessity this places an emphasis on a clearly defined reliability and accuracy and on the principal factors affecting them.

The selection of eigenvalues Γ (and the concomitant recommendation concerning experimental specimens to be used) is made from among the pure modes defined by Kosinski, Lu and Ballato in Ref.2.

Numerical results are stated for lithium niobate and lithium tantalate. The aim is to show the extent to which the pure modes can be used to replace the existing values of the material constants by new ones possessing known reliability and accuracy.

The values of all constants mentioned in this paper are stated in the frame of reference according to Ref.1. Wherever the current values of these constants are needed, they are taken from Smith and Welsh (Ref.3).

3. THE PURE MODES AND THEIR POTENTIAL

The eigenvalue problem (Eq.1) has three solutions frequently denoted Γ^A, Γ^B and Γ^C, where $\Gamma^A > \Gamma^B \geq \Gamma^C$. In turn they correspond to the quasi-longitudinal wave/vibration mode A, and the fast and slow quasi-transverse modes B and C. The eigenvalues Γ^A, Γ^B and Γ^C can be determined experimentally and used to calculate the linear material constants.

The relationship between Γ^A, Γ^B and Γ^C and the material constants is generally inconveniently complicated. An important exception occurs if one or more of them are equal to any of the diagonal elements Γ_{11}, Γ_{22} or Γ_{33} in Eq.1. Such eigenvalues are said to be associated with pure modes (Ref.2).

In application to lithium niobate and lithium tantalate and to crystals of class 3m in general the eigenvalues of the pure modes are given by

$$\begin{aligned}\Gamma_{11} = {} & (\alpha_1^2 + 0.5 * \alpha_2^2) c_{11} - 0.5 \alpha_2^2 c_{12} \\ & + 2 \alpha_2 \alpha_3 c_{14} + \alpha_3^2 c_{44} + \alpha_1^1 \alpha_3^2 e_{15}^2/\varepsilon \\ & - 4 \alpha_1^2 \alpha_2 \alpha_3 e_{22} (e_{15} + e_{31})/\varepsilon \\ & + \alpha_1^2 \alpha_3^2 e_{31} (2 e_{15} + e_{31})/\varepsilon \\ & + 4 \alpha_1^2 \alpha_2^2 e_{22}^2/\varepsilon \end{aligned} \qquad (4a)$$

$$\begin{aligned}\Gamma_{22} = {} & (0.5 \alpha_1^2 + \alpha_2^2) c_{11} - 0.5 \alpha_1^2 c_{12} \\ & - 2 \alpha_2 \alpha_3 c_{14} + \alpha_3^2 c_{44} + \alpha_2^2 \alpha_3^2 e_{15}^2/\varepsilon \\ & - 2 \alpha_2 \alpha_3 (\alpha_1^2 - \alpha_2^2) e_{22} (e_{15} + e_{31})/\varepsilon \\ & + 2 \alpha_2^2 \alpha_3^2 e_{15} e_{31}/\varepsilon + (\alpha_1^2 - \alpha_2^2)^2 e_{22}^2/\varepsilon \\ & + \alpha_2^2 \alpha_3^2 e_{31}^2/\varepsilon \end{aligned} \qquad (4b)$$

$$\begin{aligned}\Gamma_{33} = {} & \alpha_3^2 c_{33} + (\alpha_1^2 + \alpha_2^2) c_{44} \\ & + (\alpha_1^2 + \alpha_2^2)^2 e_{15}^2/\varepsilon \\ & + 2 \alpha_3^2 (\alpha_1^2 + \alpha_2^2) e_{15} e_{33}/\varepsilon \\ & + \alpha_3^4 e_{33}^2/\varepsilon . \end{aligned} \qquad (4c)$$

In transition from Eq.2 to Eqs.4a-c the tensor index notation of the material constants has been replaced by the more usual matrix index notation.

Eqs.4a-c include five elastic constants, c_{11}, c_{12}, c_{14}, c_{33}, c_{44} and eight products of piezoelectric constants e_{15}^2, $e_{22}(e_{15}+e_{31})$, $e_{31}(2e_{15}+e_{31})$, e_{22}^2, $e_{15}(e_{15}+2e_{31})$, e_{31}^2, $e_{15}.e_{33}$, e_{33}^2. This defines the upper limit of the potential of the pure modes for determining the linear material constants of any crystal of class 3m. The number of these constants actually calculable for a given crystal will depend on the availability of suitable pure modes.

The above constants will be calculated from a linear system consisting of Eqs.4a-c using linear regression. The the dielectric constants appearing in Eqs.4a-c in the denominator of the stiffening terms cannot be obtained in this manner and need to be known beforehand.

4. AVAILABLE PURE MODES

The direction cosines α_1, α_2, and α_3, in Eqs.2 and 3 may be viewed as the direction cosines of the thickness of a piezoelectric plate of orientation (xzlw)ψ/ϕ, where ψ and ϕ are two independent rotational angles defined in Ref.1. As a result the matrix elements Γ_{ij} can be recorded as functions of the material constants and the two rotational angles ψ and ϕ.

Using the sufficient conditions

$$\Gamma_{12} = \Gamma_{31} = 0 \qquad (5a)$$

$$\Gamma_{23} = \Gamma_{12} = 0 \qquad (5b)$$

$$\Gamma_{31} = \Gamma_{23} = 0 \qquad (5c)$$

resulting, respectively, in $\Gamma = \Gamma_{11}$, $\Gamma = \Gamma_{22}$, and $\Gamma = \Gamma_{33}$, the search for the pure modes was made for both lithium niobate and lithium tantalate in the entire domain of the rotational angles ψ and ϕ.

In general the definition of the pure modes depends on the material constants used for their calculation (Ref.3). However, in some cases it is possible to satisfy some of Eqs. 5a-c independently of their values; the corresponding pure modes are called here 'material-independent'. The remaining pure modes are referred to as 'material-dependent'.

The results are shown in Table I. Only those pure modes which can provide independent information about the sought-after material constants are listed .

Table I.

material-independent			
ψ	ϕ	Γ	mode
0°	0°	Γ_{11}	A
90°	$[0^\circ,180^\circ)$	Γ_{11}	C or B*)
material-dependent - lithium niobate			
ψ	ϕ	Γ	mode
90°	4.880°	Γ_{22}	A
90°	4.880°	Γ_{33}	B
material-dependent lithium tantalate			
ψ	ϕ	Γ	mode
-7.750°	36.585°	Γ_{22}	B
90°	6.550°	Γ_{22}	A
90°	6.550°	Γ_{33}	B

List of potentially useful pure modes. Some of the pure modes form a continuous locus in the (ψ,ϕ) domain, others are just isolated points. They are described by the two rotational angles ψ and ϕ; fractions of degrees, stated in the decimal form, are determined with a maximum absolute error of 9 angular seconds. The eigenvalue Γ indicates which of the conditions 5a-c is satisfied; The letters A, B, or C indicate the type of wave/vibration mode.
*) The mode (B or C) depend on ϕ and on the material.

5. LINEAR SYSTEM TO COMPUTE THE MATERIAL CONSTANTS

Each pure mode in Table I is associated with an equation

$$\Gamma = \Gamma_{11} , \qquad (6)$$

where Γ is determined experimentally and Γ_{11} is given by one of Eqs.5a-c chosen according to Table I. From among these linear equations a linear system can be selected and used to compute the linear material constants.

Three types of errors degrading the quality of the numerical formulation of the equations forming the system have been considered:

(1) random experimental errors in the eigenvalue Γ, also including random errors in the critical dimension of the specimens used;

(2) random orientation errors by which the orientation angles ψ and ϕ of experimental specimens differ from the calculated, nominal orientations of the pure modes;

(3) systematic errors in the equations constructed for the material-dependent pure modes, whose orientation angles ψ and ϕ may be inaccurate because they have been calculated using material constants of insufficient reliability; such errors generate an unknown perturbation Δ changing Eq.6 into

$$\Gamma = \Gamma_{11} + \Delta . \qquad (7)$$

The matrix of the linear system has been analyzed and the effect of the above errors assessed by extensive simulations for a large spectrum of linear systems varying in number and composition. It has resulted in the following general conclusions and recommendations:

(1) The material-independent pure modes to be used should include the pure mode defined by $\psi = 0^\circ$ and $\phi = 0^\circ$ and a selection of 6 to 8 modes from among the locus defined by $\psi = 0^\circ$ and $\phi \in [0^\circ, 180^\circ)$; insofar as possible, the choice

of the locus angles ϕ should be uniformly distributed along the interval $[0^\circ;180^\circ)$; pure modes whose eigenvalue Γ is too close to another eigenvalue should be avoided; in case that resonators be used for experimentation, pure modes too difficult to excite should be avoided as well;

(2) all the material-dependent pure modes should be used with the exception of the mode of lithium tantalate defined by the angles $\psi = -7.750^\circ$ and $\phi = 36.585^\circ$; this mode should be avoided as a potential source of intolerably large errors in the material constants to be calculated;

(3) to reduce the errors in the material constants, their calculation should be done in two steps; first solving the system based on the material-independent modes, then by solving the system based on the material-dependent modes; the determination of the material-dependent modes for the second system should be made using the reliable values of the material constants obtained by the solution of the first system.

6. RESULTS

In general, the available material-independent modes can be used to determine elastic constants c_{11}, c_{12}, c_{14}, c_{44}. The recommended selection of the modes for this purpose is listed in Table II.

Table II

lithium niobate ψ/ϕ	mode	lithium tantalate ψ/ϕ	mode
$0^\circ/0^\circ$	A	$0^\circ/0^\circ$	A
$90^\circ/25^\circ$	C	$90^\circ/10^\circ$	C
$90^\circ/55^\circ$	C	$90^\circ/40^\circ$	B
$90^\circ/85^\circ$	C	$90^\circ/70^\circ$	B
$90^\circ/115^\circ$	B	$90^\circ/100^\circ$	C
$90^\circ/145^\circ$	B	$90^\circ/130^\circ$	C
$90^\circ/175^\circ$	C	$90^\circ/160^\circ$	C

Recommended material-independent pure modes to determine elastic moduli c_{11}, c_{12}, c_{14}, and c_{44}.

The available material-dependent modes do not provide access to any isolated material constants. They can be used to determine only their combinations k_{1515} and k_{2222} defined Table III.

Table III

lithium niobate

$$k_{1515} = e_{15}^2 + 2.870613E\text{-}12\ c_{33} + 0.0146\ e_{15}.e_{33} + 0.0001\ e_{33}^2$$

$$k_{2222} = e_{22}^2 - 0.020925E\text{-}12\ c_{33} - 0.1708\ (e_{15}+ e_{31}).e_{22} + 0.0146\ e_{15}.e_{31} - 0.0001\ e_{15}.e_{33} + 0.0073\ e_{31}^2$$

lithium tantalate

$$k_{1515} = e_{15}^2 + 5.036077E\text{-}12\ c_{33} + 0.0264\ e_{15}.e_{33} + 0.0002\ e_{33}^2$$

$$k_{2222} = e_{22}^2 - 0.066393E\text{-}12\ c_{33} - 0.2296\ (e_{15}+ e_{31}).e_{22} + 0.0264\ e_{15}.e_{31} - 0.0003\ e_{15}.e_{33} + 0.0132\ e_{31}^2$$

Combinations of material constants that can be obtained from the material-dependent modes in Table IV.

The combinations k_{1515} and k_{2222} can be calculated using the linear equations obtained from the material-dependent modes in Table IV.

Table IV

lithium niobate ψ/ϕ	mode	lithium tantalate ψ/ϕ	mode
$90^\circ/4.880^\circ$	A	$90^\circ/6.550^\circ$	A
$90^\circ/4.880^\circ$	B	$90^\circ/6.550^\circ$	B

Material-dependent modes to determine combinations k_{1515} and k_{2222}.

The standard errors of both the elastic constants and the combinations of material constants depend on the size of the experimental and orientation errors. These are assumed to be normally distributed with a standard deviation of σ_E and σ_0, respectively. The standard errors of the combinations are further dependent on the quality of the material constants which had to be taken from an external source to determine the material-dependent pure modes.

The maximum absolute standard errors of the elastic constants and combinations of constants obtained for several combinations of σ_E and σ_0 are stated in Table V.

Table V

elastic constants	combinations	σ_0	σ_E
± 0.0025	± 0.10	10'	0.1%
± 0.0024	± 0.09	5'	0.1%
± 0.0008	± 0.07	10'	0.01%
± 0.0003	± 0.04	5'	0.01%

The maximum standard errors of the elastic constants c_{11}, c_{12}, c_{14} and c_{44} in 10^{11} N/m^2 and of the combinations k_{1515} and k_{2222} in C^2/m^4. Obtained for standard errors σ_0 of orientation angles ψ and ϕ and σ_E of the experimental values of Γ.

The maximum standard errors Table V have been calculated assuming that the external source provides material constants that do not differ from their true values by more than 8%.

The experimental values Γ in Eq.(6) pertaining to the pure modes listed in Tables II and IV can be determined by means of the thickness modes of plates using the lateral (Table II) or thickness excitation (Table IV).

7. CONCLUSION

The method of pure modes can be used to determine reliable values of four elastic constants and two combinations of the elastic and piezoelectric constants of lithium niobate and lithium tantalate.

Considering the size of the elastic constants, the material-independent pure modes produce results with an absolute accuracy of 0.01% while the demands on the accuracy of the orientation of the experimental samples (standard error ≤ 10') as well as experimental accuracy (standard error ≤ 0.01 %) are very modest. Better accuracy can be attained most effectively by further reducing the experimental errors.

Under the same conditions the accuracy of the combinations of material constants obtained using the material-dependent pure modes is about 0.5%. This is mainly due to the need to rely on the material constants obtained from external sources. Their increased accuracy would be the most effective way to reduce the errors here.

The fact that the described application of pure modes is not suitable for the determination of the dielectric constants is no substantial disadvantage as they can be obtained by other methods. More disappointing is the fact that pure modes, in principle, provide no access to the elastic constant c_{13}.

8. ACKNOWLEDGEMENTS

This work was supported by the Physical Science Directorate under the auspices of the U.S. Army Research Office Scientific Services Program and by the Natural Sciences and Engineering Research Council of Canada. Pauline Hruska helped with the final stages of preparation of this paper.

9. REFERENCES

[1] IEEE Standard on Piezoelectricity, Std 176-1978. New York: IEEE, 1978.

[2] Kosinski J.A., Lu Y, and Ballato A., IEEE Trans. UFFC 40(1993), 258-64.

[3] Smith R.T. and Welsh F.S., J.Appl.Phys. 42 (1971), pp.2219-30.

MEASUREMENTS OF ELASTIC AND PIEZOELECTRIC CONSTANTS OF $Li_2B_4O_7$ CRYSTAL

Waldemar Soluch
Institute of Electronic Materials Technology
Wolczynska 133, 01-919 Warsaw, Poland

ABSTRACT

Elastic and piezoelectric constants of $Li_2B_4O_7$ crystal were measured by the pulse echo overlap and resonance methods. The following elastic constants c^E_{ij} [GPa] and piezoelectric constants e_{ij} [C/m^2] were obtained: c^E_{11} = 136.3, c^E_{12} = 2, c^E_{13} = 31.6, c^E_{33} = 53.2, c^E_{44} = 56.8, c^E_{66} = 47.8, e_{15} = 0.40, e_{31} = 0.23 and e_{33} = 0.90. The c^E_{12} constant was additionally corrected by comparing the calculated and measured surface acoustic wave velocities for the 45^0 rotated X cut, Z propagating orientation.

Key words: piezoelectric crystals, elastic constants, piezoelectric constants.

1. INTRODUCTION

Lithium tetrabotate ($Li_2B_4O_7$) belongs to the 4mm class of the tetragonal crystallographic system. It has six elastic, three piezoelectric and two dielectric constants. It was found, that there are large differences between some of the elastic and piezoelectric constans measured by different authors [1 - 3]. The crystal is a very interesting substrate material for small size surface acoustic wave (SAW) resonators and filters [4 - 5], and its parameters should be known with high accuracy. Therefore it was decided to remeasure the above constants using crystals grown at the Institute of Electronic Materials Technology.

2. PROCEDURE

In accordance with the IEEE recommendation, the X, Y and Z axes of the rectangular coordinate system are chosen in such a way, that X and Y axes are parallel to the crystallographic a_1 and a_2 axes, respectively (in the 4mm clas these exes are equivalent), and the Z axis is parallel to c axis (Fig.1)[6].

From the expressions for bulk acoustic wave (BAW) velocities propagating in different

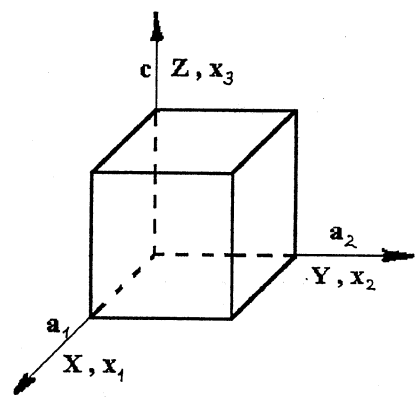


Fig.1. Rectangular coordinate systems.

orientations, it was found, that three types of cuboids and one type of plate are needed for determination all of the elastic and piezoelectric constants. The three types of cuboids are as follows: 1 - nonrotated orientation, 2- orientation rotated around the Z axis, 3 - orientation rotated around the X (or Y) axis. The angle of rotation was equal to 45^0 for both above cases. For each direction of propagation we have three types of BAWs: one longitudinal (or quasilongitudinal), and two shear (or quasishear). Therefore for each type of the cuboid we obtain 9 velocities, and together 27 velocities (Tables 1, 2 and 3). However not all of the velocities are different.
From an analysis of the analytical expressions for the velocities, it was found, that c^E_{33}, c^E_{13}, e_{13} and e_{33} constants are present only in three idependent equations. Therefore e_{33} should be determined from the Z-cut plate by the resonance method [6].

3. METHOD OF MEASUREMENT

To measure the BAW velocities, a pulse echo overlap (PEO) method was used [6-7]. The crystal sample was attached to the piezoelectric transducer (through a delay line) by

Table 1. Bulk acoustic wave velocities for the nonrotated case.

Type of wave	Direction of wave propagation and polarisation	Effective elastic constant	Velocity
L	x_1, x_1	c^E_{11}	v_1
S	x_1, x_2	c^E_{66}	v_2
S	x_1, x_3	$\bar{c}^E_{44}$	v_3
L	x_2, x_2	c^E_{11}	v_4
S	x_2, x_1	c^E_{66}	v_5
S	x_2, x_3	$\bar{c}^E_{44}$	v_6
L	x_3, x_3	$\bar{c}^E_{33}$	v_7
S	x_3, x_1	c^E_{44}	v_8
S	x_3, x_2	c^E_{44}	v_9

Table 2. Bulk acoustic waves for the 45^0 rotation around the Z axis.

Type of wave	Direction of wave propagation and polarisation	Effective elastic constant	Velocity
L	x_1, x_1	c'_{11}	v_{10}
S	x_1, x_2	c'_{66}	v_{11}
S	x_1, x_3	$\bar{c}_{44}$	v_{12}
L	x_2, x_2	c'_{11}	v_{13}
S	x_2, x_1	c'_{66}	v_{14}
S	x_2, x_3	$\bar{c}_{44}$	v_{15}
L	x_3, x_3	$\bar{c}_{33}$	v_{16}
S	x_3, x_1	c^E_{44}	v_{17}
S	x_3, x_2	c^E_{44}	v_{18}

Table 3. Bulk acoustic waves for the 45^0 rotation around the X (or Y) axis.

Type of wave	Direction of wave propagation and polarisation	Effective elastic constant	Velocity
L	x_1, x_1	c^E_{11}	v_{19}
S	$x_1, (x_2, x_3)$	c_{20}	v_{20}
S	$x_1, (x_2, x_3)$	c_{21}	v_{21}
S	x_2, x_1	c_{22}	v_{22}
QL	$x_2, (x_1, x_3)$	c_{23}	v_{23}
QS	$x_2, (x_1, x_3)$	c_{24}	v_{24}
S	x_3, x_1	c_{22}	v_{25}
QL	$x_3, (x_1, x_2)$	c_{23}	v_{26}
QS	$x_3, (x_1, x_2)$	c_{24}	v_{27}

where: L - longitudinal, S - shear, QL - quasilongitudinal, QS - quasishear,

$\bar{c}_{44} = c^E_{44} + e^2_{15}/\varepsilon^S_{11}$, $\bar{c}_{33} = c^E_{33} + e^2_{33}/\varepsilon^S_{33}$, $c'_{11} = (c^E_{11} + c^E_{12} + 2c^E_{66})/2$, $c'_{66} = (c^E_{11} - c^E_{12})/2$,

$c_{20} = (2c^E_{44} + e^2_{15}/\varepsilon^S_{11})/2$, $c_{21} = (2c^E_{66} + e^2_{15}/\varepsilon^S_{11})/2$, $c_{22} = (c^E_{44} + c^E_{66})/2$

$c_{23} = [\bar{c}'_{22} + \bar{c}'_{44} + \sqrt{(\bar{c}'_{44} - \bar{c}'_{22})^2 + 4(\bar{c}'_{24})^2}\,]/2$, $c_{24} = [\bar{c}'_{22} + \bar{c}'_{44} - \sqrt{(\bar{c}'_{44} - \bar{c}'_{22})^2 + 4(\bar{c}'_{24})^2}\,]/2$,

$\bar{c}'_{22} = c'_{22} + (e'_{22})^2/\varepsilon'_{22}$, $\bar{c}'_{44} = c'_{44} + (e'_{24})^2/\varepsilon'_{22}$, $\bar{c}'_{24} = c'_{24} + (e'_{24}e'_{22})/\varepsilon'_{22}$, $c'_{22} = (c^E_{11} + 2c^E_{13} + 4c^E_{44} + c^E_{33})/4$,

$c'_{24} = (c^E_{33} - c^E_{11})/4$, $c'_{44} = (c^E_{11} + c^E_{33} - 2c^E_{13})/4$, $e'_{22} = (2e_{15} + e_{31} + e_{33})/(2\sqrt{2})$, $e'_{24} = (e_{33} - e_{31})/(2\sqrt{2})$,

$\varepsilon'_{22} = (\varepsilon^S_{11} + \varepsilon^S_{33})/2$.

means of a coupling liquid. From the measured pulse repetition frequency f_r, the time delay between echos is determined as $t_d = 1/f_r$, and velocity $v = l_s/t_d$ where l_s is the sample length in the direction of wave propagation. The accuracy of this method is better then 0.1 %.

The piezoelectric constant e_{33} was evaluated from the measured resonance and antiresonance frequencies of the Z-cut plate [6].

4. RESULTS OF MEASUREMENT

Several samples of cuboids for each orientation, and Z-cut plates were prepared from the crystal grown by Czochralski method. The dimentions of the cuboids were equal to about 10 mm, and the thickness of the Z-cut plates, was about 0.5 mm.

The results of measurements are presented in Table 4.

It was found, that the calculated SAW velocity for the 45^0 rotated X cut, Z direction of propagation, is much more sensitive to the c^E_{12} variation than the X cut. The free surface wave velocity of the first cut was also measured [8]. Therefore the c^E_{12} constant was additionally corrected to obtain an agreement between the calculated and measured SAW velocities.

Table 4. Measured elastic and piezoelectric constants of $Li_2B_4O_7$ crystal.

c^E_{11}	c^E_{12}	c^E_{13}	c^E_{33}	c^E_{44}	c^E_{66}	e_{15}	e_{31}	e_{33}	Ref.
			[GPa]				[C/m^2]		
136.3	2	31.5	53.2	56.8	47.8	0.40	0.22	0.90	
126.1	0.5	30.0	53.9	55.0	46.0	0.36	0.19	0.89	[1]
135	3.6	33.5	56.8	58.5	46.7	0.47	0.29	0.93	[2]
135.8	- 0.3	32.0	54.2	57.1	47.7	0.39	0.25	0.88	[3]

Mass density: ρ = 2432 kg/m^3, dielectric constants: $\varepsilon^S_{11}/\varepsilon_0$ = 8.7; $\varepsilon^S_{33}/\varepsilon_0$ = 7.6 [3].

5. CONCLUSION

If we compare the values of the elastic and piezoelectric constants presented in Table 1, we see, that very large differences exist in the case of c^E_{12} constant. It is the result of an indirect calculation of the constant from the other ones. The value of the e_{15} constant, obtained in this work is in good agreement with the values obtained in [1], [3] and [9], while the value presented in [2] seems to high.

ACKOWLEDGEMENT

I would like to thank T.Lukasiewicz, J.Kisielewski, M.Swirkowicz and W.Hofman for crystals and samples fabrication, to H.Majewska for measurements and to M.Lysakowska for calculation of SAW velocities.

This work was supported by the State Committee for Scientific Research (KBN) under grant No. 8 8009 92 03.

REFERENCES

[1] N.M.Shorrocks, R.W.Whatmore, F.W.Ainger, I.M.Young , "Lithium tetraborate - a new temperature compensated piezoelectric substrate material for surface acoustic wave devices", 1981 IEEE Ultrasonics Symposium Proceedings, pp.337 - 340.

[2] M.Adachi, T.Shiosaki, H.Kobayashi, O.Ohnishi, A.Kawabata, "Temperature compensated piezoelectric lithium tetraborate crystal for high frequency surface acoustic wave and bulk wave device applications", 1985 IEEE Ultrasonics Symposium Proceedings, pp.228 - 232.

[3] J.A.Kosinski, Y.Lu, A.Ballato, "Pure-mode measurements of $Li_2B_4O_7$ material properties", IEEE Trans.on UFFC,vol.41,No.4,1994,pp.473 - 478.

[4] S.Matsumura, T.Omi, N.Yamaji, Y.Ebata, "A 45^0 X cut $Li_2B_4O_7$ single crystal substrate for SAW resonators", 1987 IEEE Ultrasonics Symposium Proceedings, pp.285 - 250.

[5] H.Abe, M.Ohmura, H.Saitou, "SAW devices on lithium tetraborate ($Li_2B_4O_7$)", 1994 IEEE International Frequency Control Symposium Proceedings, pp.289 - 295.

[6] IEEE Standard on Piezoelectricity, IEEE Trans., vol.SU-31, No.2, 1984.

[7] E.P.Papadakis, "Ultrasonic attenuation and velocity in three transformation products in steel", Journ. Appl. Phys., vol.35, No.5, 1964, pp.1474 - 1482.

[8] W.Soluch (to be published).

[9] W.Soluch, "Measurements of Bleustein-Gulyaev waves in $Li_2B_4O_7$ crystal", IEEE Trans. on UFFC, vol.42, No.5, 1995, pp.967 - 969.

CONVEX AND CONCAVE CORNER UNDERCUTTINGS IN THE MICROMACHINING OF QUARTZ AND SILICON MECHANICAL STRUCTURES

C.R. TELLIER, T. MESSAOUDI, T.G. LEBLOIS and S. DURAND*

ECOLE NATIONALE SUPERIEURE DE MECANIQUE ET DES MICROTECHNIQUES
Laboratoire de Chronométrie, Electronique et Piézoélectricité
26 Chemin de l'Epitaphe
25030 BESANCON CEDEX FRANCE

ABSTRACT

This paper is devoted to the convex and concave undercuttings encountered in the micromachining of crystals by a wet etching process. SEM micrographs of membrane and mesa micromachined on differently oriented silicon and quartz crystals are reported and analysed. 3D numerical and graphical simulations of etched structures are derived starting from the dissolution slowness surface of crystals. Emphasis is placed on theoretical etching shapes at square corners. The agreement between predicted and experimental shapes is rather good. In particular the 3D simulation gives realistic representations of bunchings at corners.
keywords : micromachining, chemical etching, silicon and quartz crystals, 3D simulation

1 - INTRODUCTION

In the past few years interest in the micromachining of crystals using an anisotropic wet etching process was revived [**1-4**]. This process is generally applied to quartz [**1, 4**] and silicon [**2, 3**] crystals in order to fabricate quartz micro-resonators [**1, 5**] or mechanical devices for quartz [**1, 5, 6**] and silicon [**7, 8**] sensors. In the case of silicon crystal most of applications concern (100) and (110) wafers. However for silicon sensors based on the piezoresistive effect (hhl) plates may be of interest [**7**] to improve the sensitivity of gauges or to design four-terminal sensors [**7, 8**]. Moreover it is also possible for quartz to define non-conventional singly or doubly rotated cuts which seem convenient for bulk waves resonators acting as thermal sensors.

In sensors applications a micro-fabrication process is generally required which passes through an anisotropic chemical etching. Such chemical etchings induce various technological problems [**2, 4**] such as a large convex undercutting (specially for silicon (hhl) plates [**4**]) or a significant lateral underetch (specially for some singly rotated quartz plates [**2**]). Development of sharp edges as well as formation of blocking facets were also observed for structures micromachined on quartz and silicon plates. Since these technological problems depend on the cut and on the direction of alignment for the inert mask it remains of prime importance to develop models which allow us to predict final etching shapes for any orientations of quartz or silicon plates and for any starting shapes of masks. The tensorial model for the anisotropic dissolution of crystals developped by C. Tellier [**9**] seems particularly suitable [**4, 10**] for the graphical and numerical simulation of 3D etching shapes encountered in micromachining. But up to now 3D simulations have been limited to structures obtained with starting circular masks. The present work deals essentially with the concave and convex undercutting at the corner of square masks.

2 - THEORETICAL CONSIDERATIONS

The 3D simulation [**4, 11**] is based on a kinematic model [**12**] for the anisotropic chemical etching. In the first stage of the simulation we calculate the displacement $\vec{P}$ of all surface elements potentially present under the mask from the analytical expression for the dissolution slowness surface as derived from a tensorial analysis.The degree of divergence for the displacements $\vec{P}$ of surface elements close together determine the final etching shape. It is at this point essential to remark that for initially convex and concave 2D profiles the displacements of successive elements are of converse nature. Consequently if for a concave profile displacements $\vec{P}_1$ and $\vec{P}_2$ (Figure 1) converge they, according to Irving's criteria [**13**] necessarily diverge for the convex profile involving similarly oriented profile elements

Now let us consider the case of a mesa structure and let us apply the Irving's criteria on the one hand to the top contour of the inert mask and on the other hand, to cross-sectional dissolution profiles under the mask. We readily notice that we work with convex and concave 2D shapes respectively. Hence it becomes difficult to predict the final etching shape by using simply the Irving's criteria which govern the displacements $\vec{P}$ of elements and which firstly, emphasize the role of extrema for the dissolution slowness $\vec{L}$, in the divergence of successive $\vec{P}$ and secondly state that for a convex (concave) shape minima (maxima) in $\vec{L}$ determine the final etching shape. However when we start with circular mask of finite radius it is usual [**4, 10**] to assume that maxima in the dissolution slowness play the most important role that is to say that concave intersections predominate. But

this assumption seems to be more questionable as soon as we are concerned with a square corner because all surface elements propagate from the same corner point 0.

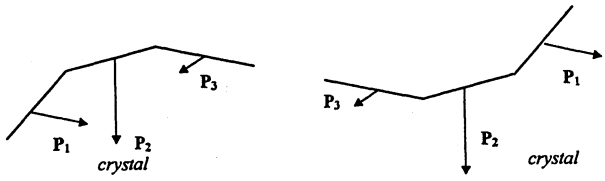


Figure 1 : Trajectories for concave and convex intersections

In this paper an attempt is made to present preliminary results on the concave and convex undercuttings at square corners in order to verify the adequation of the previous assumption on which numerical tests are based in the 3D simulation. Since the adjustment of the dissolution constants from data affects the agreement between theoretical and experimental shapes, investigations on two crystals namely silicon and quartz are reported. Effectively we have recently shown that the adequation of the proposed dissolution slowness surface for the silicon crystal (etchant : KOH 35%) is rather good [**10**].

3 - EXPERIMENTAL RESULTS

3.1 - Experimental details

(100), (hk0) and (hhl) silicon plates were cut in a silicon ingot and then mechanically polished. SiO_2-Si_3N_4 inert masks for membranes and mesa were patterned on these plates using conventional photolithographic techniques. Plates were immersed in a 35% KOH solution maintened at a constant temperature (65°C). A similar photolithographic procedure was applied to singly and doubly rotated quartz plates except that a Cr-Au thin film (≈ 0,3 µm) served as an inert mask. To prevent etch pitting Cr-Au evaporated films suffered a specific thermal recovery process. A concentrated $NH_4F.HF$ solution was used for the micromachining of structures on quartz plates.

3.2 - Micromachining of silicon structures

Figures 2 and 3 give SEM micrographs of some typical structures micromachined on a (110) and a (441) silicon plates. On these micrographs the $[1\bar{1}0]$ direction lies respectively along the horizontal and along a direction which makes an angle of 15° with the vertical axis as indicated by the arrow. In particular that for the (110) plate un-masked regions are covered by a characteristic surface texture [**2**] composed of grooves aligned along the $[1\bar{1}0]$ direction.

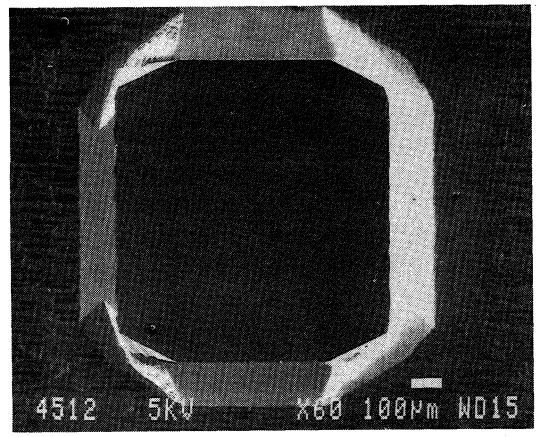


Figure 2 : A mesa etched on a (110) Si plate

For singly (hk0) and doubly (hhl) rotated silicon plates we observe that in general marked convex undercuttings develop at mesa corners even if for some doubly rotated plates undercuttings seem less pronounced than for (100) and (110) mesa. Far from corners slightly curved or flat facets denoted "f" bound mesa and membranes. In particular along <110> directions flat facets are without ambiguity associated with limiting {111} planes which etch very slowly. Frequently at membrane corners facets intersect along straight lines (see for example corners 1 and 4 on Figure 3).

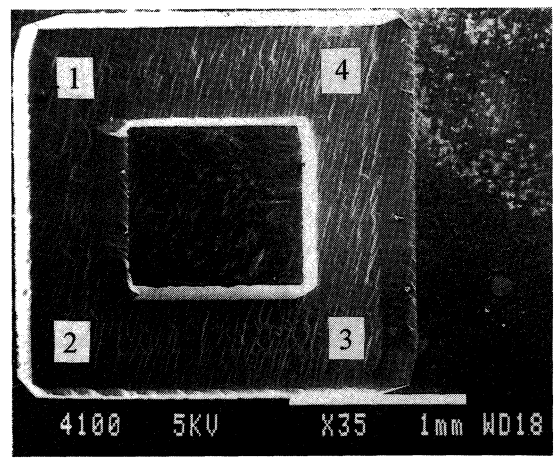


Figure 3 : A hole-mesa structure on a (441) Si plate

3.3 - Micromachining of quartz structures

Figures 4 and 5 illustrate some etching shapes obtained on a singly ($\varphi = 0°$, $\theta = 65°$) and a doubly ($\varphi = 90°$, $\theta = 70°$) quartz plate called for convenience AT+65

and Y+70 cuts respectively. On the SEM micrographs we mention a moderate pitting of the Cr - Au mask together with an unsticking of the mask edge in some regions. Nevertheless some of the observations made for silicon structures remain valids such as formation of somewhat flat facets at membrane corners. Most of these facets intersect along successive straight line segments (see for example corner 2 on Figure 4).

It is important to remark that for quartz mesa, corners are less affected by the convex undercutting than for silicon mesa. This behavior which is typical of quartz micromachining and which is characterized by a moderate bunching of kinematic dissolution "waves" will be discussed in the following section.

4 - 3D SIMULATIONS OF CONCAVE AND CONVEX UNDERCUTTINGS

4.1 - Silicon structures

Theoretical undercuttings at corners of (100) and (110) mesa, as derived from the 3D simulation, are given in Figures 6 and 7 respectively. In the case of (100) mesa we recognize the well-known undercutting resulting from the micromachining of a "square" mesa aligned along <110> directions by a 35% KOH solution [14]. Such a micromachining results in the formation of two {hll} facets which intersect the upper surface of the mesa along two segments which make an intersection angle δ. A value of about 127° was measured for δ in most of the experimental works [14]. This value is not far from the predicted value (about 131°, Figure 6).

Figures 7 indicates that for a (110) mesa the simulation leads to a convex corner undercutting characterized by two upper intersecting segments with an angle δ of about 134° in good agreement with the experimental angle. In theory one of the facet which develops under the mask at the square corner joints the {111} limiting plane (facet f_1, Figure 7) by forming a curved intersecting line in close accord with the SEM observations. But for the facet f_2 which bounds the other square side some slight departures can be depicted. Effectively in practice corner bunchings intersect this facet f_2 along two straight line segments which are more accentuated than predicted segments. Moreover the predicted cross-sectional dissolution profile (Figure 7) under the mask is composed of two lines with different slopes. This feature is conveniently illustrated by the SEM micrograph of Figure 2 where we distinguish easily two regions with different inclinaisons. Thus for the (110) mesa the impression of rather good agreement between data and predictions predominates.

Figure 8 gives theoretical convex and concave undercuttings at mesa and membrane corners on a (441) plate. Let us firstly consider the convex corner undercutting. Now the agreement between theory and experiments appears to be very good. The agreement covers the following points :
(i) the bunching at corners with in particular the presence of curved discontinuities between {111} facets and regions which cut corners (see for example corners 1 and 2)..
(ii) the geometrical arrangement of bunched facets at corners. We can easily depict this accord on the SEM image (Figure 2) in the case of corners 1 and 2.
(iii) the undercutting generating a sharp edge with an acute angle at corner 4.

Turning now our attention to some concave undercuttings (Figure 8) here again we observe no departure. Inclinations and extents of facets which bound the four sides of the membrane are similar. General shape of concave corners as well as directions of intersecting straight lines associated with limiting facets are in close agreement.

4.2 - Quartz structures

Some results related to an AT+65 and an Y+70 quartz plates are reported in Figures 9 and 10. Theoretical shapes for the AT+65 mesa do not deviate markedly from experimental shapes. Effectively on Figure 9 a facet is present at corner 1 whereas sharp edges are formed under corners 2 and 3. Comparison of Figures 4 and 9 reveals only a small deviation between theoretical and experimental extents of the curved region CR. The agreement for concave undercuttings seems also rather good and the reader can easily detect the two straight line segments which compose the intersection of limiting facets at corner 2.

In addition some theoretical diagrams related to an Y+70 mesa-membrane structure are drawn in Figure 10. Here again we only observe very small departures. The most important features which characterize experimental convex undercuttings such as formation of facets at corners 1 and 4 and the presence of sharp edges at corners 2 and 3 can be distinguished on theoretical undercuttings. It should be pointed out that simulations of concave undercuttings (Figure 10) appear to be also satisfactory.

5 - CONCLUSION

In conclusion the 3D simulation allows us to propose theoretical diagrams of concave and convex undercuttings at square corners which give satisfactory representations of experimental shapes. In particular bunchings at convex corners can be conveniently derived from the 3D simulation. Since theoretical undercuttings of converse nature have been obtained by assuming that maxima in $\vec{L}$ determine preferentially all final etching shapes. This assumption on which are based numerical tests seems reasonable even for trajectories of surface elements at convex intersections.

REFERENCES

1 - **J.S. DANEL, F. MICHEL and G. DELAPIERRE,** Sensors and Actuators, A21-23 (1990) 971-978.

2 - **C.R. TELLIER and A. BRAHIM-BOUNAB,** *ibid*, 29 (1994) 5953-5971.

3 - **K.E. PETERSEN,** Proc. IEEE, 70 (1982) 420-427.

4 - **C.R. TELLIER, T. MESSAOUDI and T. LEBLOIS,** *Proc. 10 th European Frequency and Time Forum*, Brighton, UK, 5-7 march 1996, (IEE, 1996) pp 364-369.

5 - **B. STUDER, W. ZINGG,** *ibid*, pp 14-20.

6 - **J. SODERKVIST,** Sensors and Actuators, A43 (1994) 65-71.

7 - **S. DURAND AND C.R. TELLIER,** J. Phys. III, 6 (1996) 237-266.

8 - **Y. KANDA,** Sensors and Actuators, A28 (1991) 83-91.

9 - **C. R. TELLIER and J. L. VATERKOWSKI,** J. Mater. Sci., 24 (1989) 1077-1088.

10 - **C.R. TELLIER and S. DURAND,** Sensors and Actuators (1997), to be published.

11 - **C.R. TELLIER and D. BENMESSAOUDA,** *Proc. of the 8 th European and Time Forum*, Munich, Germany, March 1994 (VDI PROJEKT, 1994) pp 245-255

12 - **K. SANGWAL,** *Etching of Crystals*, (North-Holland, Amsterdam, 1987) chap. 4.

13 - **B.A. IRVING** in *"The Electrochemistry of Semiconductors"* edited by P.J. Holmes (Academic Press, London, 1962) pp 256-289.

*** Actually with the University du Maine (Le Mans, France)**

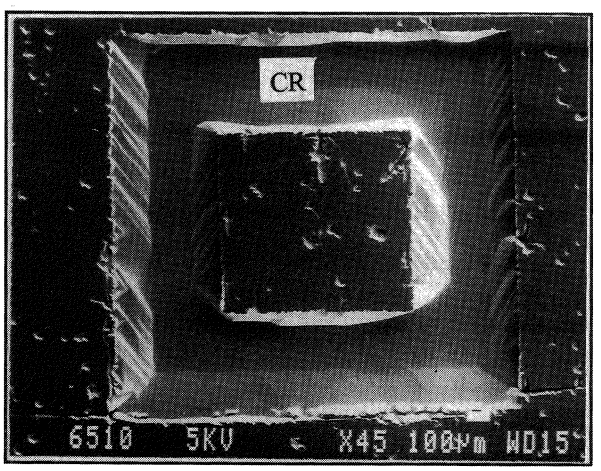


Figure 4: Hole-mesa structure on an AT+65 quartz plate

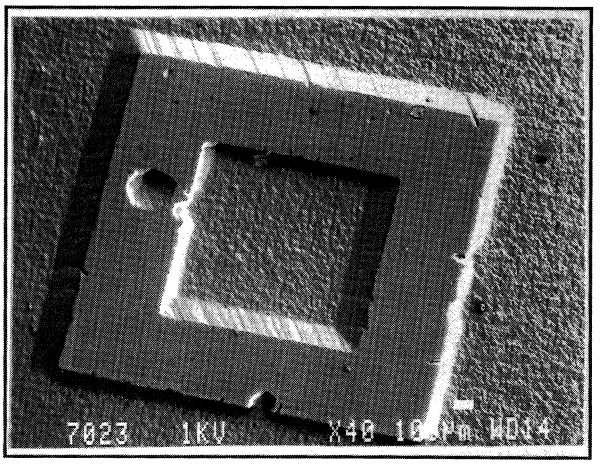


Figure 5: Mesa-hole structure on a Y+70 quartz plate

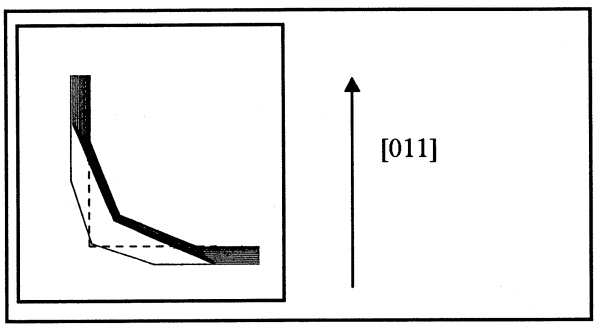


Figure 6: Predicted (100) convex undercutting

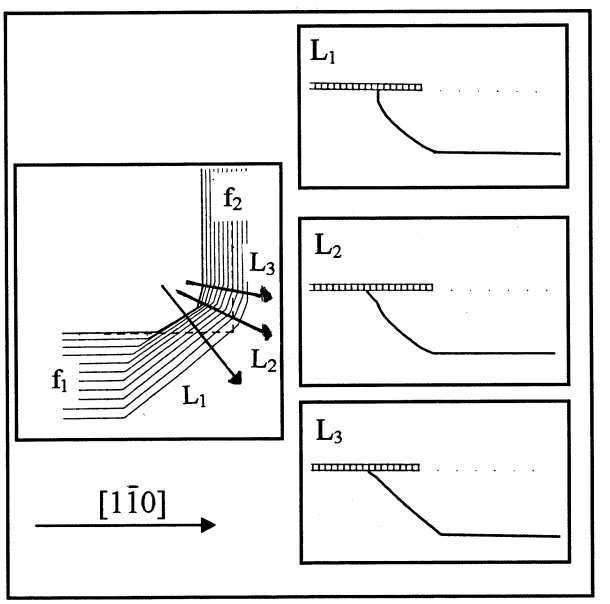


Figure 7: Theoretical (110) convex undercutting and cross-sectional profiles as indicated on diagram.

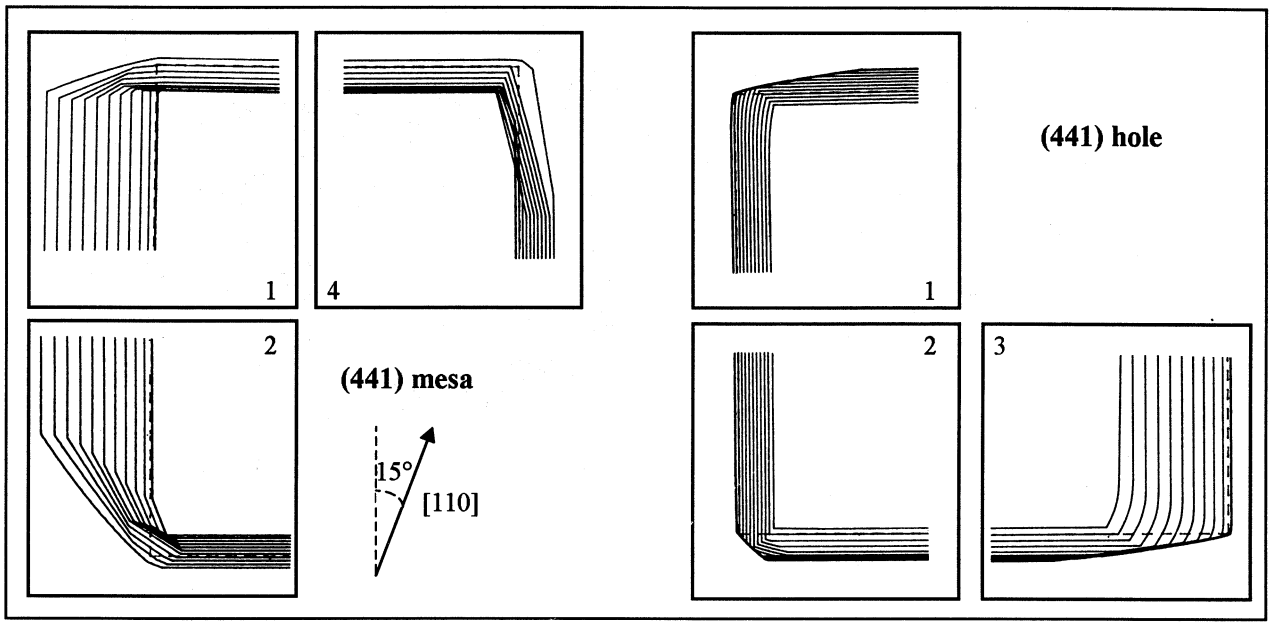


Figure 8:Theoretical undercuttings for (441) mesa and membrane

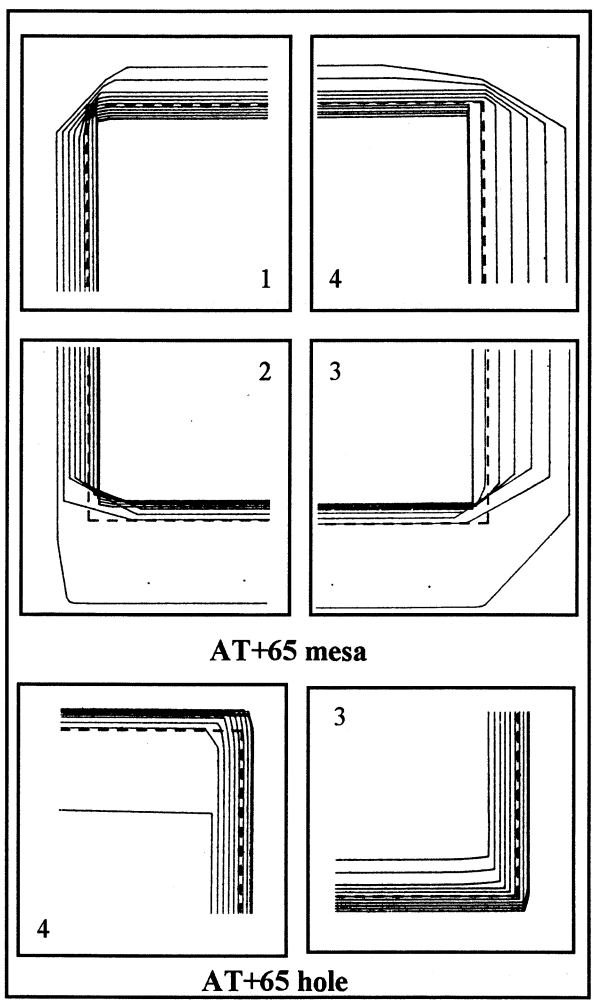


Figure 9: Theoretical undercuttings for AT+65 mesa and membrane

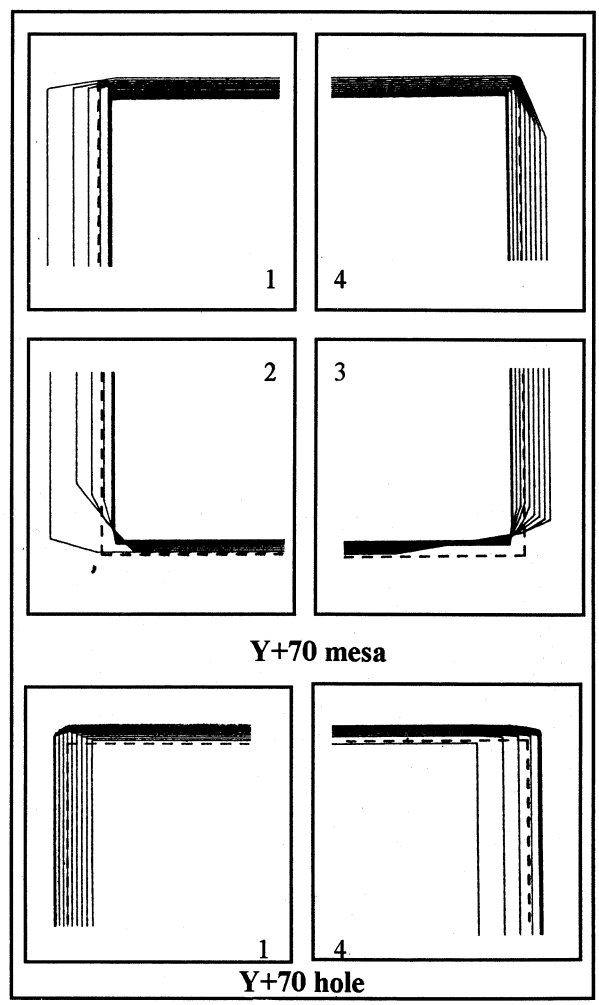


Figure 10: Theoretical undercuttings for Y+70 mesa and membrane

HIGH STABILITY RESONATOR-THERMOSTAT PACKAGED IN TO-8 VACUUM HOLDER

Igor Abramzon, Consultant, Roman Boroditsky, Valpey-Fisher Co., MA, USA

ABSTRACT

This work presents technology and test results of the Directly Heated Resonator (DHR) design, packaged in TO-8 vacuum holder. The DHR is comprised of a 10 MHz 3rd overtone SC-cut crystal plate with film heaters and thermistor, as well as a thermocontroller circuit assembled on a ceramic substrate inside the package. Very small size is accompanied by the following performance: stability vs. temperature - better than ±5 x 10-8 over -30 to +70°C range; warm-up time - 15 s for 1 x 10-7 accuracy; aging rate is about 5 x 10-10/day after 15 days of operation. OCXOs based on this resonator can be an attractive option for various applications including mobile communications, portable GPS receivers, rescue systems, etc..

INTRODUCTION

Current requirements for mobile communication and navigation devices demand oscillators with high frequency stability, which maintain low power consumption, small size, and fast warm-up time. Since TCXO performance is no longer satisfactory for a number of applications, designers concentrate on the creation of OCXOs, which are compatible in size, power consumption and warm-up time with TCXO. One of the promising technologies is the Directly Heated Resonator (DHR). It provides the unique combination of high performance in frequency stability, power, size, and warm-up time [1-5]. The goal of present work was to accommodate DHR technique in a small resonator package. We chose a TO-8 cold-weld vacuum holder for this purpose. It provides suitable housing for a 3rd overtone SC-cut crystal, sufficient number of isolated leads, and room for thermocontroller circuit.

It was shown in previous works, that utilization of power dissipated in heating transistor considerably improves steady state power consumption and frequency stability of the OCXO [3-5]. Further progress in DHR parameters can be accomplished by incorporating all of the thermocontroller circuitry into resonator package. The target of the work was to create an optimal resonator design in small volume providing lowest thermal losses and sound mechanical structure.

Another problem we faced was to attain sufficiently high vacuum level as a prerequisite of superior DHR performance. To achieve this goal we significantly modified the process of evacuating the TO-8 package.

RT DESIGN AND PROCESSES

A DHR design is outlined in fig.1. It contains a ceramic substrate, supporting an SC-cut crystal blank with deposited thin film electrodes and heaters. A miniature thermistor is also attached to the crystal surface to sense its temperature. Thermocontroller circuitry is assembled on the substrate using thick film hybrid technology. The entire assembly is placed on the TO-8 header and sealed with 0.32" high cover.

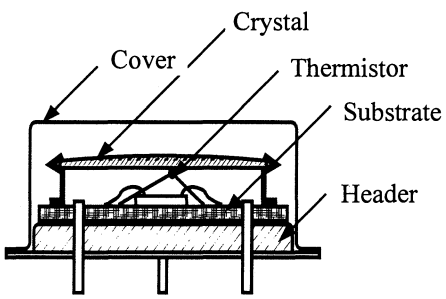


Fig.1.

The resonator has a pair of leads for input voltage, and another pair to connect the crystal to the external oscillator network. As soon as input voltage is applied to the device the thermocontroller raises the crystal temperature to the turnover point and maintains it with accuracy of ±0.5°C over the entire ambient temperature range. The higher the supply voltage the better the accuracy of temperature control. Since there is no additional means to control the resonator temperature outside the unit, the DHR is essentially a Resonator-Thermostat (RT).

This combination of directly heating the crystal plate with utilization of the power dissipated in the thermocontroller yields an excellent set of RT parameters. Very short warm-up times (<10s) can be achieved by varying combinations input voltage and heater resistance (during the design process). To minimize transient overshoot in the fast warm-up process,

we used SC-cut crystal with optimized blank geometry and heater configuration . In steady state operation most of the power dissipates in the thermocontroller circuitry. This leads to insignificant temperature gradients in the crystal plate and results in good temperature stability of the RT.

The advantages of the RT design can be attained only by properly minimizing its thermal losses. This was accomplished by optimizing the substrate configuration and achieving a high vacuum level inside the TO-8 housing. A modified cold-weld process resulted in achieving a vacuum level of less than 0.05 Pa. That resulted in higher than 600 K/W thermal resistance of the design.

RT PROTOTYPES TEST RESULTS

Some essential parameters of the developed RT are shown in Fig. 2 - 5.

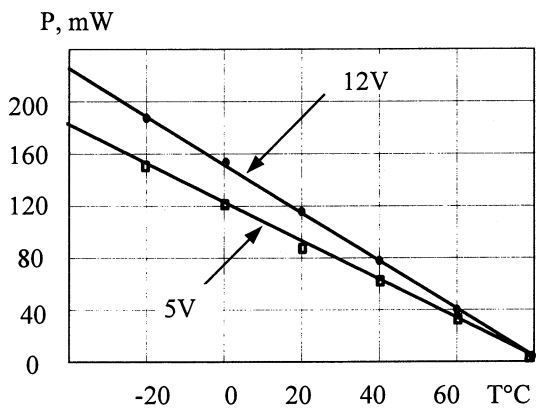


Fig.2

Power consumption, as shown on Fig.2. ranges from 30 mW to 240 mW in the temperature range of -30°C to 70°C. The values vary insignificantly within the batch of units, as a proof of process reliability. Power consumption can be lowered somewhat by increasing the resistance of the heaters. It would lead however to the degradation of the frequency stability vs. temperature performance.

As one can see from Fig.3. the frequency stability versus temperature is about ± 3E-8 over -30°C to 70°C range with 12 V supply, and slightly worse with 5 V supply. This parameter depends on heater resistance and can be further improved at the expense of power consumption of the RT.

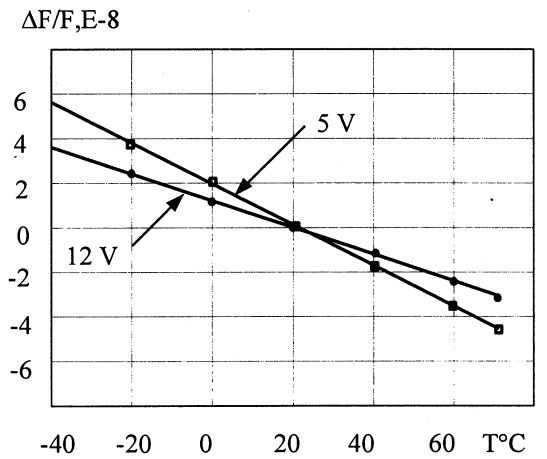


Fig.3.

Warm-up time of the RT was measured with different start-up power by varying heater resistance and supply voltage (Fig.4.). In spite of extremely fast heating process (5 to 10 °C per second) we did not observe any significant dynamic overshoots of frequency. Optimal crystal plate configuration allows for 10 s (12 V operation) to 30 s (5 V operation) frequency settling time, which is comparable to most TCXOs.

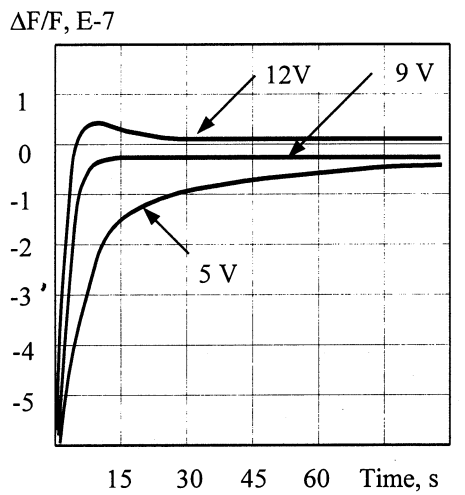


Fig.4.

Only preliminary results of the RT aging are available now. Fig.5. shows that aging slope of 5E-10 per day is reached after 15 -20 days of operation at normal conditions. Aging is expected to be 1 to 1.5E-7 for the first year. Achieving the best possible long term frequency stability will

require careful processing and design optimization.

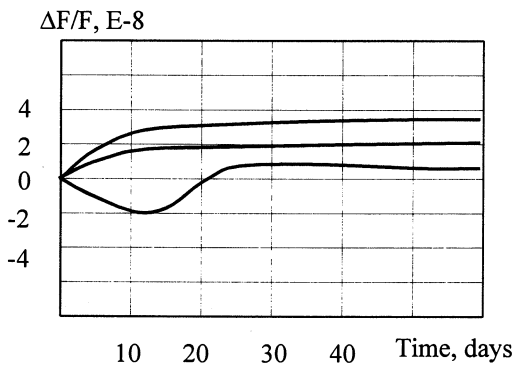


Fig.5. Typical aging curves.

We built miniature oscillators based on the developed RT. Their performance in comparison with typical modern OCXO and TCXO is shown below

Parameter	Oscillator type		
	RT	Small OCXO	Best TCXO
Frequency Stability			
over temp.,E-7	0.5	0.5	5
per 1 s, E-10	<1	<1	<10
per year, E-7	1-1.5	1	5
Power Consumption at 25°C, mW	150	500	80
Warm-up Time to 1E-7, s	15	120	10
Volume, cm^3	3	15	3

This data shows that oscillators built with the RT are similar in performance to small OCXO, but its size, power consumption and warm-up time is comparable to high stability TCXO.

CONCLUSIONS

Incorporating temperature control and stabilization circuitry inside the RT package is a good base for building a high stability oscillator whose size is compatible with TCXOs. Such an oscillator is the most attractive option for high quality mobile communications, GPS and rescue systems where TCXO performance does not satisfy frequency stability requirements. At the same time conventional OCXOs are not suitable due to large size, power consumption and long warm-up time.

The technology developed in this work yielded miniature Resonator-Thermostat, possessing an excellent combination of high frequency stability, low power consumption, and extra short warm-up time.

REFERENCES.

1. B.Long, G Weaver, "Quartz crystal oscillators with direct resonator heating", Proceeding of the 45 Annual Frequency Control Symposium, pp. 384 - 392, 1991

2. S.Galliou, M. Mourey and B. Dumlet, "A single bridge resonator", Proceeding of 8th EFTF, pp. 792 - 800, 1994.

3. I. Abramzon, A. Dikidzhi, "Improvement of characteristics of quartz resonator-thermostats with direct heating piezoelement", Proceedings of 1992 IEEE Frequency Control Symposium, pp. 499-504, 1992.

4. I. Abramzon, "OCXO design using composite-heating of the crystal resonator", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol.41, No.2, pp.284 - 287, 1994.

5. I. Abramzon, "Internal heated quartz resonators with low sensitivity to acceleration", Proceedings of 1995 IEEE Frequency Control Symposium, pp. 838 - 842, 1995.

INCREASING THE LOADED Q OF GHz RANGE SURFACE TRANSVERSE WAVE RESONATORS

Ivan D. Avramov

Institute of Solid State Physics
72 Tzarigradsko Chaussee Blvd.
1784 Sofia, Bulgaria
e-mail: iavramov@phys.acad.bg

1. ABSTRACT

This paper presents data from a systematic study on increasing the loaded Q of 1.244 GHz two-port surface transverse wave (STW) resonators, featuring an insertion loss in the 4-5 dB range, by altering the capacitive coupling of the acoustic device to the load. It is shown that, if this well known method is applied to STW devices, the loaded group delay can be doubled while the insertion loss increases only by 5 to 6 dB. A substantial increase of the device Q was achieved when the same devices were used as single-port resonators. They demonstrate typical values for the dynamic resistance in the 10-15 Ω range and the capacitance ratio exceeds 2000. Most of the devices were found to have a Q in the 15000-18000 range at 1.244 GHz. Such devices were found to be well suited for the design of miniature Pierce-type fixed frequency oscillators operating at 1.244 and 2.1 GHz and featuring a short term stability and voltage pushing sensitivity by a factor of 2-4 better than feedback oscillators using the same devices as 2-port resonators.

Keywords: STW resonators, loaded Q, Pierce-type oscillator.

2. INTRODUCTION

Recent advances in STW resonator technology in the 1 to 3 GHz range have allowed engineers to implement the unique features of STW in state-of-the-art oscillators with extremely low 1/f noise, high output power and superior RF/d.c. efficiency (Ref. 1-3). Recently, it was shown with experimental devices that the material Q for STW exceeds the material limit for surface acoustic waves (SAW) by at least 37% (Ref. 4). This has made it possible to design and build high-performance STW resonators in the 2.0-2.5 GHz range featuring a loaded Q in the 4000-6500 range and an insertion loss in the 8-17 dB range (Ref. 5). Thus the STW resonator becomes a serious competitor of the dielectric resonator and makes it possible to design miniature, low-cost fixed frequency and voltage controlled oscillators with excellent phase noise performance, high temperature stability and low vibration sensitivity in the lower GHz range.

In most oscillator applications close-to-carrier phase noise is an important design parameter. Since, at frequencies above 1 GHz, STW resonators have generally a lower 1/f noise than the active circuit in the oscillator loop (Ref. 2), the only way to improve the close-to-carrier phase noise is to increase the loaded Q (Q_L) of the acoustic device. In this case the close-to-carrier oscillator phase noise decreases by the factor $(1/Q_L)^2$ (Ref. 6). Unfortunately, it is difficult to design a SAW or STW two-port resonator with a desired loaded Q in a 50 Ω environment as typically necessary for feedback oscillator applications. However, a substantial increase in the loaded Q can be achieved by reducing the coupling of the resonator ports to the load, a technique widely used in dielectric resonator oscillators. A further improvement of the STW resonator Q can be achieved by implementing the acoustic device in a single-port resonator configuration. This paper illustrates both methods applied to 1.244 GHz STW resonators and presents experimental data on the device Q, insertion loss, change in resonant frequency and device input and output impedance as well as sidelobe suppression with coupling to the 2-port device. Furthermore, the increase in Q is illustrated by comparing the performance of Pierce type oscillators, using single-port devices, with feedback oscillators using the same devices as two-port resonators.

3. INCREASING THE LOADED Q OF 2-PORT STW RESONATORS BY REDUCING THE COUPLING TO THE LOAD

The experiments were performed with a 1.244 GHz two-port STW resonator similar to the one described in Ref. 7. This design has demonstrated a power handling ability in excess of 2W and an extremely low residual phase noise level in the -142 to -144 dBc/Hz range at 1 GHz (Ref. 2). The frequency and group delay responses of the 1.244 GHz device, used in this experimental study and measured in a 50 Ω system without coupling, are shown in Fig. 1.

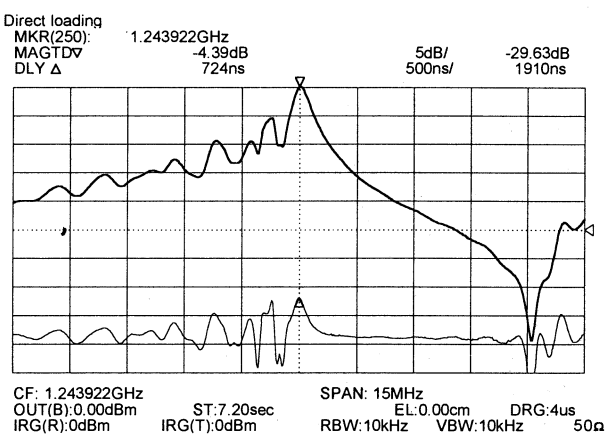


Fig. 1. Frequency response (upper curve) and group delay response (lower curve) of the device used in the experiments.

Using the relations:

$$Q_L = \pi \tau_g f_0 \quad (1)$$

$$Q_U = Q_L / [1 - 10^{(-IL/20)}] \quad (2)$$

where τ_g is the device loaded group delay, f_0 is its resonant frequency, IL is its insertion loss and Q_U is its unloaded Q, the Q_L and Q_U values of the device from Fig. 1 are calculated as 2830 and 7130 respectively. From (1) and (2) it is obvious that Q_L and Q_U are related to each other via the device insertion loss which, on the other hand, will depend on the coupling to the external load. In our case, direct coupling to the 50 Ω load of the measurement system results in an insertion loss of 4.39 dB (see Fig. 1), and a loaded Q value by a factor of 2.5 smaller than the unloaded Q. If, on the other hand, the coupling to the load becomes weak and the insertion loss increases, then Q_L approaches the Q_U value. For practical applications, requiring high loaded Q, it is reasonable to stop reducing the coupling before the insertion loss becomes too high. Here a coupling variation was performed by connecting two capacitors C_k of equal values in series to the input and output interdigital transducer (IDT) as shown in Fig. 2. The effect of these coupling capacitors on the device frequency

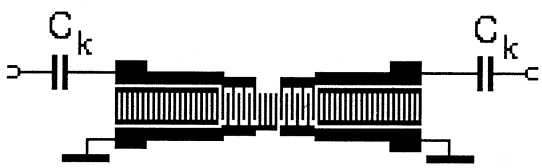


Fig. 2. Reducing the coupling to the load using series capacitors.

and group delay responses is shown in Fig. 3 for a C_k value of 1.5 pF. It is evident that the reduction of the load coupling for this capacitor value nearly doubles the loaded group delay, increases the insertion loss by 5.4 dB and improves the sidelobe suppression by about 5 dB compared to the case in which the resonator is directly connected to the 2×50Ω load of the measurement system (see Fig. 1).

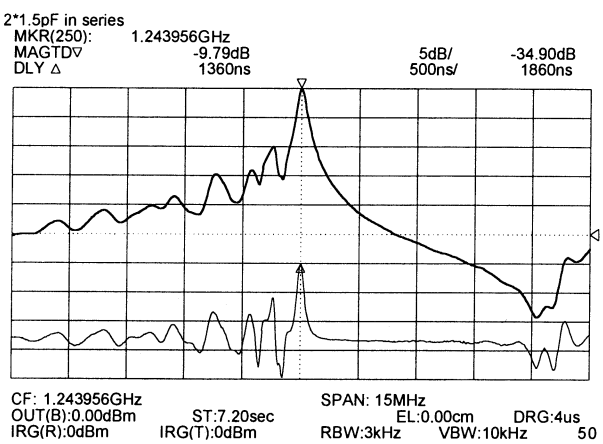


Fig. 3. Effect of 2×1.5 pF coupling capacitors on the resonator's frequency and group delay responses.

Another effect of the capacitive load coupling is the increase of the device input and output impedances which is evident from the comparison of the parameters S11 with and without coupling capacitors. A similar result was obtained also from the comparison of the parameters S22.

Finally, the capacitive load coupling via 2×1.5 pF moves slightly the resonant frequency by 24 ppm higher.

To test the feasibility of this technique for increasing the loaded Q of STW 2-port resonators, the electrical characteristics of the device from Fig. 1 were evaluated at different values of the coupling capacitors C_k starting from 3.4 pF and ending with 0.4 pF. The results are shown in Fig. 5, 6, 7 and 8. As evident from Fig. 5 and 6 both the loaded Q and the insertion loss increase exponentially with reduced coupling. The unloaded Q, calculated with equation (2), can

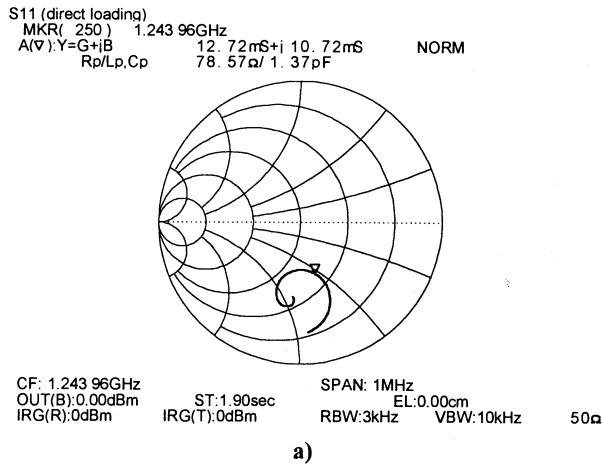


Fig. 4. Input reflection coefficient S11 of the 1.244 GHz device: a) direct load coupling; b) coupling via 2×1.5 pF.

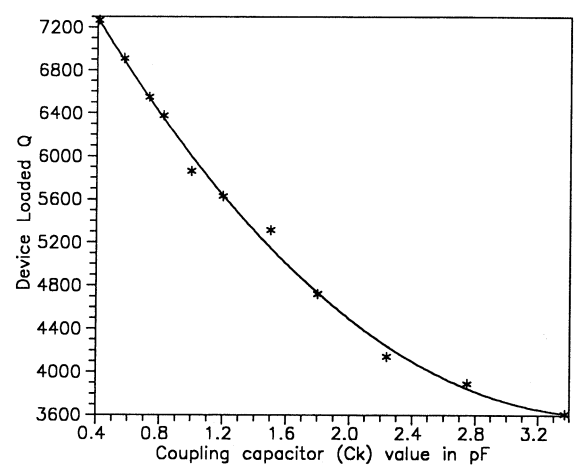


Fig. 5. Device loaded Q versus coupling capacitors' value.

be achieved only at the expense of greatly increased insertion loss (more than 23 dB in this case). On the other hand, 80% of the unloaded Q value (C_k=1.2 pF), are obtained at an insertion loss of only 11 dB which is quite acceptable for most

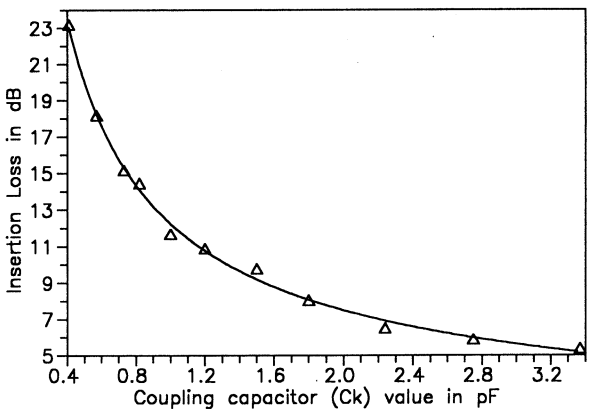


Fig. 6. Insertion loss versus coupling capacitors' value.

applications. In most cases the loaded Q is adjusted to be as high as 50% of the unloaded Q. This is obviously achieved with less than 1 dB increase in insertion loss (see Fig. 1 and compare with the insertion loss value at C_k = 3.3 pF).

Fig. 7 indicates an insignificant shift of the resonant frequency when the coupling becomes very small. This freq-

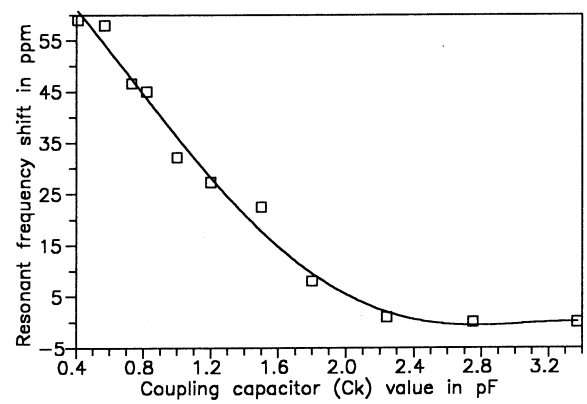


Fig. 7. Resonant frequency shift versus C_k value.

uency shift does not exceed 30 ppm for a loaded Q as high as 80% of the unloaded Q value.

Finally, Fig. 8 illustrates the behavior of the sidelobe

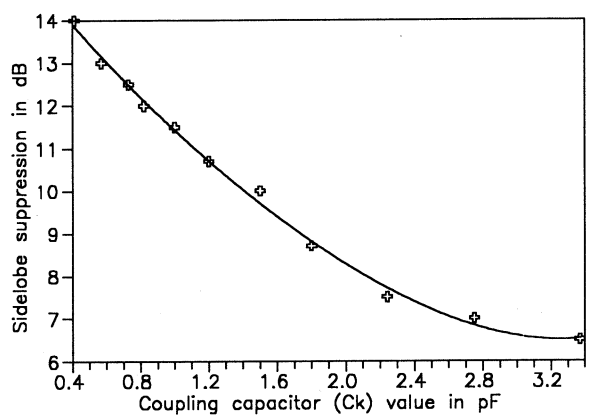


Fig. 8. Sidelobe suppression versus C_k value.

suppression with coupling. A substantial improvement of the sidelobe level occurs at weak coupling. At Q_L=80%Q_U and Q_L=50%Q_U the sidelobe level is suppressed by about 11 and 7 dB respectively. These sidelobe levels are fully sufficient for oscillator applications.

4. THE Q OF A TWO-PORT STW RESONATOR USED AS A SINGLE-PORT DEVICE

A further way of improving the resonant Q of the STW resonator from Fig. 1 is to use it as a single-port device. In this case only one of the IDT is used and the other one is left open. The equivalent electric circuit of this device is the same as of a bulk crystal resonator. It is shown in Fig. 9.

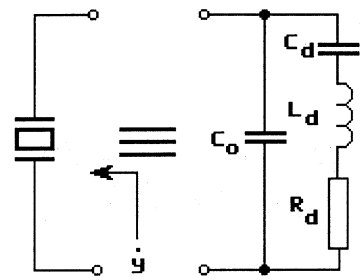


Fig. 9. Equivalent electrical circuit of the single-port device.

It is well known that a device like this is characterized by a series and parallel resonant frequency f_s and f_p. If these two frequencies and the static IDT capacitance C_o are measured, then, according to Ref. 8, the dynamic capacitance C_d can be calculated using the following relationship (Ref. 8):

$$C_d = \frac{2C_0(f_p - f_s)}{f_s} \tag{3}$$

Then the device Q can be calculated as:

$$Q = \frac{1}{2\pi f_s C_d R_d} \tag{4}$$

The series and parallel resonant frequencies of the device from Fig. 1, used as a single-port resonator were measured in a Π-type circuit embedded in a low-feedthrough fixture which was found to provide a feedthrough suppression in excess of 60 dB at 1.244 GHz. The schematic of the Π-type circuit is shown in Fig. 10.

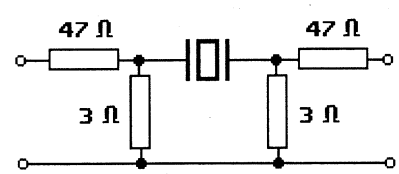


Fig. 10. Schematic of the Π-type circuit for measuring the series and parallel resonant frequencies.

The broadband and narrowband frequency and group delay responses of this device, obtained in a transmission measurement in the Π-type circuit are shown in Fig. 11 a) and b) respectively. On the fairly strong frequency response the

series and parallel resonant frequencies are clearly visible. They are located at the places of minimum and maximum insertion loss respectively. The static IDT capacitance was measured as 1.62 pF with a capacitance meter. As evident from Fig. 11 b) the distance between f_s and f_p is 290 kHz which, according to (3) corresponds to a capacitance ratio C_o/C_d of about 2140. This results in a Cd value of 0.76 fF.

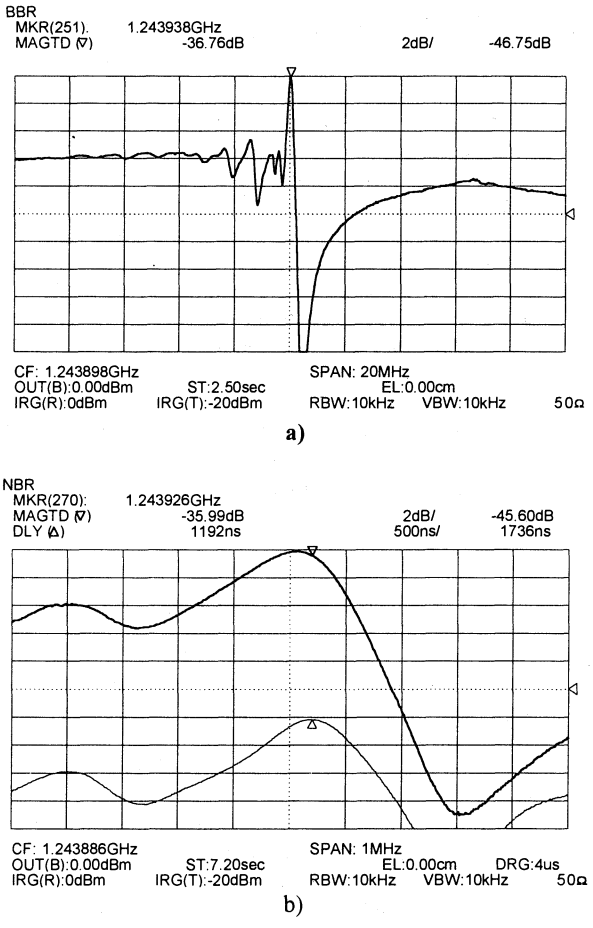


Fig. 11. Results from the transmission measurement on the single-port device in the Π-type circuit:
a) broadband frequency response,
b) narrowband frequency and group delay responses.

Finally, the dynamic resistance R_d was measured in a reflection measurement after tuning out the static IDT capacitance with a parallel inductor as shown in Fig. 12.

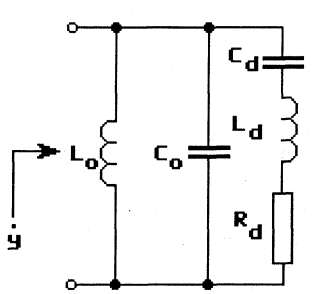


Fig. 12. Equivalent circuit of the single-port device with C_o tuned out. Setup for R_d measurement.

The result of the reflection measurement is shown in Fig. 13. It is evident that the dynamic resistance reads the value of 9.38 Ω which is obtained at the location at which the admittance plot crosses the real axis of the Smith chart. Substituting the R_d and C_d values, obtained from these two measurements in (4), for the device Q of the single-port device we obtain:

$$Q_{siport} = 17450$$

This value is by a factor of 2.5 higher than the unloaded Q of this same device characterized as a two-port resonator in Fig. 1.

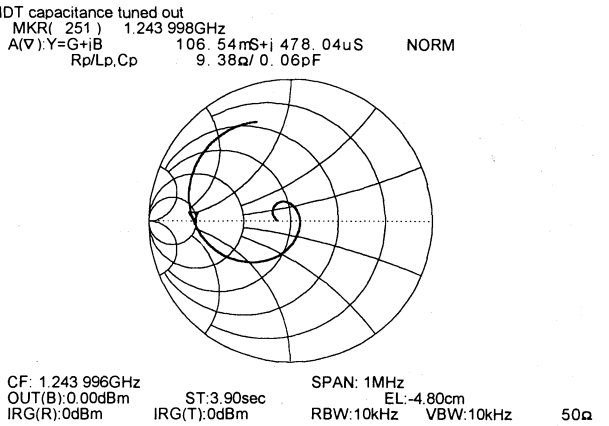


Fig. 13. Result of the reflection measurement with C_o tuned out with a parallel inductor.

5. FEEDBACK AND PIERCE-TYPE OSCILLATORS USING 2-PORT AND SINGLE-PORT STW RESONATORS

If a STW device is used to stabilize the frequency of a microwave oscillator, then the loaded Q of the acoustic device strongly influences the phase noise performance, the short term stability as well as the voltage pushing and load pulling sensitivities of the oscillator. To verify this, two types of fixed frequency oscillators, using the device from Fig. 1 were designed, built and tested. The first type is a feedback oscillator in which the STW device operates as a two-port resonator, and the second one is a Pierce-type oscillator in which the acoustic device operates as an equivalent inductor in a single-port resonator configuration. The simplified schematics of both oscillators are shown in Fig. 14. The oscillator output power with a 50 Ω load and the frequency

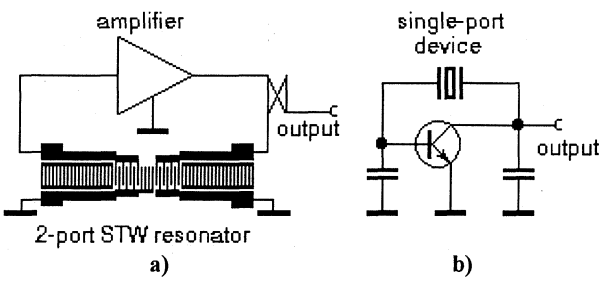


Fig. 14. Simplified schematics of the two oscillator types:
a) feedback oscillator,
b) Pierce-type oscillator

sensitivity to supply voltage variation (voltage pushing) were measured and the oscillator short term stability over a 1 s measurement time was evaluated. The results are summarized in Table 1. The comparison of both oscillators clearly indicates that the Pierce type oscillator has a much better short term stability and voltage pushing sensitivity compared to the feedback oscillator. This clearly indicates that in this oscillator the acoustic device operates at a substantially higher loaded Q than in the feedback oscillator and this is consistent with the data from the previous paragraph. Unfortunately, the better stability of the Pierce-type oscillator is at the expense of greatly reduced output power and RF/d.c. efficiency.

Parameter	Feedback STWO	Pierce type STWO
Output power	10 dBm	0 dBm
Voltage pushing	28 ppm/V	8 ppm/V
Short term stability	6×10^{-10}/s	2×10^{-10}/s

Table 1. Comparison of the measured data from both 1.244 GHz STW oscillators.

Similar results were obtained also with a 2.1 GHz STW device, characterized in Fig. 15. Although this low-Q device was designed to provide 2 MHz tuning range in a feedback-type voltage controlled oscillator (VCO), it was found to perform very well also as a single-port device in a Pierce-type oscillator. The measured data from both types 2.1 GHz oscillators are compared in Table 2.

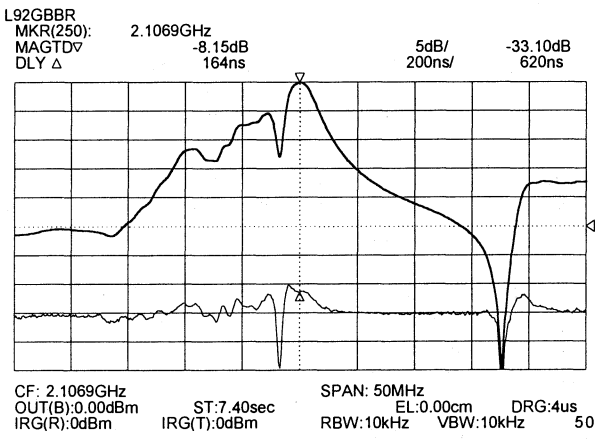


Fig. 15. Frequency and group delay responses of the 2.1 GHz device used in the oscillators.

Parameter	Feedback STWO	Pierce type STWO
Output power	5 dBm	-6 dBm
Voltage pushing	46 ppm/V	12 ppm/V
Short term stability	9×10^{-10}/s	4×10^{-10}/s

Table 2. Comparison of the measured data from both 2.1 GHz STW oscillators.

In the author's opinion the Pierce-type oscillator has a great potential as a miniature, low-noise and cost efficient microwave frequency source. In this design, because of its simplicity, it occupies an area of less than 1 cm^2. Using small size chip components it may be possible to accommodate the entire oscillator in a TO8 package.

6. SUMMARY AND CONCLUSIONS

This paper has presented results from a systematic study on increasing the loaded Q of GHz range STW resonators for oscillator applications. Two methods were investigated:

1. capacitive coupling applied to a 2-port device and
2. using the STW resonator as a single-port device.

It was shown that capacitive coupling can increase the loaded device Q to 80% of the unloaded Q value at the expense of only about 6 dB increase in device insertion loss. A device Q of 17450 was obtained with a 1.244 GHz single-port device which has otherwise an unloaded Q of 7130 when used as a 2-port device. Fixed frequency oscillators, operating at 1.244 GHz and 2.1 GHz and using both, feedback and Pierce-type designs, indicate that the Pierce-type oscillator, using the STW resonator as a single-port device, features a substantially higher short-term stability and lower voltage pushing sensitivity than the feedback concept using a 2-port device. However, this improved performance, attributed to the higher loaded Q of the single-port device, is at the expense of about 10 times less output power.

7. REFERENCES

[1] I. D. Avramov, "Extremely high loop power GHz range surface transverse wave oscillators using AB-class amplifiers", in Proc. 1993 IEEE Int. Frequency Control Symp., pp. 459-468.

[2] I. D. Avramov, F. L. Walls, T. E. Parker and G. K. Montress, "Extremely low thermal noise floor, high power oscillators using surface transverse wave devices", IEEE Trans. Ultrason., Ferroelec. Freq. Contr., vol. 43, no. 1, Jan. 1996, pp. 20-29.

[3] M. S. Cavin and R. C. Almar, "An oscillator design using low g-sensitivity, low phase noise STW devices", in Proc. 1995 IEEE Int. Frequency Control Symp., pp. 476-485.

[4] I. D. Avramov and Mei Suohai, "Surface transverse waves exceed the material Q-limit for surface acoustic waves on quartz", IEEE Trans. Ultrason., Ferroelec. Freq. Contr., vol. 43, no. 6, Nov., 1996, pp. 1133-1135.

[5] I. D. Avramov, O. Ikata, T. Matsuda, T. Nishihara and Y. Satoh, "Further improvements of surface transverse wave resonator performance in the 2.0 to 2.5 GHz range", to be presented at the 1997 IEEE Int. Frequency control Symp., Orlando, Florida, May 28 to 30, 1997.

[6] T. E. Parker, "The search for the perfect oscillator", 1996/97 Distinguished lecture for the IEEE Ultrasonics, Ferroelectrics and Frequency Control Society.

[7] I. D. Avramov, "Gigahertz range resonant devices for oscillator applications using shear horizontal acoustic waves", IEEE Trans. Ultrason., Ferroelec. Freq. Contr., vol. 40, Sept. 1993, pp. 459-468.

[8] S. I. Manolov and C. G. Tihchev, "Oscillators" in Handbook on Electronic Circuits, Part 4, V. K. Zlatarov, P. T. Petrov, Eds., Sofia, Darzhavno Izdatelstvo "Technica", 1982, pp. 223-227.

SUPPRESSION OF TRIPLE-TRANSIT SIGNALS IN BROADBAND SAW FILTERS

E.V.Bausk and I.B.Yakovkin

Institute of Semiconductor Physics of Russian Academy of Sciences
630090 Novosibirsk, Russia

The main characteristic of frequency-selective devices on surface acoustic waves (SAW) is the suppression value of undesirable signals in frequency and time domains. One of the most important spurious signal in SAW filters is the triple-transit signal (TTS) arising from multiple acoustic reflections between input and output. The simultaneous attainment of a good TTS suppression and a high out-of-band rejection in broadband SAW filters is a serious problem.

The proposed here technique has been developed for filters with passband of about 5-20% to provide an out-of-band rejection and a TTS suppression both of 50 dB and more.

Like [1], we realize TTS suppression with the help of another signal that has the opposite phase and the same amplitude. Our filter structure includes a multistrip coupler (MSC), input apodized interdigital transducer (IDT) and two output IDTs, one of that is dummy. An input signal is divided by the MSC between the dummy and output transducers that are situated in adjacent parallel acoustic tracks. The MSC provides a phase shift of 180° between TTS and a signal reflected from the dummy IDT. This shift doesn't depend on frequency. The output and dummy IDTs have identical structure with constant length of electrodes. Unlike [1] where an output transducer was apodized and a dummy IDT was unapodized, our filter structure doesn't require a complicated matching procedure of the dummy IDT in order to provide identical amplitude frequency dependences for these two reflected signals. It is enough to load the both transducers equally.

The output IDT with an unweighted time response has poor frequency selectivity and adds a significant value to stopband rejection of input transducer only in frequency regions far from its passband. We have developed an withdrawal weighting optimization algorithm to improve selectivity of broadband unapodized transducers [2]. The output withdrawal weighted IDT in couple with the input apodized transducer provide a filter out-of-band rejection of 50 dB and more.

For these devices we use lithium niobate as a substrate.

References:

1. W.A.Porter and B.Smilowitz. - IEEE Ultrasonics Symp. Proc., 1982, p.35-39.
2. E.V.Bausk and I.B.Yakovkin. - IEEE Frequency Control Symp. Proc., 1996.

AGEING STATISTICAL STUDY ON ONE HALF YEAR OF OCXO IN SC-CUT GLASS HOLDER

Olivier BIGNON, Hervé FIALAIRE, Patrick COLLET

CQE
TEKELEC TEMEX
2 Rue Robert Keller
10150 PONT SAINTE MARIE
Tel : +33 3 25 76.45 00
Fax : +33 3 25 80 34 57

1. INTRODUCTION

A number of papers have been written in the past several years about oscillator ageing.

An ageing measurement shall, to the greatest extend practical, avoid the effects of environmental sensitivities (such as temperature and barometric pressure), noise and reference error.

Ageing data can be continuous or sampled. All measurements shall extend over sufficient time that the noise inherent in the measurement is smaller than the frequency changes to be measured in cases where the ageing is small.

Consequently, the measurement of ageing thus requires ascertaining the integrity of the frequency source and the stabilisation of environmental parameters to levels below those which cause frequency changes of the order of magnitude of the observed ageing.

The purpose of this paper is to report representative ageing results from tests on 28 oscillators, from commercial oscillators based on standard SC-cut quartz resonators of current CQE production, for periods ranging up to 140 days.

These oscillators, all OCXOs, were in laboratory ambient where the internal thermal control circuit maintained a nominally constant temperature.

Ageing data was collected over a period of 90 days. To optimise the industrial manufacturing of OCXOs, we have made correlation between logarithmic and exponential fits of the ageing of long duration.

We propose to show the first results and we expected in the near future to build a new method which should avoid long ageing in bay.

2. AGEING SYSTEM DESCRIPTION

The measurement of high precision ageing (some 1.10^{-11}) can not be realised directly.

The in-house CQE system uses a Cesium Beam Frequency standard (TF 554 OA) which delivers a 10 MHz frequency compared to an output frequency and the difference is displayed on a frequency difference corrector (ADRET 4110A) (see on last page : the block diagram of ageing bank).

With an heterodyne box, the system exceeds 30 MHz and permits measurement of any frequency with high accuracy.
This system is connected to an ageing bank which control 1024 oscillators in 4 racks of 256 positions each.

Software and hardware has been built to drive up to 4 bays, divided among one master bay and 3 secondary bays. The master bay receives and multiplexes frequency signals from the 3 other bays to frequency-meter.

Each oscillator is measured during the night at the same interval and results are accumulated with this automated ageing facility described above.

3. OSCILLATOR DESCRIPTION

Each oscillator of this study provides from CQE standard production. These OCXOs have been equipped with a SC-cut 3rd overtone glass holder quartz resonator which delivers a 5 MHz frequency.

These oscillators have been specified to offer a 10 MHz output frequency with a $\pm\ 50.10^{-9}$ frequency accuracy, a stability versus temperature range of $\pm\ 1.5.10^{-9}$ over -35°C up to +70°C and a long term ageing of $\pm\ 3.10^{-10}$ /day and $\pm\ 20.10^{-9}$/first year.

4. CALCULATION METHOD OF AGEING FIT

The CQE method is based upon correlation between data using exponential and logarithmic fits. On the ageing graph the ordinate is the reduced frequency Y_k where:

$$Y_k = \frac{\Delta F_k}{F_0} = \frac{F_k - F_0}{F_0}$$

Fo is the first frequency, measured at time t and the abscissa is the elapsed time from t.

On each graph there is a dotted line and a dashed line. The dotted line represents the exponential fit and the dashed line the logarithmic fit.

4.a Exponential fit

The exponential fit of a quartz resonator is

$$\mathbf{Y_k = A.exp\ (Bt) + C}$$

where t is the time, B a negative constant, A and C two constants, which are determined from a linear correlation fit to all of the data between the days indicated.

The ageing behaviour is given by the following method :

$$\mathbf{Y_{day} = A.Bexp\ (B.t)}$$

where t is the last value measured, A and B are the values of constants determined at t time.

For the first month of operation :

$$Y_{month} = Y_{(30+t)} - Y_t$$

For the first year of operation :

$$Y_{year} = Y_{(365+t)} - Y_t$$

4.b Logarithmic fit

The logarithmic fit follows the same law as the exponential fit and the function :

$$Y_h = A' \ln(B'.t + 1) + C'$$

has been proposed for extrapolation from the initial 4 days of ageing data to periods in excess of 1 year.

The daily ageing is given by :

$$Y_{day} = A'.B' / (B'.t + 1)$$

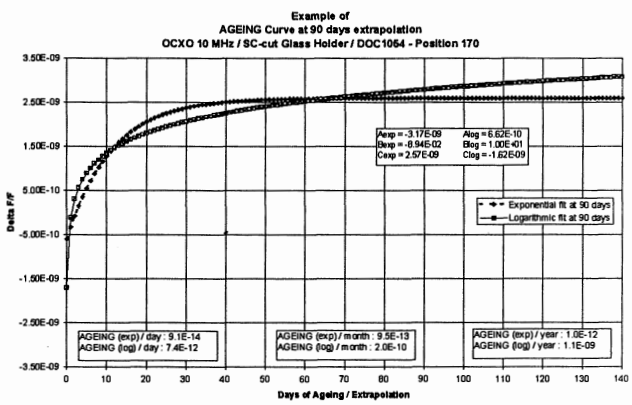
For monthly and yearly ageing, the formula are the same as for the exponential fit.

5.EXTRAPOLATION INTERPRETATION

In general, exponential extrapolation is more representative of real values.

We observe a good fit of ageing curves. This fit is all the better as measurements present a small difference with theoretical curve (little on none abrupt frequency variation from one measurement to another).

In general, for a well behaved oscillator that is kept on continuously, the exponential fit is a better approximation than logarithmic fit and the exponential extrapolation is very reliable even for short ageing, when retraces are negligible.



Quartz resonators that may provide retraces equivalent to observed ageing are poor indicator of long term ageing. In such case, a visual interpretation of the curve may be necessary to compare with calculated ageing.

6. RESULT ANALYSIS

The first analysis has been made on 28 oscillators during 140 days and we have compared the results given by exponential and logarithmic fits at 10, 30, 50, 70 and 90 days to 140 days. Ageing data at these intervals are named "predicted data" and compared to ageing data at 140 days named "actual data".

Average Measures at 140 days using Exponential fit on 28 OCXOs type DOC 1054 :

Nbre of Days	*Data*	*Stability /day*	*Stability/ month*	*Stability/ year*
10	Predicted	$1.9\ 10^{-11}$	$5.7\ 10^{-10}$	$5.7\ 10^{-9}$
	Actual	$0.7\ 10^{-11}$	$1.7\ 10^{-10}$	$1.0\ 10^{-9}$
30	Predicted	$1.2\ 10^{-11}$	$3.5\ 10^{-10}$	$3.2\ 10^{-9}$
	Actual	$0.7\ 10^{-11}$	$1.7\ 10^{-10}$	$1.0\ 10^{-9}$
50	Predicted	$1.4\ 10^{-11}$	$5.7\ 10^{-10}$	$4.1\ 10^{-9}$
	Actual	$0.7\ 10^{-11}$	$1.7\ 10^{-10}$	$1.0\ 10^{-9}$
70	Predicted	$1.6\ 10^{-11}$	$4.8\ 10^{-10}$	$4.2\ 10^{-9}$
	Actual	$0.7\ 10^{-11}$	$1.7\ 10^{-10}$	$1.0\ 10^{-9}$
90	Predicted	$1.9\ 10^{-11}$	$6.5\ 10^{-10}$	$4.6\ 10^{-9}$
	Actual	$0.7\ 10^{-11}$	$1.7\ 10^{-10}$	$1.0\ 10^{-9}$

Average Measures at 140 days using Logarithmic fit on 28 OCXOs type DOC 1054 :

Nbre of Days	*Data*	*Stability /day*	*Stability/ month*	*Stability/ year*
10	Predicted	$2.2\ 10^{-11}$	$6.7\ 10^{-10}$	$6.8\ 10^{-9}$
	Actual	$1.1\ 10^{-11}$	$3.0\ 10^{-10}$	$1.9\ 10^{-9}$
30	Predicted	$1.8\ 10^{-11}$	$5.6\ 10^{-10}$	$5.2\ 10^{-9}$
	Actual	$1.1\ 10^{-11}$	$3.0\ 10^{-10}$	$1.9\ 10^{-9}$
50	Predicted	$1.9\ 10^{-11}$	$5.7\ 10^{-10}$	$5.5\ 10^{-9}$
	Actual	$1.1\ 10^{-11}$	$3.0\ 10^{-10}$	$1.9\ 10^{-9}$
70	Predicted	$1.9\ 10^{-11}$	$5.8\ 10^{-10}$	$5.5\ 10^{-9}$
	Actual	$1.1\ 10^{-11}$	$3.0\ 10^{-10}$	$1.9\ 10^{-9}$
90	Predicted	$2.1\ 10^{-11}$	$6.2\ 10^{-10}$	$5.8\ 10^{-9}$
	Actual	$1.1\ 10^{-11}$	$3.0\ 10^{-10}$	$1.9\ 10^{-9}$

To complete this data, we have built a comparison between average deviation per day, month and year for exponential and logarithmic fits.

A review of the ageing data shows that exponential fit seems to be a better approximation than logarithmic fit. The first 30 days of ageing data is a good predictor of the long term ageing.

Figure 1

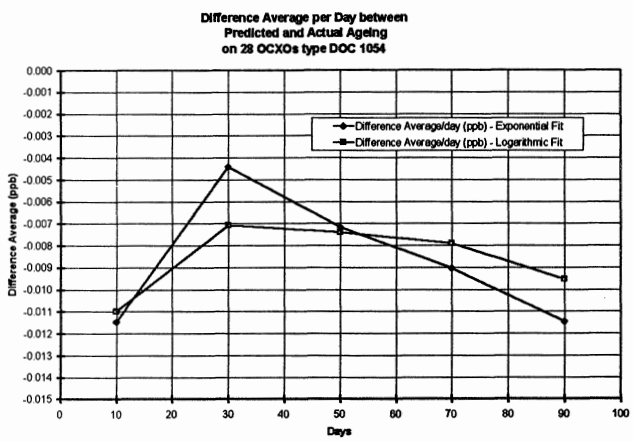


Figure 2

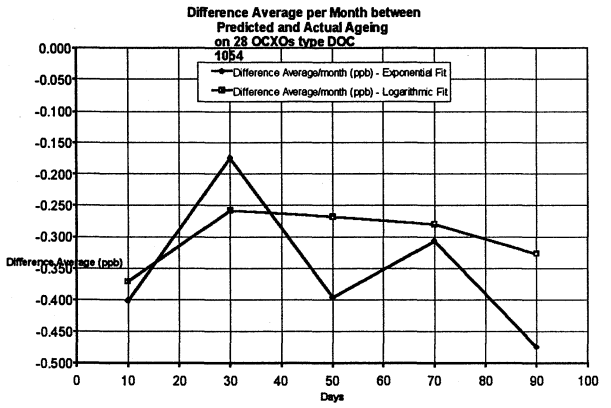
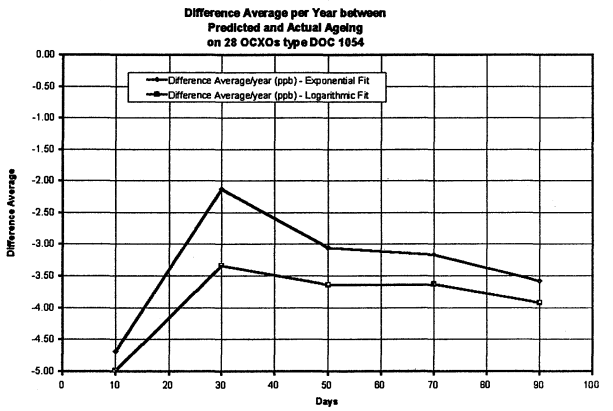


Figure 3



The cumulated curves of the average difference between predicted and daily, monthly and yearly actual ageing shows that we observed over a minimum 30 days and this result is equally obtained with an exponential fit as with logarithmic fit.

In other terms, the parameters of the ageing models obtained from the first 30 days of data may be a goof indicator of process control.

7. CONCLUSION

The obtained results show the necessity to go on with these ageing experiments to improve our knowledge on the possiblity to build a mathematic law of the ageing behaviour.

Exploration ways could be :

- must be weighted measure points and which of these? Are the last ten measurements characteristic of the behaviour and when?
- research of more complex mathematical functions to determine the better fit : e.g : new development of logarithmic and exponential functions as :

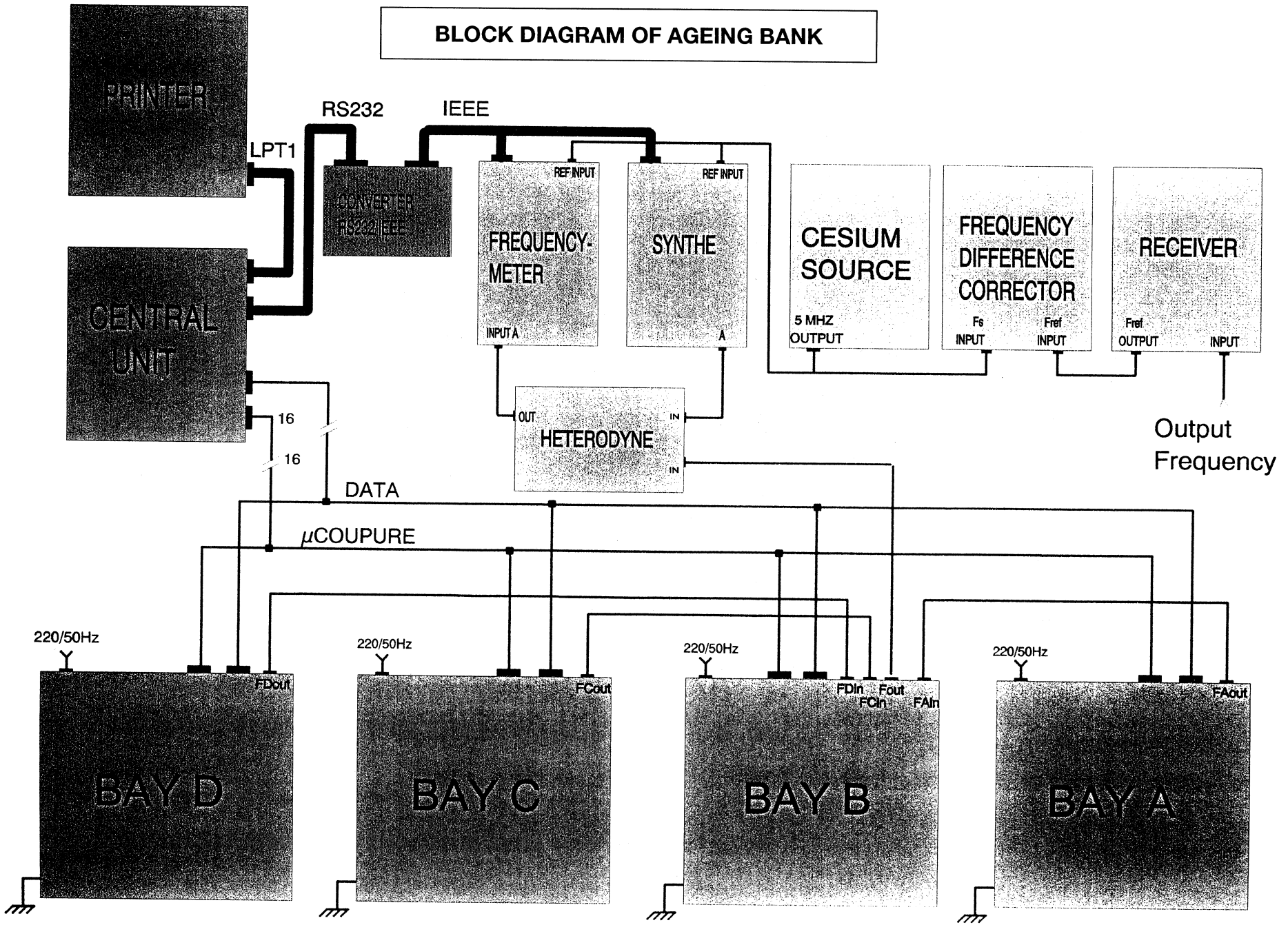
  $$Y = A \times \ln(B^2 t) + C$$

  and increasing modelisation to avoid spurious phenomena.
- continue this work over a very long period on different types of quartz resonators (AT and SC-cut) and with different types of OCXOs (medium and high stability).
- improve reliability of results and effect on approximation.

With such analysises, we would be able to build a mathematical model more and more reliable and integrate it in current software package.

8. REFERENCES

1. CQE data sheets, DOC 1054 and internal studies
2. L Odiet and K.K. Tuladhar , Compact low cost low ageing crystal oscillator for telecommunication applications and production statistics.(EFTF 1994)
3. R.L. Filler and J.R. Vig, Long term aging of oscillators (1992 IEEE Frequency Control Symposium).
4. J. Vanier, J.J. Gagnepain, W.J. Riley, F.L. Walls, M. Granveaud, Aging, Warm-up time and retrace; Important characteristics of standard frequency generators (1992 IEEE Frequency Control Symposium).


BLOCK DIAGRAM OF AGEING BANK
PRINTER
LPT1
RS232
IEEE
CONVERTER RS232/IEEE
CENTRAL UNIT
FREQUENCY-METER
REF INPUT
INPUT A
SYNTHE
REF INPUT
A
CESIUM SOURCE
5 MHZ OUTPUT
FREQUENCY DIFFERENCE CORRECTOR
Fs INPUT
Fref INPUT
RECEIVER
Fref OUTPUT
INPUT
Output Frequency
HETERODYNE
OUT
IN
IN
16
16
DATA
µCOUPURE
220/50Hz
FDout
BAY D
220/50Hz
FCout
BAY C
220/50Hz
FDin
Fout
FCin
FAin
BAY B
220/50Hz
FAout
BAY A

DRIVE LEVEL CRYSTAL OSCILLATORS CALCULATION

M. E. Bogoush, P. G. Pozdnyakov

Scientific Research Institute Phonon
44 Krasnobogatyrskaya, Moscow, Russia

The value of crystal unit excitation level influences upon its equivalent parameters (dynamical capacities) and characteristics in first turn upon the resonant frequency and temperature frequency characteristic. It is suggested in the report to characterize the excitation level of a crystal unit not by extensive values, such as dissipation power or piezoelectric current but by intensive ones - the mechanical stress or strain amplitudes. According calculations had been carried on to define the amplitudes of the mechanical stress in a crystal unit used in typical schemes of crystal oscillators. In particular crystal unit drive level was calculated for the cases of its being used in schames with active two-pole elements CCVS (current controlled voltage source), VCCS (voltage controlled current source) and with active three-pole elements VCCS, VCVS (voltage controlled voltage source).

REFERENCE

[1] V.N.Kouleshov, M.V. Kapranov, G.M. Oytkin: VIBRATIONS THEORY IN RADIOTECHNICS. Moscow, Science, 1983.

[2] PROBLEMS IN VIBRATIONS THEORY. Ed. L.V.Postnikov, V.I. Korolev. Moscow, Science, 1978.

DESIGN AND FABRICATION WIDE-BAND SAW DDL

V.F.Dmitriev and I.S.Mitrofanov

" Avangard-Elionica" Co. Ltd, 195271, St.Petersburg , Russia

ABSTRACT

This paper discusses a SAW scattering mechanism influencing the frequency response of a slanted Dispersive Delay Lines (DDL). Reflected power profiles along slanted Reflective Array (RA) has been calculated for fixed frequences within the DDL passband, with regard to the fact the incident power is subject to the partial conversion into volume waves. A correlation has been found between the DDL experimental frequency response and the calculation results. Based on our suggested analytical relations, we have revealed some characteristic features of experimental DDL frequency responses which can be associated with the SAW scattering process.

1. SPECIAL FEATURES of SAW SCATTERING in a WIDE-BAND DDL.

Scattering SAW into the bulk takes place in a DDL irrespective whether that has positive slope (up-chirp) or negative slope (down-chirp) of the dispersion curve. We will point out first that the most efficient scattering occurs where the SAW wave-length is less than twice the width of an RA element. Therefore, in the up-chirp, the effective scattering can be found in the RA area that is located upstream from the efficient SAW reflection region. In this case, the major part of SAW energy is strongly scattered by RA grooves just before being reflected, without any useful contribution into the reflection process. In the down-chirp the essential scattering occurs in the RA grooves region that lies downstream from the most effective SAW reflection region. The results of the RA grooves SAW scattering process are the increasing of the insertion loss and the changing of the frequency response shape.The unwanted scattering into the bulk by RA elements may be reduced, to same extent, through the use of a slanted DDL. Then propagation paths for transmitted SAWs will be separated in space for different frequencies with the result, that the SAW of some frequency will predominantely interact with the RA area wherein the reflection of the major SAW energy portion occurs. The SAW scattering into the bulk is an undesirable effect, but yet it may be used for equalization of up-chirp frequency response (Ref.1).

2. DESIGN PRINCIPLES of WIDE-BAND DDL.

A poor flatness of the frequency response (due to the fact, that the reflectivity increases as the square of the frequency (Ref.2)) is a characteristic feature of wide-band DDL. The most commonly used way to obtain an adequate DDL response is the RA design, wherein structural elements should be etched to variable (along the RA length) depths. It is worth noting that the response equalization technique with the use of variable aperture grooves and Inter Digital Transducers (IDT) electrodes is scarcely efficient because of the arising SAW diffraction. Moreover, a process of fabrication of DDL with variable-in-the-etch-depth grooves along the RA length is a complicated problem. However, the effects of increasing reflectivity with higher frequencies could be compensated by the increase of the SAW energy portion to be scattered, if one made a proper choice of features as well as a constant etch depth along the RA, when designing a up-chirp. It should be emphasized, that such approach is most efficient only if the slanted layout is employed, since the latter provides a means for a variation of the scattering area

location (i.e. of the insertion loss) in the RA through varying (within minor limits) the grooves aperture, the ratio between the width/pitch of grooves, and the IDT dispersion delay, as well. In actual practice, it is desirable to have a DDL with the minimum insertion loss. It is well known that the total insertion loss of a DDL consists of transducer loss, reflection loss, scattering loss, propagation loss, and diffraction loss. Both the propagation and the diffraction components can be neglected, when the IDT aperture is properly designed and frequencies are relatively low . Let us assume that the other parametrs of the IDT have been optimized (Ref.2). Then a DDL optimization for the insertion loss reduces to the proper design of the RA element etch depth. At a sufficiently low etch depth (i.e. $h/\lambda < 0.01$), the insertion loss decreases with a rise in h . However the h increase leaves the insertion loss unaltered, begining with a certain depth h . Let us note that this is true when absolute value of the reflectivity $|\Gamma| \ll 1$ and, as approximation of the first level, $|\Gamma|^2 \sim h^2/\lambda^2$. Such behaviour can derive from the fact that an increase in reflectivity of an RA element could be counterbalanced by more efficiently scattering SAWs. When it is considered that a SAW velocity change δV is increased as the etch depth increases in the vicinity of an RA element (whereas $\delta V \sim (h/\lambda)^2$), as well as that distorsions of the Phase-Frequency Characteristic grow, it makes sense that the etch depth h should be less than or equal to h_o which can be calculated for every specific DDL.

3. DESIGN PROCEDURE.

The SAW conversion loss is most pronounced in the wide-band DDL that have a long dispersive delay time. Therefore the slanted layout may be used most efficiently just in this application. The DDL design is based on the model wherein both IDTs and the RA are broken up into "channels " in accordance with any convenient for calculations principle. Since the grooves can be several thousands in number, the calculation based on the partial reflected wave superposition model may be a problem requiring several hours by AT 486DX4-100 computer. For preliminary results, it is reasonable to use the calculation method, where the response of the RA *i-th* channel defined as

$$K_{RA}^{i}(\omega) = [N_{eff}^{i}(\omega) R_{RA}^{i}(\omega)]^2 G^{i}(z) K_S^{i}(\omega), \quad (1)$$

where $N_{eff}(\omega)$ - is the effective grooves number in channel , $R_{RA}(\omega)$ - is the absolute value of the RA reflection coefficient , $G(z)$ - is the array factor (Ref.2), $K_S(\omega)$ - is the scattering loss in the *i-th* channel of the RA. The scattering loss $K_S(\omega)$ are calculated with following equation (Ref.3):

$$K_S^{i}(\omega) = \prod_{j=N_i}^{M_i} \exp[-\eta_j(\omega)(h_j/\lambda)]^2, \quad (2)$$

where the scattering loss in each *i-th* channel are summed up from the first groove of the *i-th* channel - N_i to last one M_i ; h_j - is the *j-th* groove depth. Calculation method of the frequency-dependent coefficient $\eta(\omega)$ is reported in the (Ref.3). For more accurate DDL frequency response predicting, it should be use the more complex calculation. Such calculation allowes to reveal some essential features of amplitude and phase frequency responses. Now, RA response (for *i-th* channel) is calculated using follow equation:

$$K_{Ra}^{i}(\omega) = \sum_{n=N_i}^{n=M_i} \sum_{m=N_i}^{m=M_i} \gamma_{nm} \rho_m \rho_n \xi_n \xi_m \exp[-\imath\beta(z_n + z_m + k_V x_n - k_V x_m)], \quad (3)$$

where γ_{nm} is the overfloor coefficient of *m-th* and *n-th* grooves in the *i-th* channel; ρ_m, ρ_n are the reflection coefficients, they are calculated by usual method (Ref.2); ξ_m - is the RA propagation coefficient up to

m-th groove. To take into account the SAW transform into the bulk waves,

$$\xi_m = \prod_{n=N_i}^{m} (1 - \zeta_n)[1 - (\rho_n)^2]^{1/2}, \quad (4)$$

$$\zeta_n = 1 - \exp[\eta_n(\omega)(h_n/\lambda)^2], \quad (5)$$

where $\eta_n(\omega)$ - is the coefficient which determine the part of transformed power on the *n-th* groove into bulk waves (Ref.3); z_m, x_m, z_n, x_n - are the coordinates of both the *n-th* and *m-th* groove centers respectively; $\beta = \omega/v_z$, $k_v = v_x/v_z$, v_x, v_z are the SAW velocities in the directions x and z respectively. Transmission coefficient and phase response of the DDL are determined as follow:

$$K(\omega) = (2R_g / \{[R_g + R_a(\omega)]^2 + X(\omega)^2\})^2 R_a(\omega) \sum_{i=1}^{N_K} [R_a(\omega)]^2 K_{RA}(\omega), \quad (6)$$

$$\Phi(\omega) = \operatorname{arctg}\left[\operatorname{Im}\{\sum_{i=1}^{N_K} K_{RA}(\omega)\} / \operatorname{Re}\{\sum_{i=1}^{N_K} K_{RA}(\omega)\}\right] + \Phi_{o_i}(\omega), \quad (7)$$

where $Ra(\omega)$ - is the IDT radiation resistance in the *i-th* channel (Ref.1), $X(\omega)$ - is the total reactance of the IDT, $Ra(\omega)$ - is the total radiation resistance of the IDT, $\Phi_0(\omega)$ - is the IDT phase response, Rg - is the impedance of the generator (load impedance). By the analysis of the layout, it should be determine the constant grooved depth of the RA, when a flat response (for up-chirp) should be obtained and the required change of the grooved depth along the RA (for down-chirp).

4. SIMULATION RESULTS.

The method described in the part 3 was used for design and fabrication DDLs with follows parameters: center frequency 85MHz, bandwidth 30MHz, chirp pulse duration 51μs.

The frequency responses calculation were performed on the base of presented in this work equations (2) - (6) (Fig.1a and 2a).

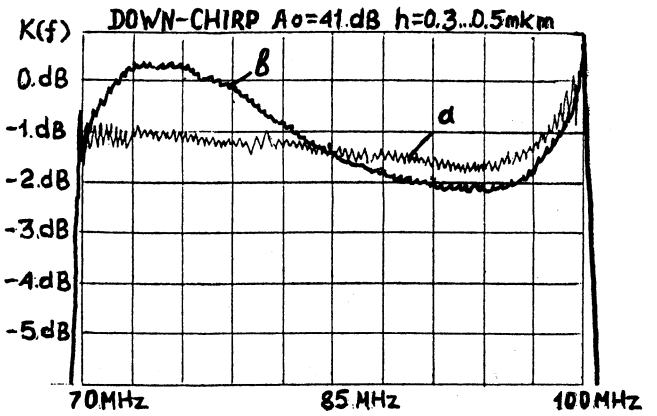


Fig. 1. Frequency response for up-chirp (F_0=85 MHz, Δf=30 MHz, τ=51 μs)
a - calculated, b - experimental

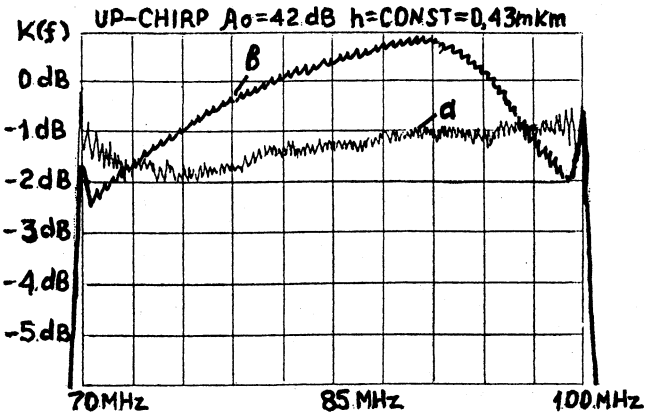


Fig. 2. Frequency responce for down-chip (f_0=85 MHz, Δf=30MHz, τ=51μs)
a - calculated, b - experimental

Flat frequency response was achieved with uniform grooves depth 0.43μm for up-chirp (experimental frequency response is presented in Fig 1b). The linear changing of the grooves depth (from 0.3μm in the vicinity of the IDT to 0.5μm in the end of the RA) was realized to smooth the down-chirp frequency response (experimental frequency response is presented in Fig.2b.
By our opinion, the coordinate curves of power reflected by RA are very useful for understanding the influence of the

scattering process on the DDL frequency responses. These curves (for DDL passband fixed frequencies, include and exclude the scattering) are presented in the Fig.3.

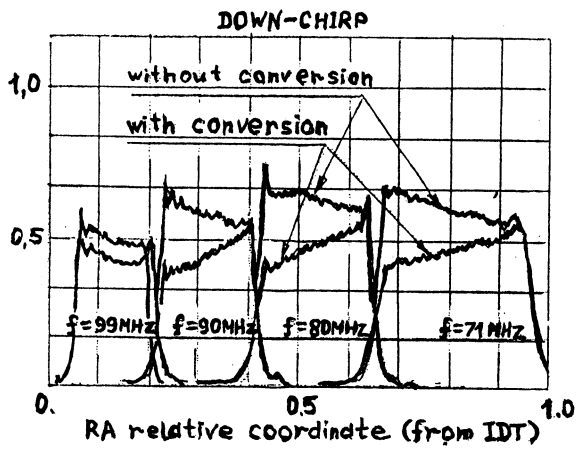


Fig. 3a. Various SAW scattering efficiency with propagation along RA (down-chirp)

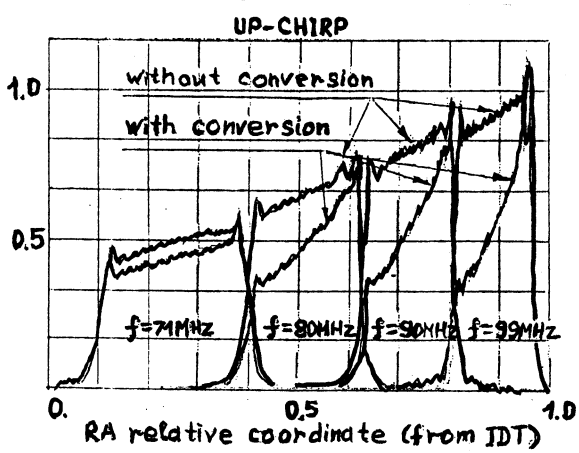


Fig. 3b. Various SAW scattering efficiency with propagation along RA (up-chirp)

The curves depicted in Fig.3 are described the various SAW scattering efficiency with propagation along RA (Fig.3a - in case of up-chirp, Fig.3b - in case of down-chirp). Fig.3b explaines the method of up-chirp frequency response equalization with constant grooves depth. With increasing of the frequency the arising of the grooves reflectivity is compensated by the arising of the SAW scattering efficiency. The calculations, which has been performed using (2) -(6), allowes to explane on experimental frequency responses features. In the case of down-chirp in the region of high frequencies (see Fig.1), the monotony character passband unflatness is caused by different efficiency of SAW scattering in the initial area of the RA (from the IDT) (see Fig.3b). Similarly, in the case of up-chirp in the region of low frequencies (see Fig.2), the monotony character bandpass unflatness is caused by different efficiency of SAW scattering in the initial area of the RA (see Fig.3a). Note that the slight "deflection" (in the high frequencies region of the response - Fig.1 and in the low frequencies region of the response - Fig.2) increases sharply because of the incorrectly choice of the RA grooves depth (over optimum). Such "deflection" can be reach 6 - 8 dB. Both DDL frequency responses theoretical predictions (calculations with equations (2)-(6)) and experimental evidence (with $h>>h_0$) confirm this fact. The frequency response ripples (close to boundaries of bandpass) is the Fresnel ripples. It can be avoided by RA aperture decreasing on the edges. However, it will demand to increase the substrate length about 10-15%.

5. THE DDL FABRICATION TECHNIQUE AND SOME RESULTS OF EXPERIMENTS

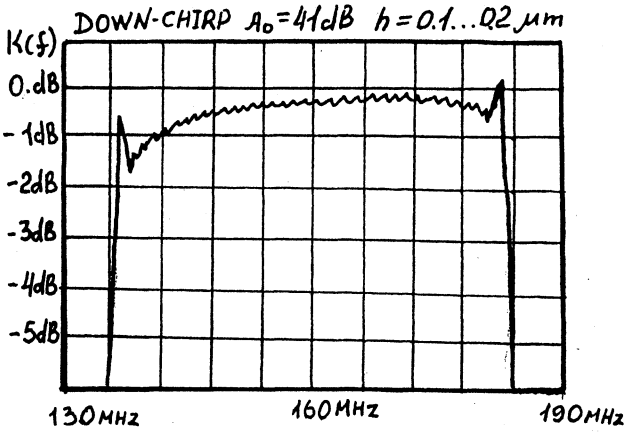
The process of fabricate DDLs is used a "single"photomask that combines the IDT and the RA layouts (Ref.1).
Dispersion delay lines have been manufactured and experimentaly tested. It has been found that the up-chirp with the uniform RA etch depth (0.43 μm) had both a fairly flat response (Fig.2b) the r.m.s. phase/frequency deviation within 2.8 degrees. It the case of down-chirp, (see Fig.1b) was achieved with the variable from 0.3 μm to 0.5 μm etch depth of elements along the RA, whereas the r.m.s. phase/frequency characteristic deviation

(from the quadratic one) did not exceed 2.5 degree.

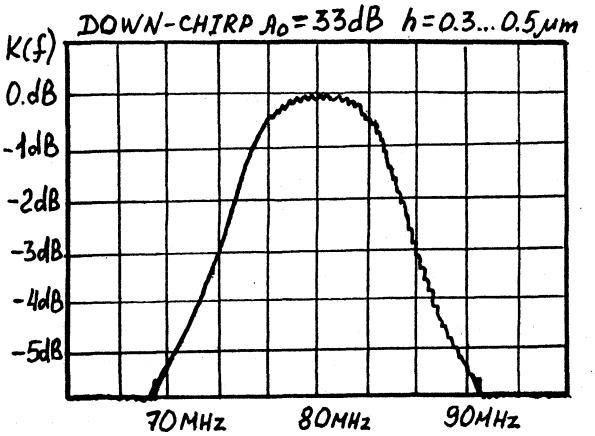
For the purpose of experiments on the chirp pulse compression, a down-chirp was designed specifically for weighted data processing in the shape of the Taylors function (compressing DDL). The DDL had the r.m.s. phase error below 1.9 degree at the operating pass band 30 MHz and the dispersion delay 51 μs.

The output chirp pulse was fed from the up-chirp immediately into the compressing DDL, while the down-chirp output first arrived at a balanced mixer, wherein the spectrum inversion was carried out, being delivered to the compressing DDL thereafter. The side-lobe level in the compressed signal (for either up-chirp or down-chirp) ranged from -28 dB to -32 dB with no phase compensation in samples.

We also design and fabricate DDLs for the expanding and compressing chirp signals with follows parameters: center frequency 160 MHz, bandwidth 45 MHz, down-chirp pulse duration 71μs (expander), no weighting; center frequency 80 MHz, bandwidth 22,5 MHz, down-chirp pulse duration 71μs (compressor), with weighting. Variation of groove depth over bandwidth is from 0.3 μm to 0.5 μm for compressor and from 0:1μm to 0.2μm for expander. For the compressor one transducer was weighted and the other was unweighted. The experimental frequency responses are presented in Fig. 4 a - expander and Fig. 4 b - compressor. In that case the output chirp pulse duration 142 μs was fed from the expander first arrived at a balanced mixer, wherein the spectrum inversion was carried out, being delivered to the compressing DDL thereafter. The side-lobe level in the compressed signal ranged from -26 dB to -32 dB with no phase compensation in samples.



a)



b)

Fig.4. Experimental frequency response DDLs (a-expander, b-compressor)

REFERENCES

1. Dmitriev V.F., Mitrofanov I.S. Equalization of SRAC filter amplitude respounse by means of SAW conversion loss control. 1995 World Congress on Ultrasonics, Berlin, September 1995.

2. D.P.Morgan. Surface - Wave Devices for Signal Processing. (Elsever, Amsterdam, 1985).

3. Dmitriev V.F. Coherent and uncoherent scattering of the Surface Acoustic Wave in the periodical reflectional structures . 1995, Sov.J.Tech. Phys, vol.65, N 8, p.111-123.

DIRECT PRINTING PHOTOLITHOGRAPHY FOR THE REALIZATION OF ACOUSTIC WAVE SENSORS

Irène ESTEBAN, Corinne DEJOUS, Dominique REBIERE, Jacques PISTRE

IXL - Université Bordeaux I (URA CNRS n°846)
351, Cours de la Libération - 33405 TALENCE Cedex - France
Phone : (33) 05.56.84.28.04 Fax : (33) 05.56.37.15.45
email : esteban@ixl.u-bordeaux.fr

ABSTRACT

Shear horizontal acoustic plate mode (SH-APM) sensors are receiving a great deal of attention in biochemical applications for few years. They consist of two interdigital transducers (IDT) deposited on a quartz plate, and of a selective coating placed on the acoustic propagation path along which the wave propagates. The sensor sensitivity to mass loading, the acousto-electric, viscosity, conductivity and temperature effects, depend strongly on the geometrical parameters of the delay line especially on the periodicity and on the number of interdigitated electrodes, on the plate thickness, on the considered mode...etc Consequently, our research on SH-APM sensors involves the realization of several devices with various geometries.

At IXL laboratory, a direct printing photolithography bench was developped by another research team [1,2,3] for ASIC's applications. After a technological adaptation, this bench can be employed to realize SH-APM prototypes. While the conventional photolithography technique requires a specific mask for each interdigital transducer design, the direct printing photolithography uses no mask, allowing the fabrication of delay lines of different designs in few quantities, with no additional cost and no delivery time. This paper reports on the technique of direct printing photolithography applied to the realization of SH-APM prototypes, and displays frequency responses of these devices.

Keywords : SH-APM sensors, technology, delay line design.

1. INTRODUCTION

In an acoustic wave sensor, the wave is excited by an input interdigital transducer, then travels along the acoustic propagation path coated by a sensitive film and reaches the output interdigital transducer (fig.1). Acoustic wave sensor detection mechanisms are based on the physico-chemical properties modification of the coating due to chemical species sorption, involving mesurable changes of the wave amplitude and velocity.

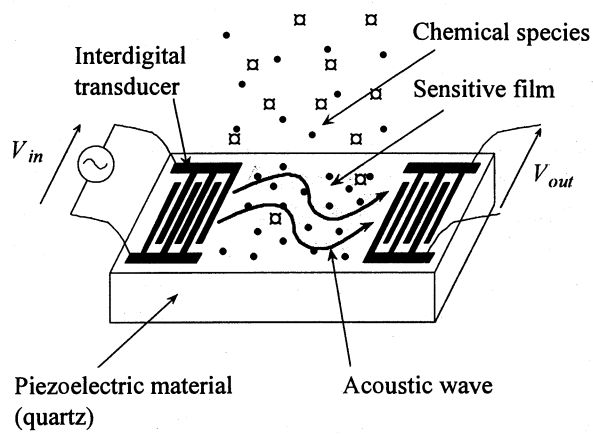


fig.1 : Schematic of an acoustic wave sensor.

In the case of an acoustic wave sensor based on SH-APM, the wave is shear horizontally polarized, in other words the mechanical displacement component is parallel to the surface of the device and normal to the direction of propagation. There is no displacement component normal to the device surface, this allows the propagation of the acoustic wave even if it is in contact with a liquid medium, without excessive energy loss. Moreover, SH plate modes propagate along the plate by multiple reflections between the top and the bottom of the plate (fig.2), and the mechanical displacements are equal on both faces of the device. As a consequence, the interdigital transducers can be separated from the liquid medium making the sensor all the more advantageous for detection in liquid environment.

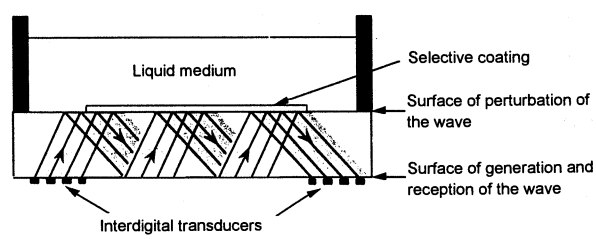


fig.2 :SH-APM devices in liquid medium.

Interactions between SH plate modes and the adjacent medium can lead to mass loading, acousto-electric,

viscosity and conductivity effects and constitute a possible way of detection [4,5].

However, these interaction mechanisms are all the more difficult to understand since there are many propagation modes. To validate theoretical studies on SH-APM propagation and to quantify the sensor sensitivity, it becomes necessary to optimize the delay line design. For that the contribution of geometrical parameters (as periodicity and the number of interdigital transducers electrodes, plate thickness, cristallographic orientation...) on the delay line electrical characteristics, must be evaluated. Consequently, our research on SH-APM sensors involves the realization of several devices with various geometries.

At IXL laboratory a direct printing photolithography bench can be employed to fabricate SH-APM prototypes of different designs. Section 2 presents the direct printing photolithography, and section 3 displays frequency responses of some prototypes realized with this technique.

2. DESCRIPTION OF THE DIRECT PRINTING PHOTOLITHOGRAPHY

The standard photolithography technique by insolation through a mask, is commonly used for the fabrication of interdigital transducers in an industrial production. The main technological steps of this planar technique are depicted in fig.3. An aluminium film is deposited on 'a quartz plate and covered by a positive photosensitive polymer. The delay line pattern is reproduced on the photoresist by insolation through a mask (1). During the developing process, the insolated photoresist is dissolved (2), the non-protected aluminium is then etched (3). Multiple identical chips are realized on a wafer and, after cuting, they are encapsulated.

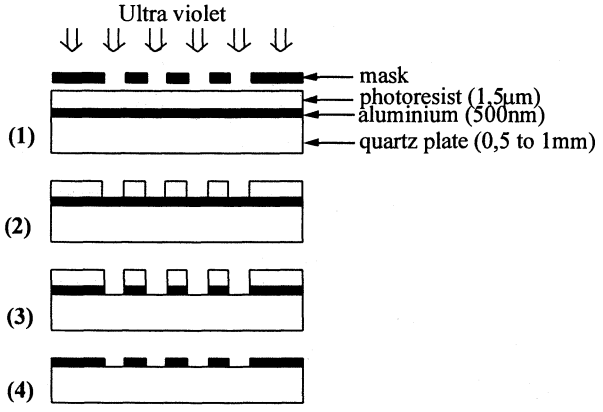


fig.3 : Technological steps of the standard photolithography by mask.

The "lift-off" technique is another photolithographic technique by mask, that provides better resolution. It consists of a positive photoresist deposited directly on the substrate and then insolated through a mask. During the developing, the insolated photoresist is dissolved. An aluminium film is deposited and finally, an appropriate chemical cleaning removes the photoresist with the covered aluminium.

The direct printing photolithography developped at IXL laboratory uses the same technological steps except the insolation one which uses no mask (fig.4).

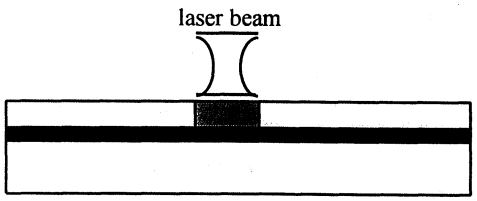


fig.4 : Insolation step by laser beam.

The displacements of the laser beam insolating the photoresist describes the delay line pattern.

On one hand, positive photoresists are not adapted for the direct printing technique due to the fact that the laser would have to sweep a too large surface. On the other hand, the use of a negative photoresist is problematic because of the lesser quality of this kind of product. That's why we used an inverted positive photoresist for the direct printing photolithography, which insured the high resolution of a positive photoresist while the substrate could be insolated as for a negative photoresist (the non-insolated areas were dissolved during the developing). The technological process of the inverted positive photoresist is described in [3] by M.Fathi.

The synoptic of the laser bench is shown in fig.5. The pattern is realized on a workstation using C.A.D software (OPUS). This generates a file which allows a micro-computer to control the X-Y laser beam displacements insolating the photoresist.

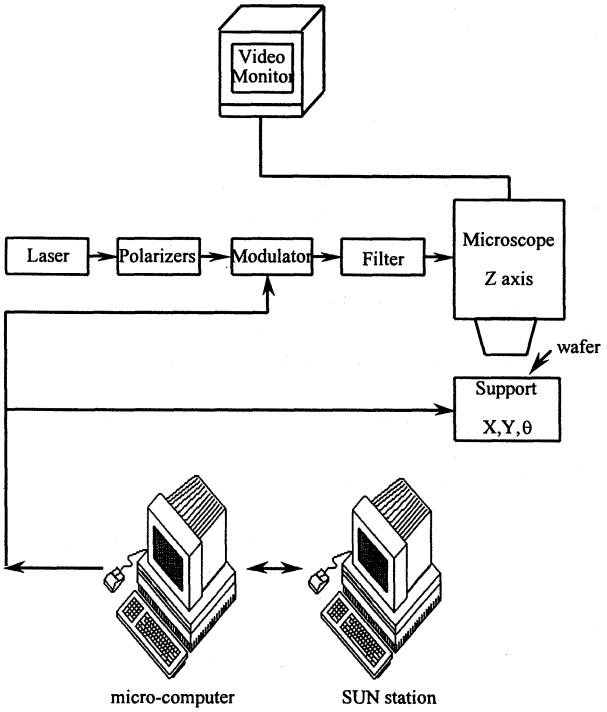


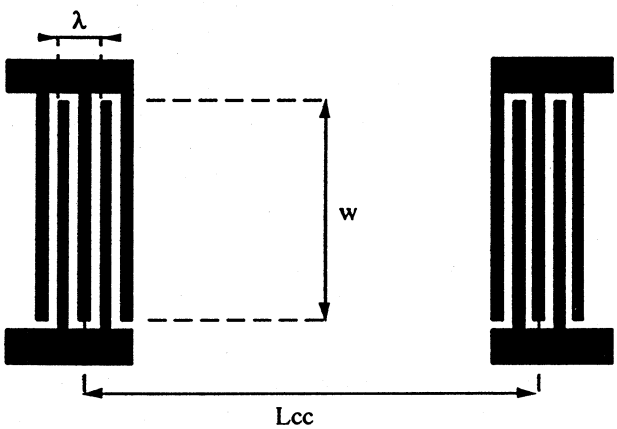
fig.5 : Synoptic of the direct printing photolithography bench.

The insolation source is an Helium-Cadmium laser of 442nm wavelength. The laser power can be adjusted by

polarizers placed on the optical path. An acousto-optical modulator is used to switch the laser beam. A system of associated mirrors lead the laser beam vertically through a microscope objective to the wafer surface. The latter, placed on a rotating socket, can be manually aligned in the X and Y directions (both, the X and Y directions and the wafer are displayed on a video monitor).

The step, from a technological point of view, is the line width after insolation and developing. When both the laser beam velocity and the technological process of the photoresist are determined, the step depends only on the laser power and can be fixed between 1 and 10μm.

Some technological adjustments were necessary to realize delay lines with the direct printing technique. Geometrical parameters of interdigital transducers (fig.6) are indeed critical for the fabrication of such devices.



λ	periodicity of intergitated electrodes,
Lcc	center to center distance between transducers,
w	electrodes aperture,
N	number of electrodes pairs.

fig.6 : Geometrical parameters of an acoustic wave delay line.

The major problem we encountered was when the laser power was too high : the non-insolated photoresist could not be correctly developed. This problem was solved by reducing the laser power, and hence reducing the technological step. This involved for instance, that the laser made 5 adjacent paths of 2μm to realize an electrode of 10 μm width, and consequently, the time to realize a delay line increased.

Another problem may result of the inverted positive photoresist properties which seem not to be homogeneous. We assumed that the use of a positive photoresist could provide better results.

Finally, the great versatility of the technological equipments in a university laboratory is not as well suited to the realization of such devices, as it is in micro-electronics industries.

Nevertheless, we are able now to realize SH-APM prototypes, the aim being to test and to improve the SH-APM delay line electrical characteristics.

3. EXPERIMENTAL DEVICES

Several devices were realized on quartz plates, with the direct printing photolithography. The geometrical characteristics of four of them are summarized in table 1.

The frequency response of device n°3 given by the network analyzer is shown in figure 7.

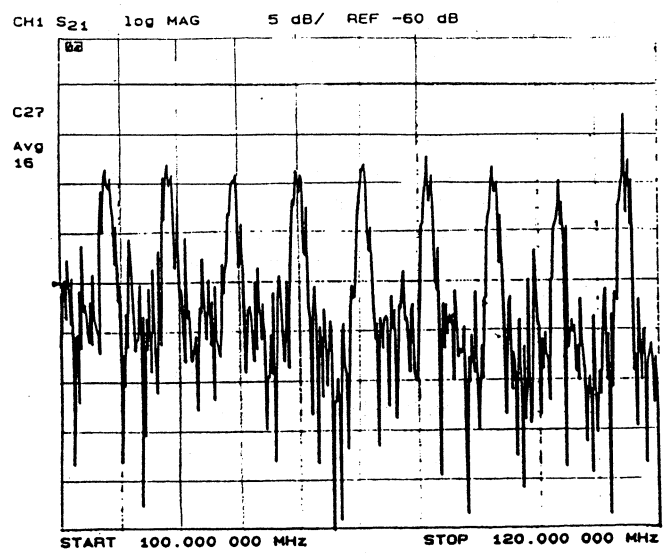


fig.7 : Frequency response of device n°3

The realization and electrical tests of such prototypes will allow us to study the influence of each geometrical parameter. These experimental results as well as the theoretical modelizations, will allow us to optimize a delay line for a desired application. For example, with devices n°2 and n°4, we can study the influence of the electrodes number on the quality coefficient of a SH-APM oscillator. The bandpass of each mode was measured at a constant phase. We found respectively a bandpass of 0,344MHz and 0,258MHz for the modes 82MHz and 118MHz, for device n°2, and a bandpass of 0,273MHz and 0,203MHz for the modes 82MHz and 118MHz, for device n°4. These results show, as expected, that device n°4 exhibits a narrower bandpass than device n°2 due to a great number of electrodes.

4. CONCLUSION

The direct printing photolithography can be used for the realization of acoustic wave delay lines. With regard to cost and time, it can advantageously replace the standard photolithography technique that needs a mask for each delay line pattern. However, the fabrication of one dual delay line by laser needs about one hour, which makes this technique only suitable for prototypes.

Therefore, the direct printing photolithography is an original way that suits well to the realization of SH-APM prototypes aimed for example at optimizing the acousto-electric characteristics of sensors (bandpass of modes, frequency stability, sensitivity...).

Device number	Cristallographic orientation. Euler angles: (0,Θ,90°)	Plate thickness h (mm)	Periodicity of electrodes λ (μm)	Number of pairs of electrodes N	Center to center distance between transducers Lcc/λ	First mode frequency (MHz)
1	(0,22°,90°)	0,5	40	40	200	83
2	(0,32°,90°)	1	60	40	70	55
3	(0,32°,90°)	1	60	60	70	55
4	(0,32°,90°)	1	60	70	70	55

table 1 : Geometrical parameters of SH-APM devices realized by the direct printing technique.

Acknowledgements

This work was supported by the French Ministry of Defence, through a specific program from the Direction des Recherches et Études Techniques (DRET) and especially from the Centre d'Études du Bouchet (CEB).

References :

[1] A.Fleury, "Etude et mise au point d'une station de photolithographie par faisceau laser Hélium Cadmium pour la personnalisation de circuits microélectroniques", Thèse de doctorat, Université Bordeaux I, Juillet 1992.

[2] M.Fathi, "Application de la gravure directe à la personnalisation des circuits intégrés et à la fabrication de modules d'interconnexion multipuces en technologie couches minces", Thèse de doctorat, Université Bordeaux I, Février 1996.

[3] A.Fleury, E.Saint Christophe, H.Frémont, M.Fathi, G.N'Kaoua and Y.Danto, "Laser beam lithography for direct patterning of interconnections on prediffused ASIC's", J.Phys. III France 5, 1455-1467(1995).

[4] Niemczyck T.M, Martin S.J, Frye G.C, Ricco A.J, " Aco ustoelectric interaction of plate modes with solutions ", J.Appl.Phys.64(10), 15 November 1988.

[5] Martin S.J, Ricco A.J, " Sensing in liquids using acoustic plate mode devices", 1987, IEEE InternationalElectr on. Devices Mtg Proc., Washington, USA, p290-293.

AC-Spice Simulation of Quartz Crystal Oscillators with the Negative Resistance Model

Detlef Göhrig
Tele Quarz Group, D-74924 Neckarbischofsheim, Germany

This article deals with the possibility of crystal-oscillator simulation with Berkley Spice 2G6. The simulation method based on the „negative resistance model" works with the *AC-analysis* of Spice. Absolute frequency, frequency deviation, check of the start-up condition and yield simulation are possible. For the extended simulation method, however, PSpice or another Spice simulator with extended functions are required.

1.General

Beside the often used open loop simulation [2], "Colpitts-" and "Clapp-" Oscillators can be simulated by means of the "negative resistance analysis". These two oscillator types can be described as 2 two-ports.

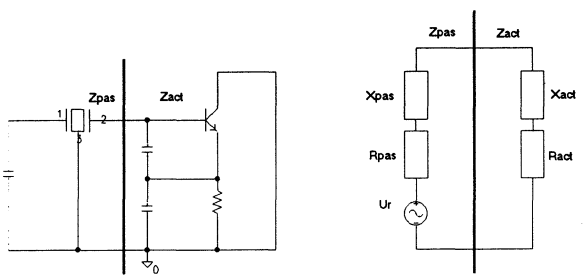


figure 1 active and passive part of the oscillator circuitl

The two port Z_{pas} describes the passive part of the oscillator, Z_{act} the active one. In this case the passive part of the oscillator has an impedance with positive real part. Thus it describes an attenuator in the signal theory. If the active part is an amplifier and has one feed back to its input the real part of the input impedance can be positive or negative. In case that the real part is positive, it also will be an attenuator and the magnitude of the gain is smaller than one. On the other side if the real part is negative, it will be an amplifier.

By connecting the passive and active part (see figure 1) the circuit can work as oscillator if the sum of the real parts of the whole impedance is negative. In this case the noise voltage can be amplified. This means for the start-up condition:

$$|R_{act}| > |R_{pas}| \qquad \text{sgn } R_{act} = -\text{sgn } R_{pas}.$$

With increasing signal power, the transistor operates at its self limitation, so the input resistance value is getting smaller. Only if the sum of both real parts is zero, the oscillator is in a steady state operation with the following condition :

$$|R_{act}| = |R_{pas}| \text{ und } X_{pas} = -X_{act}$$

This formula can be solved graphically [1] [5]. The circle diagram of the impedance of the active and passive circuit is shown in diagram 2. For the passive part of a crystal oscillator the locus of the complex impedance is a function of frequency and has only a very weak dependence of the signal power. On the other hand the active part at oscillator frequency is nearly independent of the frequency, and is a function of the signal power. The two impedance curves at one diagramm intersect at certain points, which yield information about the oscillator frequency.

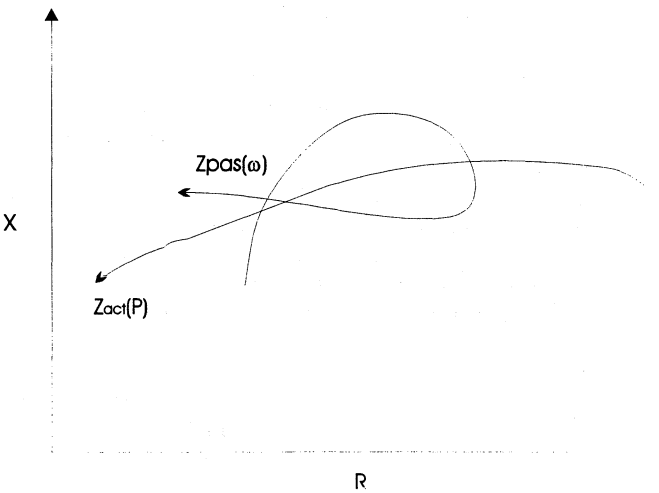


figure 2 The impedance $Z_{pas}(\varpi)$ and $Z_{act}(P)$ at one diagram

Taking into consideration the possibilities of an *AC-simulation* in PSpice it reveals that a power sweep of the active part is not possible.On the other hand the passive part is described very well with a simple quartz crystal equivalent circuit (3). Opposed to the start-up condition the steady state operation cannot be simulated. But this can be reached by summing Z_{act} and Z_{pas}. If $Z_{total} = Z_{akt} + Z_{pas}$ is depicted as a function of frequency, the oscillator frequency is represented as the point of intersection with the x-axis. The passive and active part can be defined by a bi-section of the crystal unit and the transistor. At this point two AC-current sources are introduced (see figure 3). The AC-power of these sources is less important, because the Spice *AC-simulation* is independent of the signal power. For simplicity a current value of 1A is chosen. From "Ohm's law" R=U/I and with I=1 the voltage value at the node of the current source is equal to the impedance value of the circuit.

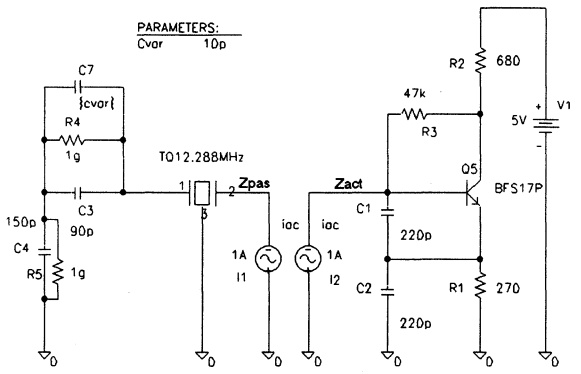


figure 3 typical oscillator circuit with help circuit

The crystal equivalent circuit used in figure 3 was generated by using the software tool TQSlib (download from server *www.telequarz.de*) [3]. This Public domain Software was developed by Tele Quarz Group, and it generates a Quartz Spice model alternatively from measured values or typical values (R_1, C_1, C_0 nominal load capacitance or frequency deviation)

The capabilities of this simulation method shall be shown by an example:

- A 12.288MHz Colpitts oscillator, with a fundamental mode crystal (R_1=8.01Ω; C_1= 19.49fF; C_0=4.81pF (incl. case capacitance) and C_{load}=39pF)

The frequency bandwidth of the first AC-analysis should be very wide, (i.e. MHz).With this simulation the start-up condition of the oscillator is checked. For that purpose the real part of the active circuit must be negative and the absolut value must be at least two times larger than the real part of the passive circuit in order to achieve reliable operation. PSpice has one software Scope named "*Probe*" to check this. It is very convenient to name the nodes at the current sources as Z_{act} and Z_{pas}. Another helpful tool is the macro function in *Probe*. For our circuit one macro is defined as:

$$Z_{total}=V(Z_{pas})+V(Z_{act})$$

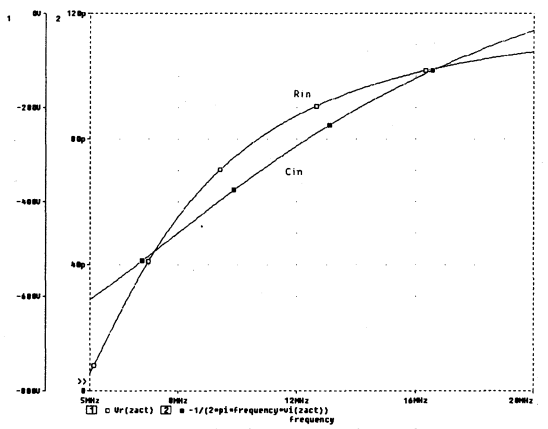


figure 4 input impedance of the active part

The magnitude of the real part at the chosen frequency bandwidth is very wide and therefore the oscillator will start up. Due to the small difference between the working frequency and the nominal frequency (only a few ppm) a small frequency bandwith should be chosen for the next simulation run. Simulation steps should be more than 101points.

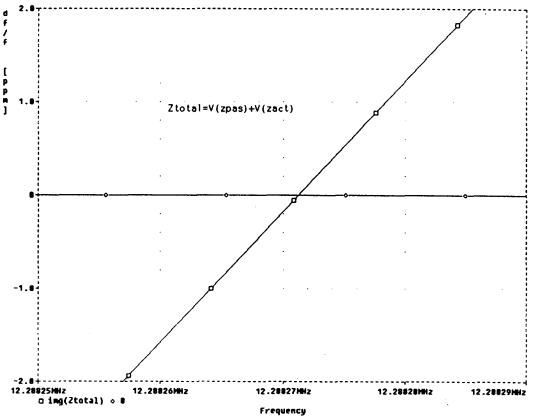


figure 5 imaginary part of Z_{total} cross the x-axis

Figure 5 shows the point of intersection with the x-axis of the imaginary part of Z_{total}, which is the start-up frequency of the oscillator.
Parametric Analysis of PSpice is used to simulate the tuning range of the oscillator. For this purpose the value of the chosen part is chosen as global variable, for example a capacitance or a resistance. This variable can be varied by means of the *.Step* instruction. From this a family of parametric curves results in *Probe*.
A better representation can be achieved with the *Performance-Analysis*. A *goal function* must be defined. This function searches the point of intersection with the x-axis. It is named:

IMNULL(1) = x1
{1|Search forward Level (0) !1;}

Usually the absolute frequency (in MHz) is of minor interest, the relative frequency deviation [in ppm] to the nominal frequency is preferred instead. The following macro will be helpful:

Freq=(Imnull(img(Z_{total}))-ref_freq)/ref_freq;
ref_freq= XXXX; XXXX subscripts the reference frequency

Thus it will be possible to display the relative frequency shift depending on the variation of the value of one component, for example the tuning capacitor.

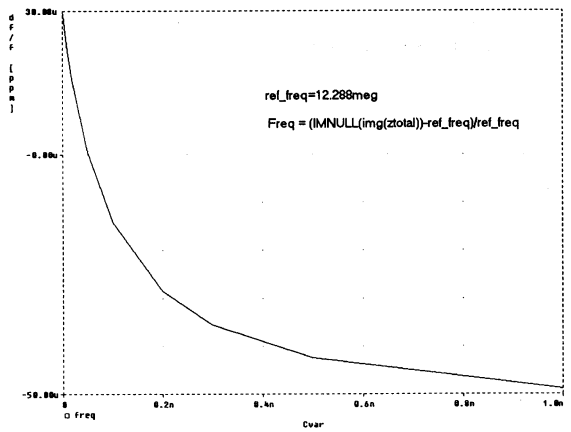


figure 6 possible frequency deviation depending on the tuning capacitor.

With this method the frequency shift depending on different values or different spice parameters of active component and analysis of the tolerance is possible. *Monte Carlo-* and *Worst Case-analysis* are tools in PSpice to make this very easy. Before starting the Monte Carlo analysis it is necessary to define a voltage controled voltage source, because this analysis needs one reference output voltage or current

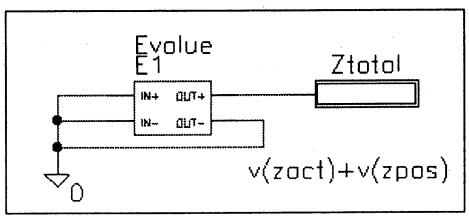


figure 7 help circuit to generate v(ztotal)

Figure 8 shows the Monte-Carlo-analysis of the sample crystal oscillator of figure 3. For the first run the tolerance range of the capacitors was 5% and the crystal has a frequency deviation of ± 10ppm.

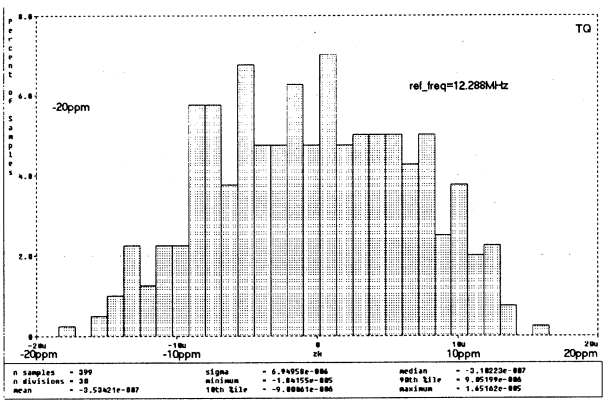


figure 8 Monte-Carlo analysis of the oscillator frequency with 5% capacitors

At the second run the crystal frequency deviation is the same but the capacitor tolerance decreased to 2%. The frequency tolerance of the oscillator was now 22ppm, an improvement of 10ppm.

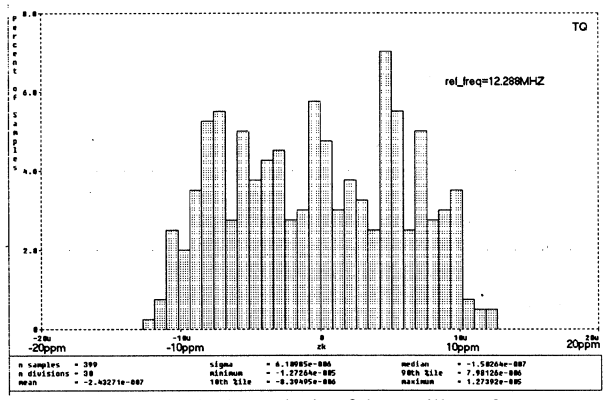


figure 9 Monte-Carlo analysis of the oscillator frequency with 2% capacitors

Normally the results of statistical analysis are a gaussian function, but with *Probe* it is not possible to show more than 400 simulation results in one histogram.

2. Comparison with measured data

The quality of a simulation can be proofed by comparing the simulation results with the experimental dates.

The simulated frequency is 12.2882712 MHz. This is 22ppm above the nominal frequency of 12.288MHz. A real tested oscillator mounted with the element values according to the simulation one has a frequency deviation of -2.9 ppm to the nominal frequency. Reasons for the difference between measured and simulated frequency are the small signal simulation method itself and that the components are assumed to be ideal.
The quality of the simulation results can be improved further by using real component models, i.e. equivalent circuits for the real components including the parasitic effects. A further improvement is to measure the value of the real component and simulate with the measured data set. But the uncertainty of the used transistor model still exists and is the most important reason for the difference between the simulation and measurement. The input impedance is a good example for this. In the simulation the impedance of the active part is:

real part:	-207 Ohm
imaginary part:	79pF

The signal power for the transistor Spice parameters measurement is usually unknown. Thus the measured values aren't directly comparable with the simulated ones. Nevertheless an estimation of the results is possible. The real impedance of the active circuit is measured with a signal power of -15dBm.

real part:	-243 Ohm
imaginary part:	69pF

Taking into consideration the simplified simulation models, this is good correspondence between measured and simulated values. Thus this method is suitable for crystal oscillator simulations.

Usually the relative frequency shift is more interesting than the absolute one. In this case the simulation results give very

good answers. This is very helpful to test the influence of some components to the oscillator frequency, for example the influence of the adjustment element C7 in figure 3.

For a capacitor C7 change from 10pF to 1nF the following frequency changes are achieved:

simulation:	78ppm
measured:	74ppm

3. Disadvantages

The disadvantage of this oscillator simulation method is the dependence from the accuracy of the semiconductor models. This problem increase with higher operation frequency, for example above 100MHz . Actually most public Spice BJT models are models of the die. But at a frequency above 100MHz it is necessary to take into account also the housing of the transistor [4].
Normally the standard Spice models of passive components are assumed as ideal. It is more useful to describe these components with their equivalent circuit diagrams or with something comparable, like a model with ABM-sources. The ideal inductors cause most problems.

However it isn't possible to simulate the temperature behaviour with the commonly used AT-crystals because the frequency vs. temperature corresponds to a third order polynominal function. A simulation with standard elements of Spice is therefore not possible.

4. Advanced solution with S-Parameter BJT model

The exact way is to simulate with the S-Parameters of the transistor at its operating point. This is not possible in Spice 2G6. Using *ABM-source*, PSpice has the possibility to simulate with S-Parameter [6] [8]. Another improvement of the accuracy of the simulation is to consider the signal power depending of the S-Parameter [7]. Figure 10 shows the structure of the S-Parameter BJT model.

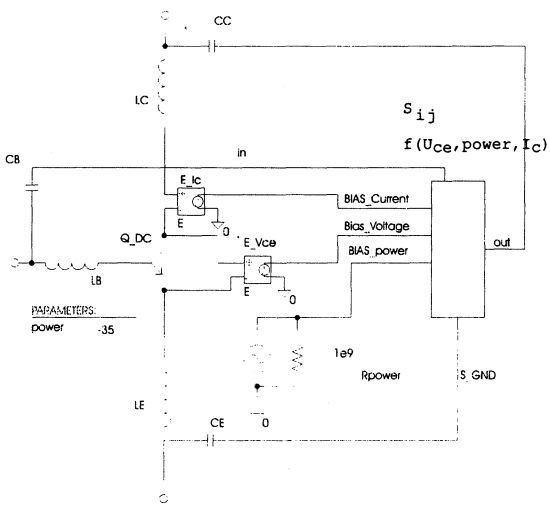


figure 10 S-Parameter BJT model depending on the signal power

The signal power must be defined as a parameter, as the detection of the signal power is not possible. Due to this fact a simulation with more than one transistor is not useful, but with one transistor it works very well.

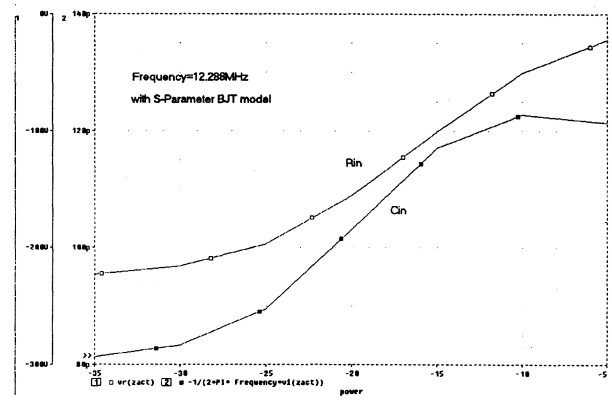


figure 11 Input impedance of the active part depending of power

Figure 11 shows the real and imaginary part depending on the signal power at the oscillator frequency. Now it is very easy to show the frequency deviation versus power (figure 12)

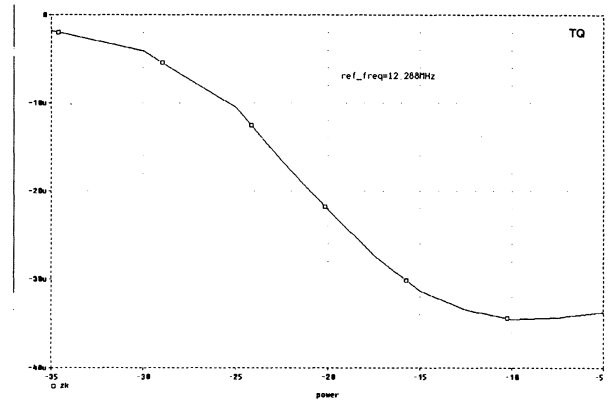


figure 12 frequency deviation depending of power

At a simulation with this model of the example in figure 3, the frequency is -11,39ppm. The measured value was -2.9ppm. With this model it is possible to simulate the start-up condition and the steady state operation with an Spice *AC-simulation*. The disadvantage of this model is the much longer simulation time, caused by the memory demand of more than 1MB for one transistor model.

5. Summary

It's possible to simulate a crystal oscillator with an *AC-simulation* based on the negative resistance model. This simulation method allows to check the frequency tuning by means of a *parametric analysis*. It's also possible to simulate the yield of an oscillator circuit with the *Worst-Case* and *Monte-Carlo* analysis.

6. References

[1] Gerhard Lohninger: Oszillatordesign in der Hochfrequenztechnik; Elektronik pp. 70-84

[2] T. Kien Truong: SPICE Techniques for Analyzing Quartz Crystal Oscillators; RF Design September 1995 pp. 28-35

[3] Tele Quarz GmbH: TQSlib.Exe; April 1996
Homepage: http://www.telequarz.de

[4] Siemens: Datenblatt SIEGET BFS405; Edition A03,04/94 pp. 8

[5] Edward C. Niehenke: RF and Oscillator Design; Westinghouse Electric Corporation Baltimore, MD 21203 pp. 4-11

[6]C.C. Timmermann: Exact S-Parameter and Noise Parameter-Based Semiconductor-Models for SPICE; IEEE Circuit & Devices Vol. 11, Nr.5 Sept 95

[7] T.Schano: Großsignal S-Parameter-Modell zur Simulation hochstabiler Quarzoszillatoren in PSpice; Tele Quarz GmbH, Diplomarbeit WS1995

[8] C.C. Timmermann, A. Bursian: S-Z-Y-Parameterschnittstelle für Pspice; Design & Elektronik 28.9.1993

PERFORMANCE LIMITS ON DIELECTRIC LOADED HTS RESONATORS AS FREQUENCY STANDARDS

Ling Hao, John Gallop, Farhat Abbas & Clifford C Hodge

Basic Metrology Group, National Physical Laboratory, Queens Rd., Teddington, TW11 0LW, UK

ABSTRACT

We assess the potential of HTS-shielded dielectric loaded resonators as practical frequency standards and reference elements for low phase noise oscillators. These circuits are being developed as potential flywheel microwave frequency standards for possible use in conjunction with the next generation of microwave frequency standards such as the caesium fountain. Here we report progress on developments of such high Q resonators shielded with relatively cheap, large area thick films of YBCO, operated at temperatures below 77 K. Q values as high as $3x10^6$ are reported. The resonator has been integrated with low noise amplifier, filter and phase shifter to provide a loop oscillator whose phase noise has been determined.

keywords: superconductor shielded dielectric resonator frequency standard

1. INTRODUCTION

Fig. 1a shows schematically the simplest design for a HTS shielded dielectric resonator. Here the dielectric "puck" (a short cylindrical element) is sandwiched between two planar thin or thick films of a suitable HTS material. For microwave applications this tends to be either $YBa_2Cu_3O_{7-\delta}$ or TlBaCaCuO since these materials have demonstrated lower microwave surface resistance than any of the other wide range of HTS materials. Using this design Q values as high as 10^7 at 5.6 GHz and $T = 60$ K have been achieved by Shen et al. [2] (usually the TE_{011} mode is used in these miniature dielectric resonators). However the temperature variation of the resonant frequency is certainly not negligible. Perhaps more important than the Q value for the ultimate performance of such standards is the insensitivity of the resonant frequency f to environmental changes in temperature or vibration, for example. We report the use of a composite dielectric puck configuration which address this issue, providing a turning point in the $f(T)$ behaviour within a convenient operating temperature range.

2. TEMPERATURE DEPENDENCE OF THE RESONANT FREQUENCY OF A DIELECTRIC RESONATOR

The temperature dependence of the resonator frequency f(T) arises from three main contributions: the variation of the superconducting penetration depth $\lambda(T)$, the linear thermal expansion of the dielectric pucks and other dimension-determining materials $\alpha_i(T)$ and the temperature dependence of the puck permittivities $\varepsilon_i(T)$. We have reported measurements of $\lambda(T)$ [4] which show that, for reduced temperatures $t = T/T_c \leq 0.2$, the fractional variation of the penetration depth is less than 0.05% K^{-1} for YBCO. The contribution that this would make to the frequency shift of the resonator depends in general on the volume-to-superconductor surface ratio or, more explicitly, on its geometry factor (see section 3).

At higher temperatures, such that $t > 0.5$, the superconductor contribution begins to dominate over other temperature coefficients of frequency. This is the situation for an HTS shielded dielectric puck resonator cooled in a single stage Stirling cycle refrigerator to around 60 K. We propose the use of a compound puck geometry, as indicated schematically in fig. 1b, to reduce the overall temperature variation of f(T).

In order to produce a f(T) curve which has a turning point at a convenient operating temperature (such as 60 K) it is necessary to have two significant contributions to $f(T)$ with opposing signs for df/dT. We have suggested elsewhere [5] that such a turning point may be produced in a sapphire puck at high temperature by deliberately doping the material with a paramagnetic impurity (for example chromium) in order to increase the magnitude of the paramagnetic susceptibility contribution. The composite resonator design shown schematically in fig. 1b also allows temperature compensation. First, the presence of low permittivity regions above and below the dielectric puck act as "dielectric mirrors" in tending to confine the electromagnetic energy within the puck, thereby increasing the effective geometry factor for the superconductor plates and thus decreasing the influence of $\lambda(T)$ on $f(T)$. Furthermore there exist low loss dielectrics which have the opposite sign for $\varepsilon(T)$ compared with most materials such as sapphire (for

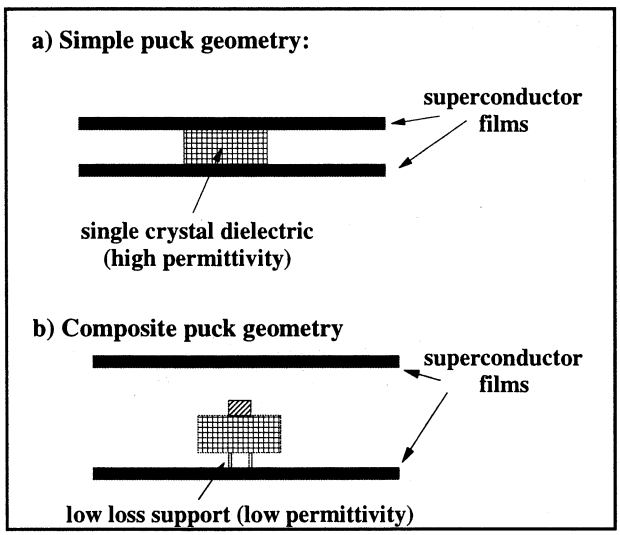


Fig. 1 Schematic of: a) simple puck resonator b) compound puck resonator with two different dielectrics with opposite sign of dε/dT.

example $SrTiO_3$ [6] or TiO_2 (rutile) [7]) so that use of two suitable complementary dielectrics in a composite resonator may allow temperature compensation in a useful temperature range, without the need for doping with a possible reduction of Q value. We have made measurements on such a composite puck resonator incorporating both sapphire and rutile dielectrics. The resonant mode structure is quite complex.

3. GEOMETRY FACTOR CONSIDERATIONS

The concept of a *geometry factor* corresponding to each component of a compound resonator is useful here in quantifying its design and performance. Thus the effect of the dielectric or conducting properties of a particular element on the properties of the *k*th mode of a compound resonator depend not only on the material properties but also on the relative proportion of the total stored electromagnetic field energy which is contained within that element. This is a function of both the dimensions of the element and the specific spatial distribution of the electromagnetic field for the *k*th mode.

To calculate the effect of the various contributions to the temperature dependence of the resonator frequency it is necessary to assign a geometry factor Γ_i to each element *i* of the compound resonator where Γ_i is given by the integral of the electromagnetic stored energy over the volum occupied by the ith component, normalised by the total stored energy in the resonator. Thus $f_k(T)$ can be written in the following form:

$$f(T) = f(0)\left[\Gamma_{sap}\left[\frac{\epsilon_{sap}(0)}{\epsilon_{sap}(T)}\right]^{1/2} + \Gamma_{rut}\left[\frac{\epsilon_{rut}(0)}{\epsilon_{rut}(T)}\right]^{1/2} + \Gamma_s\frac{\lambda(T)}{\lambda(0)}\right] \quad (1)$$

where f(0) is frequency for the *k*th mode at T = 0, Γ_{sap} is the geometry factor for the sapphire puck, Γ_{rut} that for the rutile puck and Γ_s that arising from the effect of the superconducting thick films. The relevant temperature dependent materials parameters, the puck permittivities $\varepsilon_{sap}(T)$ and $\varepsilon_{rut}(T)$ and superconducting penetration depth $\lambda(T)$ are included specifically in (1) and considered below. There are additional reactance contributions such as that from residual electromagnetic fields at the copper housing which contains the compound puck and superconductor films and that from the thin-walled quartz spacer. However the Γ_i representing these additional losses can be made vanishingly small and we neglect them here.

4. TEMPERATURE COMPENSATION

For resonant structures as complex as those shown schematically in fig. 1b analytical solutions for the geometry factors do not exist so that it is difficult to investigate *a priori* the influence of the various resonator components on the frequency and especially the temperature dependence. However if the temperature dependence of the permittivities of the dielectric components and the superconductor penetration depth are known then these may be used to deduce the geometry factors from experimental measurements of the resonator characteristics.

The steps in the process are as follows: first measure the complex permittivities $\varepsilon_i(T)$ of the individual component dielectrics and the penetration depth $\lambda(T)$ of any superconductor in subsidiary experiments involving known simple geometries such as simple puck resonators or parallel plate resonators [8]. Next measure the temperature dependence of the frequency $f_k(T)$ for the compound puck resonator for one or more modes, denoted by k. Then provided that no two of the $\varepsilon_i(T)$ and $\lambda(T)$ have the same functional form for their temperature variations these parameters may be used as inputs to a generalised mathematical fitting routine which will allow the geometry factors Γ_{ik} to be derived according to (1).

Use of low loss quartz spacers and a copper radiation confinement enclosure has enabled us to measure the temperature dependence of the permittivity of single crystal sapphire which has been found to have a temperature dependent component varying, between 15 K and 80 K, as

$$\Delta\epsilon_{sap}(T) = 2.5x10^{-11}T^4 \qquad (2)$$

for several samples of single crystal sapphire from different sources [3]. In addition a similar structure has been used to measure the temperature dependence of the permittivity of a single crystal of rutile (TiO_2). The c axis of the crystal structure is aligned with the z axis of the puck and for TE modes, on which we have concentrated in this work, we have shown that a fourth order polynomial in T provides an excellent fit to $\varepsilon_{rut}(T)$ over the entire temperature range from 10 K to 300 K. The best fit to our data is provided by the following expression:

$$\epsilon_{rut} = a_0 + a_1T + a_2T^2 + a_3T^3 + a_4T^4$$

where the coefficients in the polynomial expansion are listed below: $a_0 = 113.446$, $a_1 = 0.043$, $a_2 = -0.002$, $a_3 = 7.724x10^{-6}$, $a_4 = -1.072x10^{-8}$. The results are in good agreement with measurements made on a similar sample of rutile [7].

A parallel plate resonator has been made, using a low loss quartz spacer with two 20 mm x 20 mm square Yttria stabilised zirconia substrates coated with thick film YBCO [9]. By measuring the resonant frequency as a function of temperature, especially in the temperature range from 60 K to 92 K we have shown that the surface reactance of these films can be closely fitted by a simple two fluid model for the temperature variation of the London penetration depth $\lambda(T)$ with $\lambda(0) = 1.6$ μm:

$$\lambda(T) = \frac{\lambda(0)}{(1-(T/T_c)^4)^{1/2}}$$

5. TEMPERATURE VARIATION ANALYSIS

Using measurements made for $f_k(T)$ for a number of geometries involving the same sapphire and rutile pucks as well as the same thick film YBCO samples we have evaluated the geometry factors Γ_i for each configuration.

Figs. 2a & 2b show the data with model fits given by the dotted curves. It is clear that the data are well fitted by the model. Table 1 summarises the main results of fitting (1) to the experimental curves. Case 1 consists of a sapphire puck and thick film HTS shielding plates. Case 2 includes a thin-walled quartz tube spacer to separate the sapphire puck from the HTS films (the spacer has little direct effect on the resonant frequency). Cases 3-5 involve the addition of a small (3.5mm dia) rutile puck mounted centrally on top of the sapphire

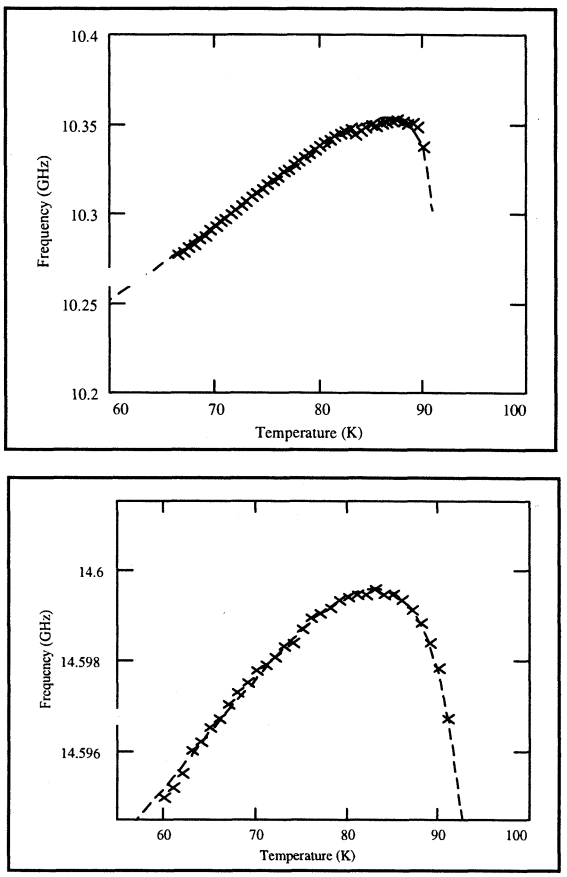


Fig. 2 Frequency vs. temperature for two puck configurations (a) rutile/sapphire puck with quartz spacer & HTS thick films (entry #4 in Table 1), (b) sapphire puck with quartz spacer (entry #1).

puck (12mm dia). Importantly for practical applications three modes shown here exhibit a turning point in the f(T) behaviour. For entry #5 in the table the geometry factor for the rutile puck is very large and the rapid temperature dependence of ε_{rut} dominates f(T) until the temperature approaches close to T_c when competition between $\varepsilon_{rut}(T)$ and $\lambda(T)$ produces a turning point at T = 87 K.

In cases 3 and 4 a turning point is produced at lower temperatures and inspection of the Γ_i indicate that for these two df/dT = 0 due to competition between the two temperature dependent permittivities. At these lower temperatures the variation of $\lambda(T)$ produces a negligible effect, accentuated by the smaller value of Γ_s than for the other three configurations. Inspection of the geometry factors indicate that for entry #2 in Table 1 $\Gamma_{rut} \sim 0$. This is reassuring confirmation that the method is able to separate the effects of different puck components correctly since for that configuration no rutile component was included in the resonator. The final column in Table 1 evaluates the second derivative $d^2f(T)/dT^2$ which is also an important parameter when temperature stability is being considered.

TABLE I

FREQUENCIES AND GEOMETRY FACTORS FOR A NUMBER OF MODES OF THE RUTILE/SAPPHIRE COMPOUND PUCK SYSTEM DESCRIBED IN FIG. 1. THE TEMPERATURE DEPENDENCES FOR THESE MODES ARE SHOWN IN FIG. 2

Case	f(0)	Γ_{sap}	Γ_{rut}	Γ_{HTS}	T_{TP} (K)	d^2f/dT^2
1	9.987	1.002	-.002	3.34×10^{-4}	N/A	
2	11.857	0.987	0.003	0.001	N/A	
3	10.402	0.483	0.523	0.006	82.0	-6.2×10^{-4}
4	14.588	0.961	0.039	6.39×10^{-14}	82.0	-3.6×10^{-5}
5	10.151	0.267	0.736	0.003	87.0	-1.8×10^{-3}

Thus although a turning point may be present it is inevitable that it will not be possible to stabilise the temperature of the compound resonator precisely at this feature.The parameter $d^2f(T)/dT^2$ indicates the error which will arise if the turning point can only be approached and maintained with a precision of δT. Note that for entry #4 the curvature is more than an order of magnitude smaller than for any of the other cases presented.

6. EXPERIMENTAL Q VALUES

The total loaded quality factor Q_l for the composite resonator can be written in the following form:

$$\frac{1}{Q_l} = \frac{1}{Q_{d_1}} + \frac{1}{Q_{d_2}} + \frac{1}{Q_c} + \frac{1}{Q_e}$$

where $1/Q_{dl}$,$1/Q_{d2}$ are proportional to the microwave power dissipation in dielectrics 1 and 2 of the composite puck, $1/Q_c$ is the loss in the superconductor and $1/Q_e$ represents the loss in all other components such as radiated power, and dissipation in the metal (typically copper) container.

The geometry factors derived in the above analysis may then be used to deconvolve the contributions from various loss mechanisms in the resonator. Thus the dielectric loss tangent $\tan\delta_s$ for sapphire is assumed to vary as $T^{4.75}$ whereas the loss tangent for rutile $\tan\delta_{rut}$ is taken from a non-linear least squares fit to the data from ref. [7]. The geometry factors for the dielectrics can then be calculated as described above in section 3. The temperature dependent $\tan\delta$ values may then be used to calculate the dielectric loss contributions to the loaded Q_l. This is shown as a dotted curve in fig. 3 which also shows the loaded $Q_l(T)$.

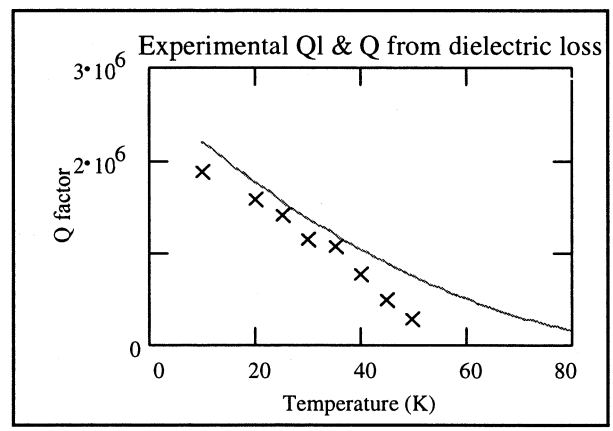


Fig. 3 Experimental loaded Q_l values (x) and theoretical dielectric losses (full curve)

Finally, substracting the dielectric losses from the inverse loaded Q yields the loss in the superconductor and the Q value associated with the YBCO thick films as a function of T is plotted in fig. 4.

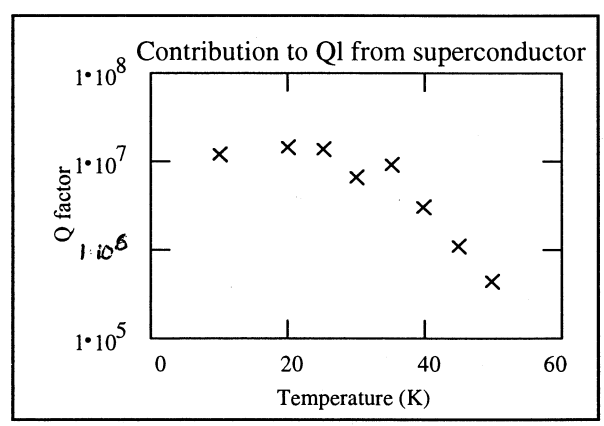


Fig. 4 Variation of superconductor Q with temperature.

Note that the superconductor Q is much higher than that from the dielectrics up to T > 45 K.

7. LOOP OSCILLATOR

One of the high Q resonator mode presented in the previous section has been used as the reference element for a simple loop oscillator. Positive feedback between output and input couplers of the reonator housing is provided by a low noise broad band amplifer via a narrow band tunable filter and an adjustable phase shifter. The microwave ouptut from the resulting oscillator has been mixed with a low phase noise synthesised-source as local oscillator, tuned to the identical frequency but in quadrature. The phase noise spectral density $S_y(f)$ of the resulting mixer output (essentially at d.c.) has been measured as a function of the carrier offset frequency f. The data is shown in fig. 5. The dotted line is a prediction from the well-known Leeson model [10]. The results are preliminary since the synthesised local oscillator source is believed to generate much of the measured phase noise.

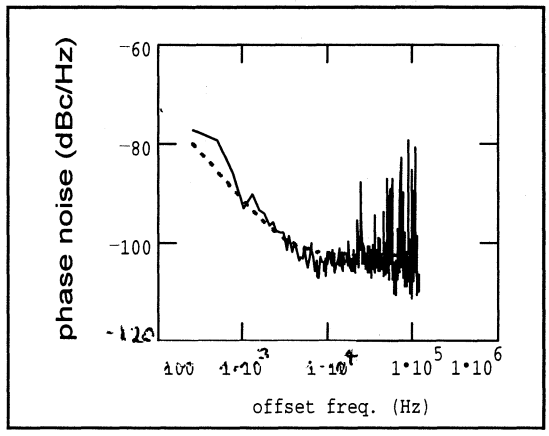


Fig. 5 Phase noise close to carrier for loop oscillator and synthesised microwave local oscillator

In future work we plan to compare two similar high Q HTS/dielectric referenced oscillators with optimised output power levels to achieve a more realistic estimate of the ultimate phase noise performance achievable with this type of device.

8. CONCLUSIONS

If the temperature of the resonator can be controlled to 0.1 mK within 1 mK of the turning point for the f(T) curve shown in fig. 2b this implies that the variation in resonant frequency should be less than 5 mHz or a fractional stability of 3 parts in 10^{14} even at long averaging times. This acievement would provide a microwave frequency standard with much better performance than the best 10 MHz quartz crystals are able to attain. Further questions concerning mechanical stability, and the ability to achieve temperature compensation and high loaded Q value at the same temperature remain to be answered but these preliminary results on in compound puck resonators are very promising. The use of these frequency references in flywheel frequency standards for the next generation of microwave atomic fountain or ion-trap frequency standards is being actively researched.

REFERENCES

[1] M. E. Costa, D. G. Blair, M. J. Buckingham, A. J. Giles, S. K. Jones, A. N. Luiten, P. J. Turner, A. C. Young and P. Hong, "A sapphire oscillator for VLBI radio astronomy," *Meas. Sci. Technol.*, Vol. 3, pp. 718-22, 1992.

[2] Z. Y. Shen, C. Wilker, P. Pang, W. L. Holstein, D. Face and D. J. Kountz, "High Tc superconductor-sapphire microwave resonator with extremely high Q values up to 90K," *IEEE Trans. MTT*, Vol. 40, pp. 2424-32, 1992.

[3] J. C. Gallop, C. D. Langham and F. Abbas, "Dielectric loaded HTS resonators as frequency standards and low phase noise oscillators," *Proc. 10th European Frequency & Time Forum,* IEE Conference Publication 418, pp 270-3, 1996.

[4] J. H. Clark, G. B. Donaldson, J. C. Gallop and R. M. Bowman, "Temperature dependence of surface impedance of PLD YBCO films," *IEEE Trans. Appl. Supercond.,* Vol. 5, pp 2555-8, 1995.

[5] J. C. Gallop, C. D. Langham, L. Hao and F. Abbas, "Dielectric loaded HTS resonators as frequency standards and low phase noise oscillators," submitted to *IEEE Trans. Instrum. Meas.*

[6] E. Sawaguchi, A. Kikuchi and Y. Kodera, "Dielectric constant of strontium titanate at low temperatures," *J. Phys. Soc. Japan,* Vol. 17, p. 1666, 1962.

[7] N. Klein, N. Tellmann, C. Zuccaro, P. Swiatek and H. Schulz, "YBCO shielded high permittivity dielectric resonators for oscillators and filters," *Applied Superconductivity,* Vol. 2, pp 743-8, 1995.

[8] R. C. Taber, "A parallel plate resonator technique for microwave loss measurements on superconductors," *Rev. Sci. Instrum.*, Vol. 61, pp. 2200-6, 1990.

[9] T. W. Button and N. McN Alford, "High Q YBCO cavitites," *Appl. Phys. Lett.*, Vol. 60, pp. 1378-80, 1992.

[10] V B Braginsky, V S Ilchenko & K S Bagdassarov, *Phys. Lett. A* 120 pp 300-5, 1987

[11] D B Leeson, *IEEE Proceedings* 54 pp 329-30, 1966

MODE-DEPENDENT NONLINEAR BEHAVIORS IN QUARTZ RESONATORS

atushi iobe and yutaka abe

Department of Applied Physics,
Hokkaido University, Sapporo 060, JAPAN

ABSTRACT
Nonlinear resonant behaviors in quartz plates are uniquely dependent on oscillation-modes. The physical origin of those behaviors is due to the difference of the dominant higher-order elastic constants in the respective modes.
The acoustic bistability in AT-cut resonators, and the shift of the resonant peak, and the narrowing of resonance width in X-cut resonators, are investigated numerically. The possibility of a chaotic current instability in a heavily damped AT-cut quartz resonator is also explored.
(Key words: Nonlinear effect of quartz resonators, Acoustic bistability)

I.INTRODUCTION

In the past, jump of the acoustic amplitude with frequency in AT- and BT-cut quartz resonators have been reported by various authors[1].. The quasi-static behavior of the acoustic bistability in a shear-mode quartz resonator has been investigated by Gagnepain and Besson[2]. However, the dynamic response of the acoustic bistability and nonlinear behavior of X-cut quartz has not been explored yet and remains to be solved.

In this paper, we will report the numerical results of nonlinear response of both AT-cut and X-cut quartz resonators using a simplified model. The chaotic current flow in heavily damped AT-cut resonators is also investigated numerically. Several applications of those nonlinear resonators are suggested.

2. SIMULATIONS

2.1 AT-cut Quartz Resonators

A rapid change of the current amplitude in a certain narrow range of the excitation frequency I an AT-cut quartz resonators can be experimentally observed with a moderate excitation level. This Frequency-jump phenomenon is accompanied by a large hysteresis.

Nonlinear wave equation for the AT-cut quartz resonator is given by

$$\frac{\partial^2 u}{\partial t^2} = v_d^2\left[1+\frac{1}{2}\left(\frac{c_{6666}^D}{c_{66}^D}\right)\left(\frac{\partial u}{\partial y}\right)^2\right]\left(\frac{\partial^2 u}{\partial y^2}\right) + \frac{r_{66}}{\rho}\left(\frac{\partial^3 u}{\partial t \partial^2 y}\right) \quad (1)$$

where c_{66}^D is the second-order elastic constant at constant electric field(SOE), c_{6666}^D is the fourth-order elastic constant (FOE), r_{66} is the linear damping constant, v_d is the sound velocity, and ρ is the mass density of the quartz. Introducing the equivalent nonlinear inductance and nonlinear capacitance , which

is expressed as

$$L(i) = L_0(1+\alpha i^2)$$
$$\frac{1}{C(i)} = \frac{1}{C_0}(1+\beta i^2) \qquad (2),$$

the current-voltage characteristic of the piezoelectric branch of the AT-cut quartz resonator is given by

$$\frac{d^2 i}{dt^2} + \theta\frac{di}{dt} + \omega_0^2 i + \Gamma i^3 = A\cos(\omega t) \qquad (3)$$

where

$$\theta = \frac{R_0}{L_0},\quad \Gamma = \omega_0(\beta - 3\alpha),\quad A = \frac{\omega_0 V}{L_0}.$$

The relationships between the circuit parameters and the material constants in the AT-cut quartz resonator are:

$$L_0 = \frac{\rho d^2}{8e_{26}^2 S},\quad C_0 = \frac{8e_{26}^2}{\pi^2 c_{66} d}$$
$$R_0 = \frac{\pi^2 r_{66} d}{8e_{26}^2 S},\quad Q_0 = \frac{L_0\omega_0}{R_0}$$
$$\alpha = -\frac{3}{256}\frac{c_{6666}^D}{c_{66}^D}\frac{\pi^2(1+Q_0)}{e_{26}^2\omega_0 S}$$
$$\beta = \frac{3}{256}\frac{c_{6666}^D}{c_{66}^D}\frac{\pi^2(1-Q_0)}{e_{26}^2\omega_0 S} \qquad (4).$$

Here, d is the thickness of the resonator, S is the area of the resonator, and e_{26} is the piezoelectric constant. The equation (3) belongs to the famous Duffing's equation in nonlinear mechanics. It is very difficult to measure FOE with non-resonant method experimentally, due to its smaller value compared with the second-order elastic constants. Eq(4) directly indicates that the nonlinear effect due to FOE is amplified by the order of Q_0 under a resonant condition and this fact makes the measurements of FOE extremely easier compared with the non-resonant method.

The numerically simulated results using the above equations are shown in Fig. 1 (a). The corresponding experimental results are also shown in Fig. 1(b) for comparison.

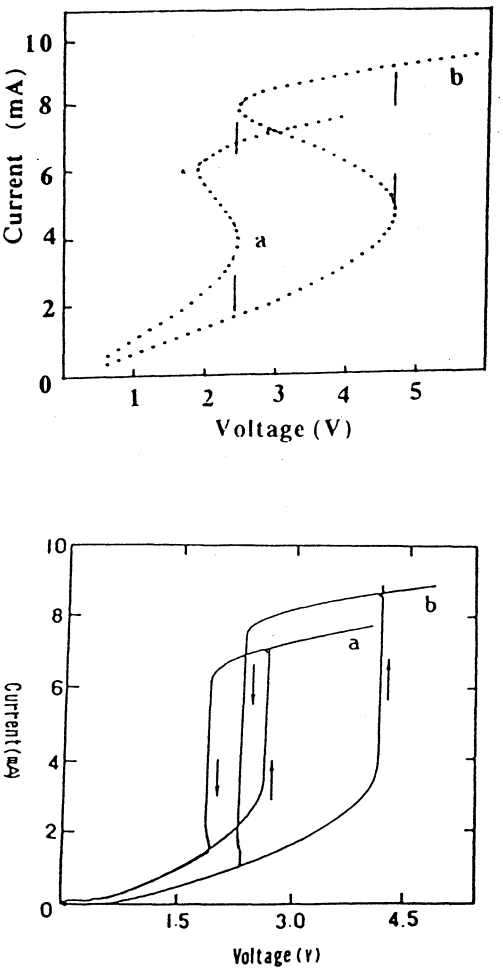


Fig.1 (a) simulated voltage-current (V-I)characteristics for a AT-cut quartz resonator. (b) experimentally obtained V-I characteristics. a: driving frequency is 4.193600 MHz. b: driving frequency is 4.193668 MHz.

As shown in Fig.1 , the simulated results are quantitatively very good agreement with the experiments. It should be noticed that an extremely small change in the driving frequency introduces a drastic change in the voltage-current characteristics, as explicitly shown in Fig.1. In the next section, we will examine the possibility of acoustic switching using a AT-cut quartz resonator.

2.2 Nonlinear Dynamic Responses of AT-cut Quartz Resonators

The nonlinear responses of the AT-cut quartz resonators described in the previous section were calculated under a quasi-static variation of the physical parameter such as the amplitude of the external voltage. We examine time-dependent responses of the resonators under the following condition: The amplitude of the external voltage is modulated linearly so that the rate of the modulation is much slower than the period of the external frequency.

$$A(t) = \lambda t, \quad 0 \le t < t_1$$
$$A(t) = -\lambda t, \quad t_1 \le t < t_2$$
$$1/\lambda << 1/\omega \qquad (5).$$

Several examples of the simulations are shown in Fig. 2 with the experimental results. It is seen that 1) the transition rate between the bistable states becomes slower as the scanning rate λ is increased, and 2) the width of the hysteresis loop becomes narrower as the value of λ is increased. The simulated results show very good agreement with the experimental results.

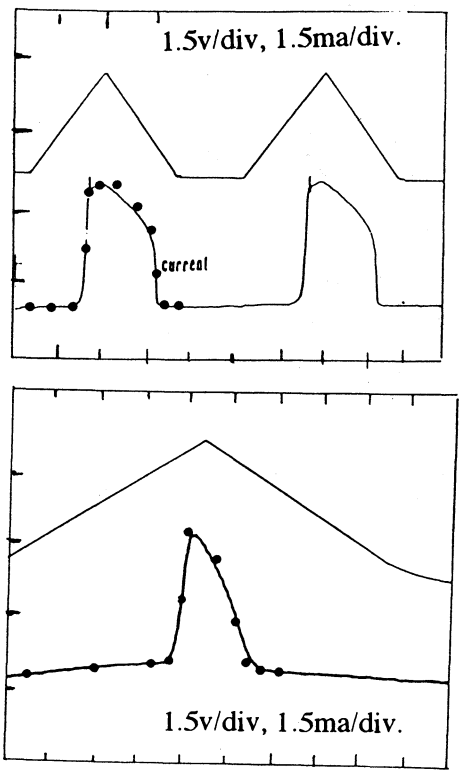


Fig.2 Time-dependent response of the AT-cut resonator. Upper trace: applied voltage. Lower trace: current response. (a) sweep seed: 5 seconds/div. (b) sweep speed: 0.1seconds/div.

The above investigations indicate that in order to realize a very fast acoustic switching using the acoustic bistability some improvements for compensating the slowness of the dynamic responses are necessary. If we applied the external voltage which is very close to the transition voltage as a bias voltage, the rise time becomes 3 times faster than the case without the bias voltage. However, a great improvement of the switching time can not realized by this method. The increase of the linear damping coefficient is much more effective for the fast acoustic switching, but in this case it is necessary relatively large power of the external excitation, and consequently heating of the resonators introduces another difficulty.

according to our simulations, it is suggested that we can realize the acoustic switching time of 10 nanoseconds with a 1 Hz AT-cut quartz resonator with $Q_0 = 10^4$ under 20W excitation level.

2.3 Nonlinear Responses of X-cut Quartz Resonators

The nonlinear behaviors in a ***longitudinal mode*** quartz resonator are very different from the ones of the ***shear-mode*** resonator. Within a moderate excitation levels, the X-cut resonator behaves almost same as a linear resonator. This means that the effect of FOE in the X-cut resonator is very weak and can not play any significant role. However, when the excitation level is further increased, the resonance peak starts to shift into the lower frequency region, and simultaneously the half width of the resonance peak becomes narrower

Even with very high excitation level, no amplitude bifurcation is observed. In order to explain the above experimental facts, we consider the following wave equation:

$$\rho \frac{\partial^2 u}{\partial t^2} = c_{11}^D \left(\frac{\partial u}{\partial x} \right)^2 + (3c_{11}^D + c_{111}^D) \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial x^2} + r_{11} \left(\frac{\partial^2 u}{\partial t \partial x} \right) \qquad (6)$$

Here, c_{111}^D is the third-order elastic constant (TOE). using the same approach mentioned in the section 2.2, we obtain the equivalent circuit equation of the following form:

$$\frac{d^2 i}{dt^2} + \gamma \frac{di}{dt} + i - \eta i^2 = F_0 \cos(\omega t + \phi) \qquad (7).$$

The term ηi^2 expresses the net nonlinear effect due to TOE. When we apply the first-order asymptotic approximation [3] for the equation (7) , the approximate solution does not involves any nonlinear term, which explains the actual experimental results of very weak nonlinearity in the X-cut resonator.

The physical origin of the narrowing of the half width of the resonance peak and the shift of the resonant peak under the intense excitation is due to the higher-order interaction of the amplitude of the fundamental frequency with the one of the second harmonics. The occurrence of this interaction is easily seen by assuming the response current is

$$i(t) = A_1 \cos(\omega t) + A_2 (\cos 2\omega t) \qquad (8)$$

and substituting (8) into (7).

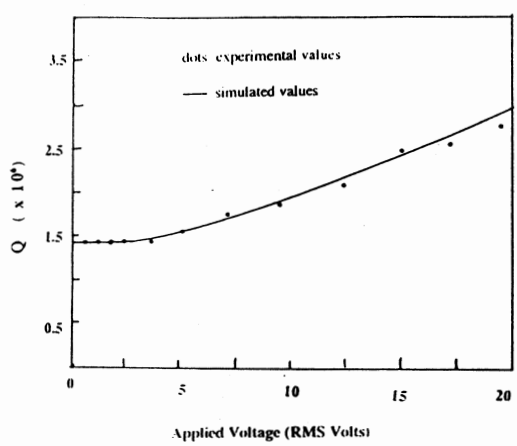


Fig.3 Q-values as a function of applied voltage. Solid line : Simulated results. dots: experimental values. The resnant frequency at low excitation level is 2.02648MHz.

The numerical results are shown in Fig. 3(a), (b). We did not examine the long-term stability of the resonant frequency under the above high level of excitation, which remains to be solved. However, the short-term stability of the resonant frequency in our laboratory system is quite excellent.

3. Chaotic Currents in a Heavily-damped AT-cut Resonator

The chaotic behaviors of Duffing's oscillator with large damping constant is extensively investigated in the past [3]. Experimentally we sometimes observe a random noisy signals when a AT-cut quartz resonator is attached to a bulk material of large acoustic absorption. It was considered for a long time that this anomalous signals were due to an imperfect adhesion between the resonator and the material. However, there exists the possibility that the occasionally observed random signal under a high excitation level and under a experimental circumstance of very low noises are closely correlated with the chaotic currents occurs in a heavily damped AT-cut resonator.

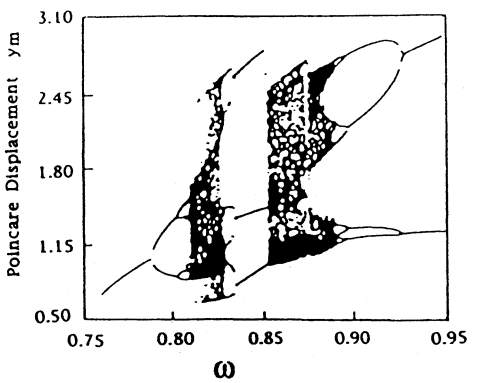


Fig. 4 Bifurcation diagram. This is the second bifurcation cascade between the third and the 5th harmonics.

A typical example of the numerical results of the bifurcation diagram is shown in Fig. 5. The above numerical results indicates that the chaotic states can occurs under a practically obtainable level of the excitation.

4. Conclusion

We have investigated nonlinear properties of quartz resonators using numerical simulations based on the simple models. The present investigations reveal that the nonlinear properties of quartz resonators are strongly depend on the mode of acoustic oscillation. It is found that the difference in the nonlinear behaviors between AT-cut shear-mode resonators and X-cut longitudinal-mode resonators is due to the difference in the dominant higher-order elastic constant. In the AT-cut resonators, FOE plays essential role, while in X-cut resonators, TOE is predominant origin for the nonlinear resonant currents. Acoustic switch using the acoustic bistability in AT-cut resonators is suggested.

The chaotic current instability in heavily-damped AT-cut resonators are also investigated numerically. The simulated results indicate that the chaotic instability can observe within a practically attainable excitation level.

References

[1] A. Seed, Proc.4th Int. Congress. Acoustics,(Copenhagen, 1962), C. Franx, Proc. Annu. Freq. Control. Symp. 21, 436 (1967).
[2] J. J. Ganepain and R. Besson, Phys. Acoustics, XI 245 (Academic Press, 1975).
[3] J. Guckenheimer and P. Holmes, Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector fields, Springer-Vellag, 1983.

Wideband HFF IF Filter for GSM Mobile Radio System

Osamu ISHII, Hideyuki KANNO and Tsuyoshi OHSHIMA

Toyo Communication Equipment Co.,Ltd.
1-1, Koyato 2-chome, Samukawa-machi, Kanagawa-pref., Japan

Abstract

This paper presents results obtained wideband 71MHz and 250MHz intermediate frequency (IF) filters which are small with low loss for the Pan European digital cellular radio system (GSM / DCS-1800). The filters are High Frequency Fundamental Monolithic Crystal Filter (HFF-MCF) using chemical etching the central section of a crystal substrate on one side. They are housed in small-sized ceramic packages with external dimensions of 5.2 × 4.5 × 1.1mm^3.

The HFF-MCF with a center frequency of 71MHz developed for GSM IF filter is designed in the balanced drive structure by a construction dividing the reverse side electrode into two input/output terminals. This filter has a 2-pole construction, with 3dB bandwidth of ±90kHz, group delay distortion of 0.7 μ sec at $f_0 \pm$ 80kHz, insertion loss of 0.9dB and stopband attenuation values of 14dB, 23dB and 40dB at $f_0 \pm$ 200kHz, $f_0 \pm$ 400kHz and $f_0 \pm$ 800kHz respectively.

For DCS-1800 base station applications requiring high stopband attenuation value, we present a cascade of two 2-pole HFF-MCFs at 250MHz, with 3dB bandwidth of ± 80kHz and stopband attenuation values of 30dB, 70dB and 80dB at $f_0 \pm$200kHz, $f_0 \pm$ 800kHz and $f_0 \pm$4MHz respectively.

1. Introduction

With the rapid spread of mobile communications typified by mobile phones, tremendous progress is being made with the digitalizing of systems, higher operating frequencies and the reduction in size and weight of handsets. Demands for small-sized IF filters of higher frequency and wider bandwidth as a part of the handsets are therefore stimulated still significantly. Filters with center frequency range of 45 to 250MHz, passband widths of ± 80 to ± 100kHz and sharp selectivity are particularly required for IF filters in GSM and DCS-1800 (which have the frequency band transposed to 1.8GHz using the same system as GSM) predominantly used in Europe.

With this system, transversal design surface acoustic wave (SAW) filters are now widely used which had excellent group delay characteristics and wide passband width. In order to give sharp selectivity to the SAW filters, some additional interdigital transducer (IDT) terminals are to be required which inevitably results in the increase of chip dimensions. However, there are problems such as the deterioration of insertion loss due to intrinsic bilateral characteristics of the SAW filters, and the increase of chip dimensions accompanying the increased number of interdigital transducer (IDT) terminals to improve suppression in the area of neighbour channel ($f_0 \pm$ 200kHz) sufficiently.

SAW filters using Z-path design with reduced chip dimensions were proposed by Machui et al in 1992 [1], and Resonant Single Phase Unidirectional Transducer (RSPUDT) design SAW filters with improvements in insertion loss proposed by Ventura et al in 1994 [2], but these haven't managed to achieve properties satisfying all the requirements mentioned above.

On the other hand, HFF-MCF which was developed by us in 1992 [3], in the way of employing the chemical etching processing technique has enabled to construct IF filters in a small size and has offered low insertion loss, and it was designed in the unbalanced drive type for simplifying the construction of electrode. As the integrated circuits (IC) now used for GSM receiver are the so-called balanced drive type, it is still necessary for us to cope with the difference.

Increasing the MCF frequency with chemical etching process techniques, and widening the passband width with a precise electrode pattern construction using photolithographic techniques, we produced 71MHz band IF filters for GSM use. Regarding balanced drive structure, we examined filter characteristics of a construction with the reverse side electrode divided into two for input and output terminals. Further increases in frequency were achieved by an additional

etching processes to adjust the wafer thickness in accordance with the target frequency. An evaluation of the experimentally produced 250MHz band filters for DCS-1800 systems is also included in this paper.

2. HFF-MCF Structure

In general, MCF has a number of advantages.

(1) Miniaturization is easily realized in the design of MCF because it has a very simple construction utilizing acoustic coupling between two electrodes.

(2) Low insertion loss and narrow passband width can be achieved effectively because Q value of crystal resonator is relativity high.

(3) Frequency temperature characteristics is excellent in wide temperature range.

Accordingly, if wide passband width and the design of the balanced drive structure are realized, HFF-MCF having low insertion loss is considered quite adequate to the small-sized IF filters for GSM/DCS-1800 systems.

Fig.1 shows the structure of the HFF-MCF with which higher frequency in fundamental mode is achieved by thinning the central section from one side with chemical etching of an AT cut crystal substrate.

To achieve a wide passband width in MCF, we have to make distance split electrode extremely small, just a few micron meters. Such a filter cannot be fabricated in conventional vacuum evaporation processes. However, band widening was attempted in the photolithographic construction with high precision, and as a result, an MCF with a fractional bandwidth of up to approximately 0.3% is to be achievable.

The chemically etched side of the crystal substrate all forms a electrode as a ground electrode, and so it can also be used as an unbalanced drive filter in the same way as a conventional MCF. Then, by splitting this full-face electrode, including the ceramic package terminal, into two input/output electrodes to configure the electrode pattern using each of these divided electrodes as input/output electrodes respectively, it is possible to simply obtain a balanced drive filter even with HFF-MCF manufacturing. The structure of balanced drive HFF-MCF is shown in fig.2.

3. Manufacturing Process

A simplified process sequence for the HFF-MCF manufacture is shown in fig.3. High productivity in the mass production is achieved by batch processing performed in the wafer state from electrode pattern formation through frequency adjustment. In this manufacturing method, it is important to ensure uniform thickness for each of the elements on the wafer. Adjustment of thickness is made to within approximately ± 30nm by additional and individually etched on each element after the main etching process. With this method, HFF-MCF with a center frequency of 150MHz can be produced with the precision of frequency within ± 30ppm.

In case of attempting frequency increase beyond 150MHz, it is difficult to achieve the desired filter characteristics further, because the thickness of vibration area is extremely thin, and as the amount of frequency adjustment by evaporation after a single individual etching process excessively increase, compared to the thickness of vibration area. However, by repeating individual etching using the dilution etchant for high-precision thickness adjustment etching, the thickness distribution for individual elements in the wafer was reduced to within ± 10nm, and this makes the filter characteristics demanded possible. This process corresponds to the area within the dashed line in fig.3.

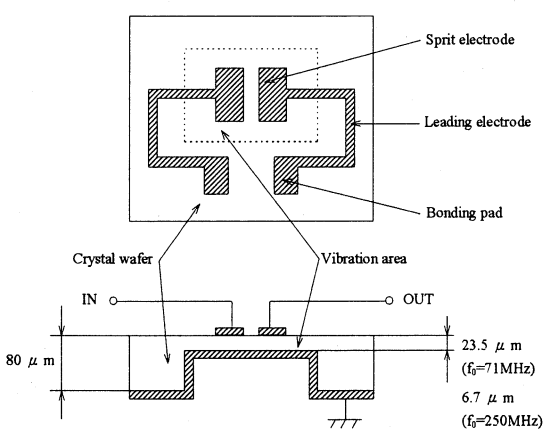


Fig.1 Structure of HFF-MCF

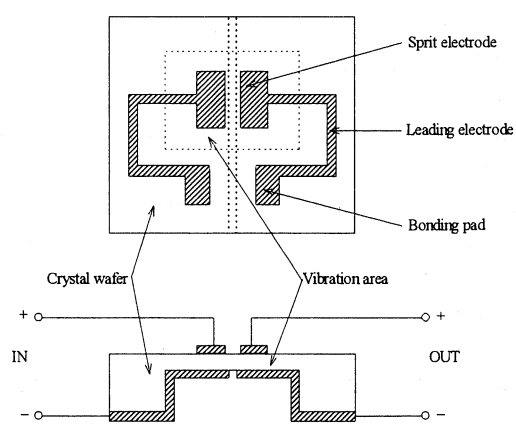


Fig.2 Structure of balanced drive HFF-MCF

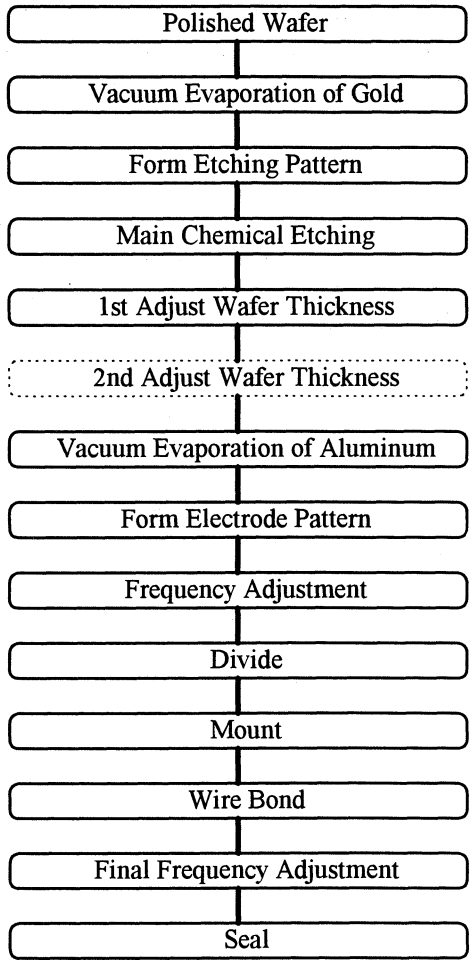


Fig.3 Simplified process sequence

4. 71MHz Band IF Filter for GSM

A balanced drive filter was produced by splitting the reverse side electrode of the unbalanced type 2-pole HFF-MCF using a laser-beam cutting equipment.

Its passband characteristics and group delay characteristics are shown in fig.4 and its stopband attenuation characteristics are shown in fig.5. This filter has 3dB bandwidth of $\pm$90kHz, group delay distortion of 0.7 μ sec at f_0 $\pm$80kHz, insertion loss of 0.9dB and stopband attenuation values of 14dB, 23dB and 40dB at $\pm$200kHz, $\pm$400kHz and $\pm$800kHz from center frequency respectively. It is still necessary to improve the stopband attenuation value at $\pm$1MHz. In this regard, now we continue some additional development to be completed in the near future.

The terminal impedance is 2.5kΩ // -2.0pF for input and output. It is housed in small-sized ceramic package with external dimensions of 5.2 $\times$ 4.5 $\times$ 1.1mm^3 (volume : 0.03cc) that is one tenth as light as those of transversal SAW IF filters now used in GSM system. An external view of the filter is shown in fig.6.

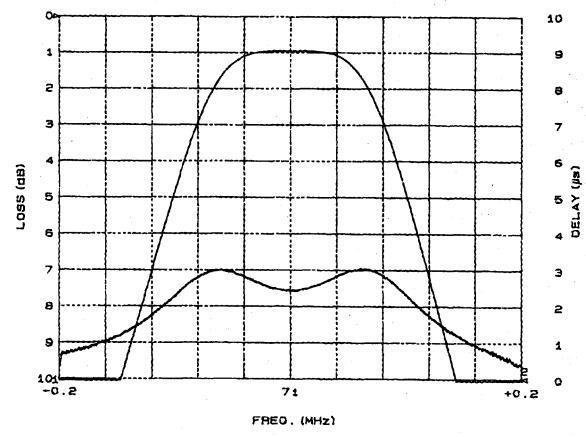


Fig.4 Passband characteristics and group delay characteristics of the balanced drive 71MHz 2-pole HFF-MCF

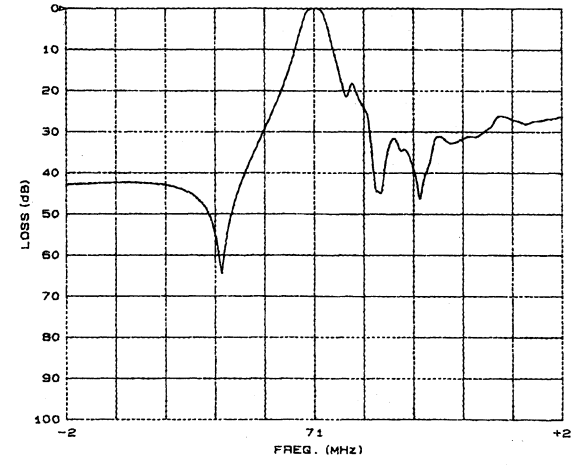


Fig.5 Stopband attenuation characteristics of the balanced drive 71MHz 2-pole HFF-MCF

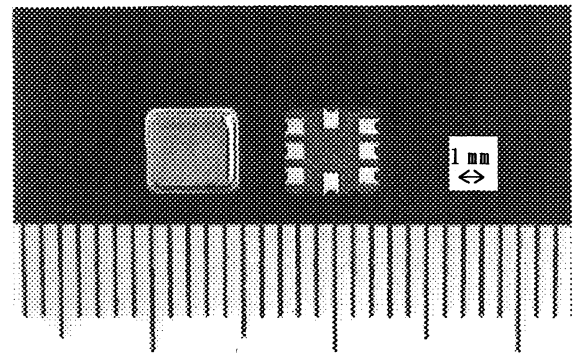


Fig.6 Photograph of the balanced drive 71MHz 2-pole HFF-MCF

5. 250MHz Band IF Filter for DCS-1800

A 250MHz band HFF-MCF for IF filter in DCS-1800 was produced using a production process with an addition of individual etching processes. The characteristics of the unbalanced type 2-pole HFF-MCF at 250MHz are shown in figs.7 and 8. This filter has 3dB bandwidth and insertion loss similar to that of the 71MHz one. It is housed in the package common to that of the 71MHz one.

In the 4-pole configuration in which two elements are connected in cascade, the filter configuration and the measuring circuit are shown in fig.9, and its characteristics are shown in figs.10 and 11. Its 3dB bandwidth is $\pm$80kHz, group delay distortion of 3.7 μ sec at f_0 $\pm$80kHz, insertion loss of 3.4dB and stopband attenuation values is 30dB, 70dB and 80dB at $\pm$200kHz, $\pm$800kHz and $\pm$4MHz from center frequency respectively. The terminal impedance is as low as 200 Ω //1pF for input and output, and the interstage capacitance is 1pF.

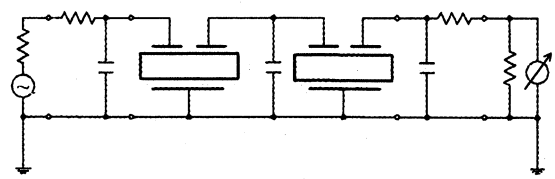


Fig.9 Circuit configuration and measured circuit of the unbalanced drive 250MHz 4-pole HFF-MCF

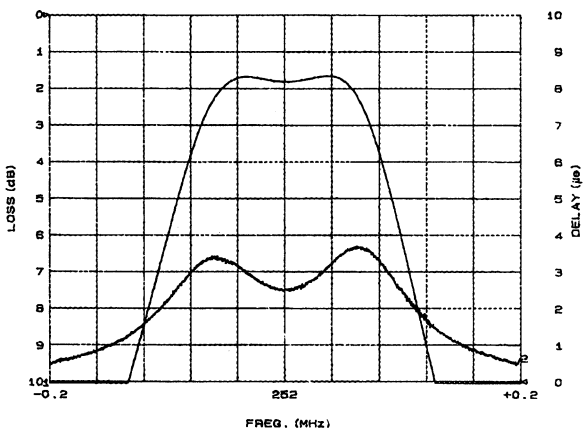


Fig.7 Passband characteristics and group delay characteristics of the unbalanced drive 250MHz 2-pole HFF-MCF

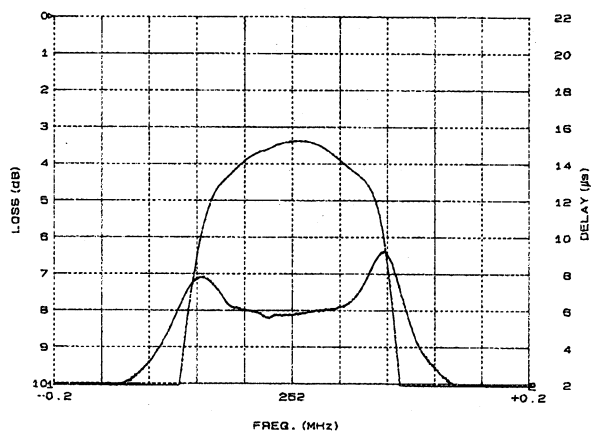


Fig.10 Passband characteristics and group delay characteristics of the balanced drive 250MHz 4-pole HFF-MCF

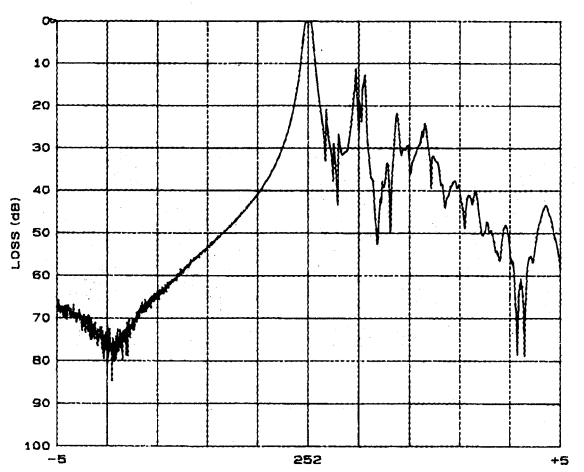


Fig.8 Stopband attenuation characteristics of the unbalanced drive 250MHz 2-pole HFF-MCF

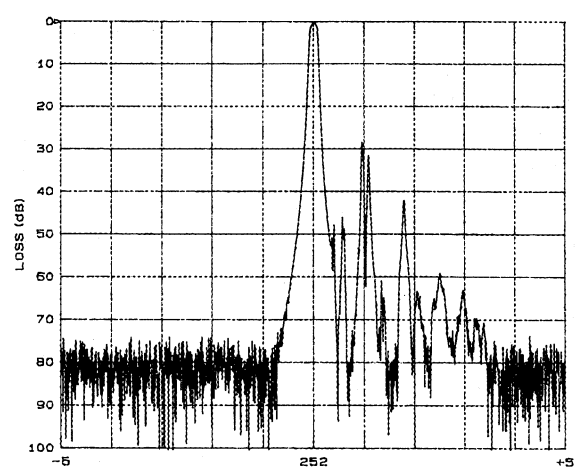


Fig.11 Stopband attenuation characteristics of the balanced drive 250MHz 4-pole HFF-MCF

A large amount of spuriouses occur at the high end of the center frequency, but this is because the configuration uses HFF-MCF elements which have identical electrode dimensions. By constructing it with two elements with differing electrode dimensions, it is considered possible to suppress the spuriouses as above.

6. Conclusion

We developed a 71MHz band IF filter for GSM in the HFF-MCF construction, which has achieved high frequency in fundamental mode using chemical etching process techniques and extended passband width using photolithographic techniques. In the design utilizing the way of dividing the reverse side electrode into input/output terminals, the filter is designed in the balanced drive structure, and as a result, it is constructed in small size (1/7～1/10 in volume) and light weight, with insertion loss (e.g. from 5～10dB to 0.9dB) and suppression in the area of neighbour channel ($f_0 \pm 200$kHz) better than transversal SAW IF filters currently used for GSM .

In the 4-pole configuration in which two 2-pole HFF-MCFs at 250MHz are connected in cascade, we realized IF filters for DCS-1800 base stations requiring sharp selectivity and high stopband attenuation.

References

[1] J.Machui, W.Ruile, "Z-Path IF-Filters for Mobile Telephones", IEEE Ultrason. Symp. Proc., pp.147-150, 1992.

[2] P.Ventura, M.Solal, P.Dufilie, J.M.Hode and F.Roux, "A New Concept in SPUDT Design : the RSPUDT (Resonant SPUDT)", IEEE Ultrason. Symp. Proc., pp.1-6, 1994.

[3] O.Ishii, T.Morita, T.Saito and Y.Nakazawa, "High Frequency Fundamental Resonators and Filters fabricated by Batch Process using Chemical Etching", IEEE Int. Freq. Contr. Symp. Proc., pp.818-826, 1995.

MASTER OSCILLATOR FOR LONG DURATION SPACE APPLICATIONS

Alain Jeanmaire, Pascal Rochat

Tekelec Neuchâtel Time SA - Neuchâtel - Switzerland

1. ABSTRACT

This paper presents a new product for use in the space industry. TNT's « Master Oscillator for Long Duration Space Applications » is a small sized temperature stabilised hi-performance crystal oscillator.
The development of this oscillator has taken into account a wide range of parameters specific for the space environment. The Master Oscillator and its tested performances are described hereafter.

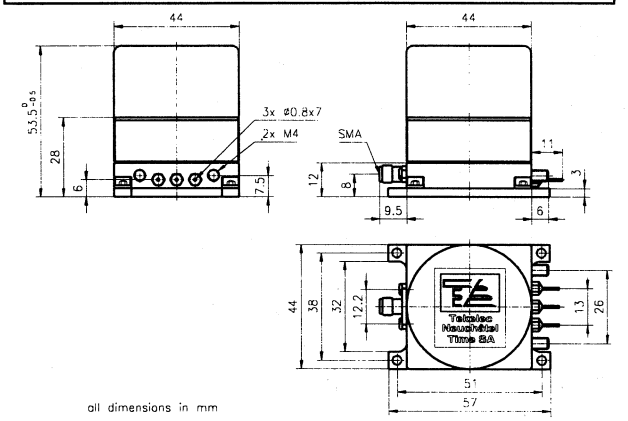
2. INTRODUCTION

All space programs require some kind of stable frequency generating device for successful operation. The required performance naturally depends on the application itself. The final choice will be influenced by spacecraft accommodation (mass, power, thermal etc.) and mission scenario (telecommunication, Earth Observation, Science, etc.). Good short term frequency stability is required for satellite and deep space tracking, communication and navigation. Precise time is essential to precision of navigation. Synchronisation plays a critical role in digital telecommunication systems, as it ensures the transfer of information with a minimum error rate.

3. IDENTIFICATION OF THE CRITICAL ELEMENTS FOR A SPACE ULTRA-STABLE FREQUENCY SOURCE.

These elements are presented in the table below

QUICK REFERENCE



4. GENERAL COMMENTS AND DESIGN GOALS

- For potential applications, the Master Oscillator has to be less cumbersome than the other types of oscillators, without sacrificing its intrinsic performances
- During the development phase of our model, a particular attention has been given to reduce the dimensions (57x44x53.5mm) and the weight (0.2kg).
- Particular attention has been given to get a good radiation shielding and a very high MTBF.

Main parameter	Best choice	+	Advantages / - disadvantages	Space applications
Very long term frequency stability (years)	Cesium	+	Accuracy $\pm 1 \times 10^{-12}$	Navigation systems
		+	Long term stab over years	
		-	Limited life time (5 to 10 yrs)	
		-	Weight	
		-	Expensive cost	
Long term frequency stability over days	Rubidium	+	Freq. stability over day: $<5 \times 10^{-14}$ /day	Navigation, micro-satellite telecom systems and one-way Doppler positioning.
		+	Reduced size and weight	
		-	Very long term stability	
Performance / size ratio and short term frequency stability	Master Oscillator	+	Lightness, 200 grams	Telecom and one-way Doppler positioning
		+	Excellent freq. stability for 1-100sec.	
		+	Very small size, 0.15 L	
		+	Low cost	
		-	Long term ageing	

5. MAIN NEW FEATURES OF TNT'S MO

During the bread boarding phase of TNT's MO, a special care has been given to the choice of the parts, in order to get the lowest phase noise as possible. Then, the electronics design has been optimised using low noise transistors and efficient automatic gain control of the oscillating loop.
After two phases of mechanical and thermal design and simulation, the final MO configuration has been fixed to have the best thermal stability for the specified low power consumption.
The choice of the material allows an important radiation shielding of the crystal.
This component oriented space crystal oscillator can be directly used inside an instrument. It can support any depressurisation down to 10^{-9} bars with very low outgassing.
The Master Oscillator does not radiate electro-magnetic field at any other frequency than 10MHz.
The Master Oscillator has been designed for operation, with full stability performances, even under pressurised or vacuum environment. This unique feature allows the MO to be tested within the payload under atmospheric or pressurised environment.

6. ORIGINAL SOLUTIONS

TNT's Master Oscillator is the result of a General Supporting Technology Program (GSTP) of the European Space Agency (ESA). This means that the latest applicable technology has been used in order to make a state-of-the-art oscillator. This is the key which helped to reach the requested high performance in such a small volume and mass.
The temperature sensitivity has been one of the parameters which required the largest numbers of special solutions. The use of a Stress Compensated crystal resonator working at Turn Over Point was the unique solution. The power consumption has been reduced by filling the hermetic enclosure with low thermal conductivity gas. This relatively high pressure gas also eliminates any possible outgassing effects from components and/or materials within the enclosure. Consequently, this configuration stabilises the thermal resistance between the two active ovens.
The temperature stabilisation of the critical components and the optimised AGC was an important factor to the performance. The low phase noise performance is the result of a transistor type selection.
The radiation tolerance has been improved by increasing the radiation shielding to the maximal allowed mass (200 grams).
TNT's Master Oscillator has no DC/DC converters and can be used as a component inside an instrument, directly connected to the instrument supply. It provides an ultra-stable voltage reference output for frequency adjustment associated with the payload or crystal ageing compensation.

7. BLOCK DIAGRAM

The block diagram of figure 1 shows the working principle of the oscillator.
It is particularly interesting to note the following points:

- The configuration has four levels:
  1) External interface, which is not hermetic and accessible by the user at anytime.
  2) Hermetic enclosure, also called Master Oscillator Core, which is vacuum proof sealed.
  3) First active oven, at stabilised temperature.
  4) Second active oven, with a better temperature stabilisation (± 0.01°C over the temperature range).
- The frequency and the temperature can be adjusted directly on the external interface.
- The "two supplies" or "single supply" option can be selected on the interface.
- The temperature references are thermally stabilised.
- The power limitation can be adjusted on the interface.
- The internal voltage reference is in the inner oven for the best thermal stabilisation (<0.01ppm/°C).

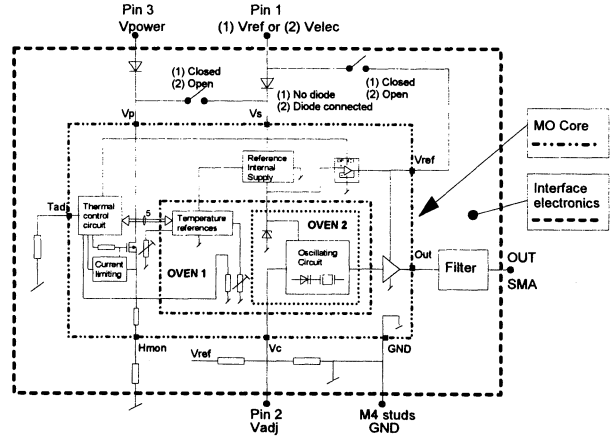


Figure I: Functional diagram of the Master Oscillator.

8. TEST RESULTS

8.1 INTRODUCTION

During the development phase, two bread boards and three Engineering Models have been manufactured. A third Engineering Model has also been manufactured for testing alternative crystal supplier source.
ESA has selected the EM#2, which contains a CQE 10MHz crystal for the Qualification model because a BVA Q007 SC 3rd overtone crystal capable of surviving the vibration specification was not available within schedule of the program. The results presented in this section have been tested on the qualification model equipped with a CQE crystal.

8.2 PHASE NOISE

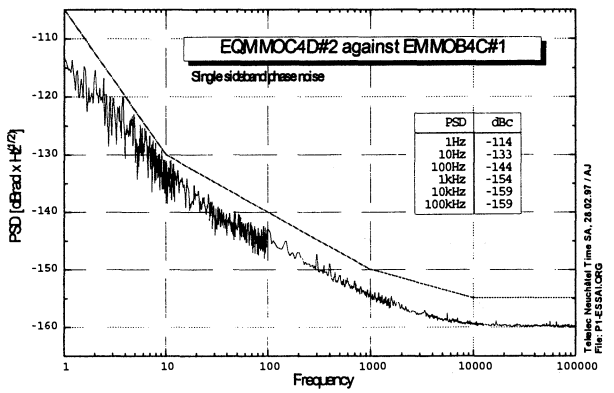


Figure II: Phase noise measurement of the EQM.

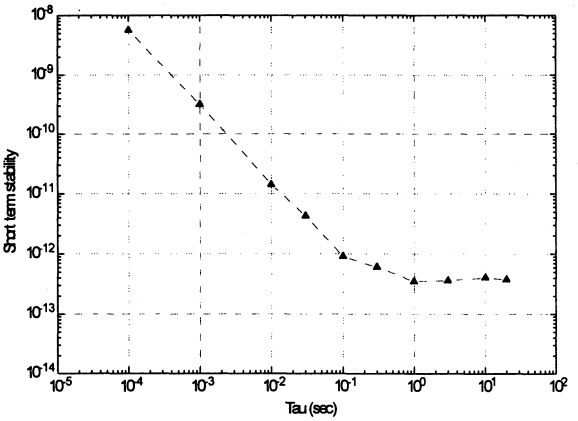
Distance to 10MHz	EQM dBc
1 Hz	-114
10 Hz	-133
100 Hz	-144
1 kHz	-154
10 kHz	-159
100 kHz	-159

Table II: Phase noise results of EQM

This is a typical phase noise curve which has been measured with a Stanford SR760 FFT analyser.

This low phase noise has been measured against another Master Oscillator (EM#1).
The theoretical phase noise of the MO is devoid of spuriouses while there are no other frequencies than 10MHz inside the oscillator (no DC-DC converters for example).

8.3 SHORT TERM STABILITY



Here is a typical short term stability curve computed by the FSMMS system of ON from a set of phase noise measurement data. The curve is quite identical for all three models.

8.4 LONG TERM STABILITY

The long term curve presented here often has been measured on the EM#3 over a three months period through an automatic measurement system. This curve gives the long term figure at the date of this publication release.

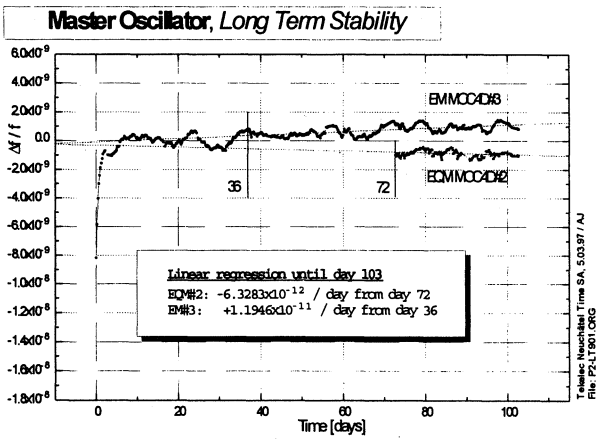


Figure IV: Long term measurement of the EM MO#3

The measured long term figure is -6.3×10^{-12}/day.
Please note that this measurement has been made on EM#3, at atmospheric pressure (oscillator not sealed). Better results are expected on hermetic oscillators.
No reliable ageing figure is available for the EQM as it has not been measured over a long enough period.
Depending on the fit, the slope is between 1.9×10^{-11} and 0.43×10^{-11}/day.

8.5 THERMAL STABILITY

Reducing the temperature sensitivity to $5x10^{-10}$ over a temperature range of -15/+60°C in such a small volume and low power has been the biggest challenge of the project. The cut of the crystal is a critical factor for the temperature sensitivity. Stress Compensated (SC) cut crystals have the advantage of offering a turn-over point (TOP) in the frequency versus temperature curve. If the crystal is maintained at TOP, small temperature variation will have a quasi-null frequency variation. To reach the $2x10^{-12}$ /°C goal, a temperature variation of less than 0.02°C when stabilised is necessary in order to compensated possible drifts of the temperature reference during the oscillator's life. The success of the temperature stabilisation within this small range is the result of many thermal tests and simulation with computer models. The final configuration is a trade-off between the distribution of thermal conductance, insulation and heating power between the two active ovens and the external housing.

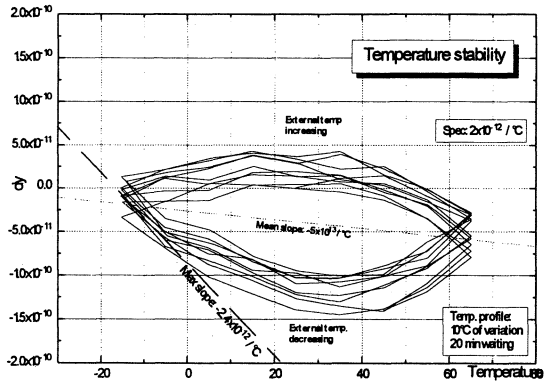


Figure V: frequency temperature sensitivity curve of one EM MO. Triangular cycles from -15°C to 60°C.

The residual hysteresis is due to the speed of the measurement: the waiting time between each point was only 20 minutes. Another measurement, made on the EQM, is presented in figure VI. The hysteresis is smaller, whereas the waiting time after temperature variation was longer.

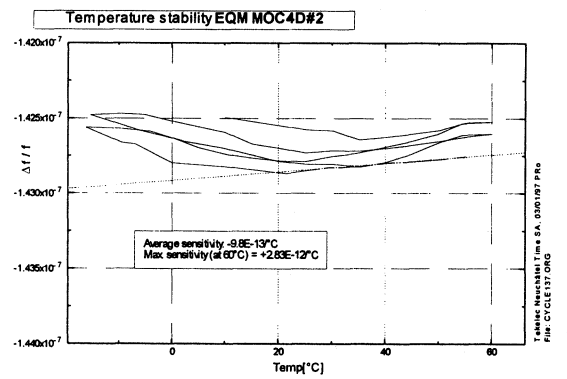


Fig. VI, an other temp. stabilty measurement

The temperature adjustment is defined by two resistors in the external interface. Therefore, it can be changed at anytime, even after the sealing of the hermetic enclosure. To regain the maximal thermal performance, the temperature can be readjusted by changing the two resistors. The adjustment value is determined by an automatic TOP adjustment program.

For qualification, the Master Oscillator is also tested according to a thermal vacuum qualification profile as show in figure VI.

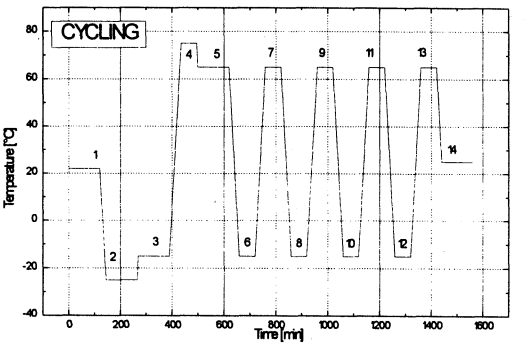


Figure VII: Thermal vacuum temperature profile.
The profile is defined in order to confront the Master Oscillator with all possible configurations.
During the test, the power is stopped and switched ON four times. Start: 1/ 2/ 3/ 5, Stop: 1/ 2/ 3/ 14.

8.6 RADIATION SENSITIVITY

Two models have been submitted to a Co_{60} low dose radiation test, 1 Rad/h at CNES / Toulouse.
The measured sensitivity of Master Oscillators with CQE crystal is typically: $\Delta f / f < 6.0x10^{-12}$ /Rad.
It is important to remind that the crystal resonators is more sensitive to low radiation dose rate.
Radiation shielding has been calculated for min. 10 years of operation in geostationnary orbit.

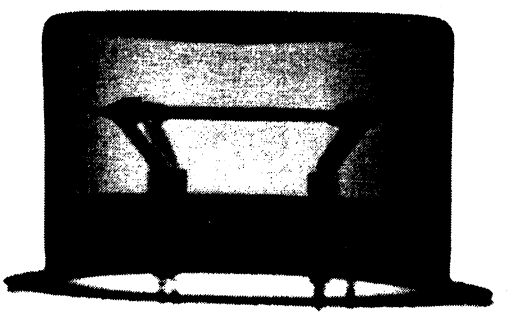


Figure VIII: Radiography of a CQE crystal resonator.

This radiography shows how the crystal plate is mounted inside its H40 enclosure.

9. PERFORMANCE SUMMARY

Type MO4D-SC	
Parameter	Value
Dimensions	44 x 54 x 57 mm see drawing in page 1
Output signal frequency	10MHz
Freq. long term stability, 1st year years after	 <$2x10^{-11}$ per day <$1x10^{-8}$ per year
Freq. short term stability	<$5x10^{-13}$ (1-100sec)
Stability over full temp. range	<$2x10^{-10}$
Frequency adjustment	~2 Hz
SSB phase noise assuming 10MHz carrier 1 Hz 10 Hz 100 Hz 1000 Hz 10000 Hz	 -114 dBc -133 dBc -144 dBc -154 dBc -159 dBc
Output signal level range	5 dBm ± 1dBm
Output impedance	50 Ω
Harmonics	-45 dBc
Spurious signals	-120 dB
Load impedance	50 Ω
Warm-up time to nominal performance	120minutes $5x10^{-9}$
Power consumption during warm-up	5 W
Nominal power consumption, vacuum, at 25°C	 2.5 W
Mass (Master Oscillator)	200 g
Volume	0.15 dm^3
Power supply	12 - 18 V
Connectors 1) Power / TC 2) RF Output 3) Case ground	3 solderable pins SMA 2 x M4 stud
Mechanical interface	flat base plate
Radiation sensitivity	$6.0x10^{-12}$ /Rad(Si)
Life time	15 years
Pressure	Amb $\rightarrow 10^{-9}$ bars

Type MO4D-SC		
Parameter		Value
Max. baseplate operating temp.		60 °C
Min. baseplate operating temp.		-15 °C
Non operating temperature		-40 to 70 °C
First natural resonance		> 2kHz
Random Vibration tested, with axis perpendicular to the mounting plane. During	20 - 80 Hz 80 - 350 Hz 350 - 443 Hz 443 - 600Hz 0.6 - 2Hz	+6 dB/oct 0.8 g^2/Hz -6 dB/oct 0.5 g^2/Hz -6 dB/oct 120 sec/axis
Random Vibration tested, with axis parallel to the mounting plane. During	20 - 80 Hz 80 - 350 Hz 350 - 443 Hz 443 - 950 Hz 0.95 - 2 kHz	+6 dB/oct 0.32 g^2/Hz -6 dB/oct 0.2 g^2/Hz -6 dB/oct 120 sec/axis
Sinusoidal vibration	5 - 19 Hz 19 - 80 Hz 80 - 100 Hz	11 mm 0-peak 16 g 8 g

Remarks:

When better short or long term stability performances are required, the MO could then be equipped with a 5MHz 3rd overtone crystal resonator.
The warm-up maximal power can be adjusted on demand for faster warm-up or reduced maximal power.
The internal voltages are relative to a reference bridge, so the power voltage requirement can be adjusted.

10. ACKNOWLEDGEMENT

- This work has been supported by a ESA/GSTP programme. We would like to thank Mr. S. Feltham for his special support, help and advice.
- We also would like to thank Mr. Brunnet and Mr. Guillemot of CNES for their help in performing the radiation test measurement on two MO.
- A very special thank you to Mr. G. Busca at the Observatoire Cantonal de Neuchâtel without whom this project would not have been possible.

11. REFERENCES

- R. Norton, "Performance of Ultra-stable Quartz Oscillator Using BVA Resonators", Proceedings 8th EFTF, pp. 457-465, March 1994
- Emmons, D.W. Martin and R.M. Garvey, "Satellite Master Oscillators", Proceedings of the 1995 IEEE Internal Frequency Control Symposium, pp. 600-613, June 1995
- Vig, "Quartz Crystal Resonators & Oscillators for Frequency Control & Timing Applications", Jan. 1988
- G. Busca, P. Rochat and B. Schlueter, "Study of potential applications for atomic clocks in space", ESA study, Observatoire de Neuchâtel, Aug. 1994.

DUAL-MODE CRYSTAL OSCILLATOR

Anatoly V. Kosykh, Sergey A. Zavialov

Omsk State Engineering University. Russia

Dual-mode oscillator is an oscillator based on simultaneous excitation of resonator two own modes. SC-cut resonators with B- and C- modes excitation are used as dual-mode most often. Each resonator mode has its own frequency vs. temperature dependence. When a resonator is being excited on two modes simultaneously, the mode frequency ratio is uniquely determined by temperature. This unique relation allows to use C- and B- mode frequency ratio as a resonator proper thermosensor. This thermosensor has full spatial coincidence with a vibrated resonator plate and, therefore, it is an ideal resonator temperature measuring device. Due to quartz nature such sensor is very precise and stable. From the above reasoning dual-mode resonator is an ideal base for thermocompensated (TCXO) and ovenized (OCXO) oscillators, in the cases when we need to measure the quartz plate proper temperature with maximum accuracy.

The concepts of dual-mode oscillators design have been discussed in this work. Considerable attention has been given to the constant B-mode excitation level keeping under external factors effects and resonator motional resistance variation. The goal of the work submitted here is development of suitable for production dual-mode oscillators of small size intended for DTCXO or OCXO. The following criteria have been observed during oscillator development:

- minimum power consumption;
- minimum sensitivity to supply voltage variations;
- minimum of spurious spectrum components in output C-mode signal;
- the absence of large size elements unaflectable to minimization (e.g tuned inductance coils);
- the stable dual-mode oscillation with range adjustable B-mode level ability must be provided by the oscillator circuit;
- possibility of micro-chip realization.

When dual mode oscillator use in DTCXO of a traditional type (on the basis of voltage controlled oscillator) it is required to supply frequency shift in limits up to 100 ppm. It is a great problem, as under frequency tuning the condition of amplitudes balance is essentially change. It may be a oscillation failure when controlling voltage or external temperature has reach a some value.

The circuit of the dual-mode oscillator, satisfying to set requirements is offered and investigated. The modified SC-cut resonator was used in the oscillator. The oscillator represents a modified Colitis circuit with the additional dual- mode stabilization units. The considered oscillator was used in DTCXO of 0.1 ppm in an -40 ... + 85 OC interval of temperatures. The results of the oscillator circuit experimental investigation submitted in this work are:

- the frequency versus controlling voltage dependence ;
- output signal spectrograms;
- diagrams of regime instability.

COM PARAMETER EXTRACTION FOR STW RESONATOR DESIGN

Svetlana Krasnikova*, Benjamin P. Abbott, Rodolfo C. Almar

Sawtek,Inc., P.O. Box 609501, Orlando FL USA 32860-9501

*Formerly with the Institute "Dalniaya sviaz", St.Petersburg, Russia

ABSTRACT

This paper discusses a technique for extracting coupling-of-mode parameters for surface transverse acoustic wave (STW) devices. The empirical values of COM parameters are presented, along with a comparison of the theoretical resonator performances calculated using these values and measured responses. To ensure the reliability of the parameter extraction results, other devices in addition to synchronous one-port resonators were used as test structures. Based on the parameters extracted from measurements for different values of metal thickness, expressions relating COM parameters to the metal thickness were obtained .

Keywords: COM parameters for STW devices

1. INTRODUCTION

Recently, much attention has been paid to the design techniques of surface transverse wave (STW) resonators. The growing interest in these devices arises from their advantages over conventional SAW resonators (higher velocity and power handling capability, good temperature stability).

The modeling of STW devices has been discussed in a large number of papers (Refs. 1-7). In publications (Refs.1-5), the COM formalism for STW is developed. In (Ref.3), the derivation of coupling-of-mode equations for STW is considered. A dispersion relation in parameteric form is obtained in (Ref.4). In (Ref.2), a closed form dispersion relation is derived and the reflection and transmission coefficients for a grating are obtained, while for the IDT area the conventional SAW COM formalism is used. It seems to be more consistent to model all parts of a device using the same formalism. The simple form of the closed form dispersion relation (Ref. 4) makes it possible to implement it (or its approximated version) in the conventional SAW COM formalism (Refs.1,5).

Although alternative approaches to the STW devices modeling are also proposed (Refs.6,7), the coupling-of-mode theory is proven to be a very convenient instrument for the STW resonator design. When employing the COM formalism, the exact values of model parameters and their variation with metal thickness are of great importance since these values are responsible for the accuracy of the analysis. There are some analytical methods for estimating COM parameters for STW, but experimental techniques still remain more reliable. Previously it was found, that COM parameters can be obtained from the measured admittance of synchronous one-port resonators. In these devices the resonance is always placed on the left edge of the stopband. In STWs, energy trapping varies with frequency, and it is of interest to examine the resonator structures with arbitrary position of the resonance in the reflector stopband.

In this work, model parameters were obtained experimentally from measurements of responses of synchronous and non-synchronous one-port resonators, long transducers and two-port resonators.

2. COM MODEL FOR STW

For modeling STW devices, the COM formalism was used in which the conventional SAW dispersion relation is replaced by the dispersion relation more appropriate for STW. This approach produces results similar to the approach described in (Ref. 1). However, the parameter extraction procedure is simplified.

The dispersion relation is written in the form:

$$D = \sqrt{\Delta^2 - \kappa_1^2} \qquad (1)$$

where

$$\Delta = \frac{2\pi}{V} f - \frac{\pi}{p} - j\gamma_1$$

V - STW velocity

p - period of the grating

$\gamma_1 = \frac{\gamma}{2p}$ - attenuation coefficient

$\kappa_1 = \left|\frac{\kappa}{2p}\right|$ - reflection coefficient

f - frequency

Although the form of (1) is exactly the same as in the conventional COM model there is a difference between dispersion relations for STW and SAW :

·An additional frequency dependent attenuation appears above the cut-off frequency in the STW dispersion relation caused by SSBW↔STW interactions.

·The reflection coefficient κ is pure imaginary for STW and its value is quadratically proportional to $\frac{h}{\lambda}$.

The above features follow from the comparison of (1) and the dispersion relation derived in (Ref. 2):

$$D = \sqrt{\Delta_1^2 - \tfrac{1}{4}(\varepsilon_1^2 + \eta_1\sqrt{\varepsilon_1^2 - \eta_1^2 - 4\Delta_1})} \qquad (2)$$

where

$\Delta_1 = \frac{2\pi}{V_s} f - \frac{\pi}{p}$

$\varepsilon_1 = \sqrt{\frac{\pi}{p}}\,\varepsilon$

$\eta_1 = \sqrt{\frac{\pi}{p}}\,\eta$

V_s - SSBW velocity

The quantities ε, η may be considered as some phenomenological parameters (Ref. 2) or they may be expressed in terms of material constants and metal thickness (Ref. 5).

The analogy of (1) and (2) becomes clear if the following substitutions are made:

$$\kappa = j\pi\varepsilon\left(2\eta + \varepsilon - \frac{\varepsilon^2\eta}{\eta+2\varepsilon}\right) \qquad (3)$$

$$V = V_s(1 - \Delta_V\kappa) \qquad (4)$$

$$\gamma = \begin{cases} \gamma_0, & f < f_u \\ \gamma_0 + \frac{\gamma_B\eta_B\sqrt{f-f_u}}{f-f_u+\eta_B^2}, & f > f_u \end{cases} \qquad (5)$$

where

$f_u = f_0(1 + \Delta_u)$

$f_0 = \frac{V_0}{2p}$

Quantities Δ_V, f_u, γ_B, η_B and f_u may by expressed in terms of κ.

Formula (3) shows a quadratic dependence of κ on metal thickness since ε and η are linearly proportional to the metal thickness (Ref. 5). The frequency dependence of γ is described by (5).

In this work κ, V and γ_0 are considered as the model parameters to be extracted from measurements. The fourth parameter of the COM model, the transduction coefficient α is required for the modeling of the IDT area. In the case of STW this parameter also depends on metal thickness (Ref.1)

3. PARAMETER EXTRACTION TECHNIQUE

COM parameters have been evaluated from the measurements of amplitudes of $S_{11}^{meas}(f)$ and $S_{12}^{meas}(f)$ for one- and two-port resonators respectively.

To extract COM-parameters from measurements the following function was defined:

$$F(\kappa, \gamma_0, \alpha, V) = \sum_{n=1}^{N}\left(\left|S_{ik}^{calc}(\kappa, \gamma_0, \alpha, V, f_n)\right| - \left|S_{ik}^{meas}(f_n)\right|\right)^2$$

A set of parameters corresponding to a minimum of $F(\kappa, \gamma_0, \alpha, V)$ was considered as a "true" set of COM parameters for a given metal thickness. The search for a minimum starts at an initial set of parameters estimated theoretically.

The frequency responses of resonators $S_{11}^{calc}(f)$ and $S_{12}^{calc}(f)$ were calculated in a common way by cascading P-matrices of gratings and transducers with parasitic components taken into account. Note that the imaginary reflection coefficient κ results in the asymmetry of P-matrices.

4. DESCRIPTION OF TEST STRUCTURES.

The design parameters of the one-port and two-port devices are shown in tables 1 and 2, respectively. The IDT period is 5.06μ in all devices , the number of electrodes in the gratings of the one-port devices is 500, the aperture of the two-port resonators is 100 wavelengths except for devices #12 (50 wavelengths) and #13 (25 wavelengths), the gap between the reflector and the IDT in the two-port devices is half a wavelength. All devices were fabricated on quartz with metal thicknesses of 1000A, 1500A, 2000A, 2500A and 3000A.

Table 1.

#	Period, microns	Number of electrodes in IDT	Aperture, wavelengths
1 (IDT)	-	200	100
2	5.06	200	100
3 (IDT)	-	200	50
4 (IDT)	-	200	25
5	5.06	200	50
6	5.06	200	25
7	5.09	200	100
8 (IDT)	-	100	100
9	5.06	100	100
10	5.12	100	100

Table 2.

#	Reflector		Num. of electr. in IDT	IDT-center grating gap, wavelengths	Center grating	
	Period, microns	Num. of el.			Period, microns	Num. of el.
11	5.09	300	100	0.514	5.09	90
12	5.09	300	100	0.514	5.09	90
13	5.09	300	100	0.51	5.09	90
14	5.1	300	90	0.524	5.1	30
15	5.12	300	80	0.534	5.12	10
16	5.08	400	180	0.504	5.08	180
17	5.13	490	45	0.5	5.06	32
18	5.06	600	150	0.617	4.8	11
19	5.06	600	150	0.617	4.8	100

5. EXPERIMENTAL RESULTS AND DISCUSSION

In most cases a set of COM parameters was obtained such that the calculated response fits well the measured one in the frequency range below the cut-off frequency for STW↔SSBW interactions. Examples of responses calculated with these extracted quantities are shown in Fig.1 and 2. Above the cut-off frequency there is a significant discrepancy between the experimental and the theoretical results. This shows that the model does not work well in the upper frequency range, which can be explained by the following reasons:

·The frequency dependence of the attenuation resulting from STW dispersion relation does not fully account for SSBW↔STW interactions.

·The electromechanical coupling coefficient, which was supposed to be constant, in fact tends to diminish with frequency in the upper frequency range.

·The contribution of the SSBW generation into the IDT admittance is not taken into account.

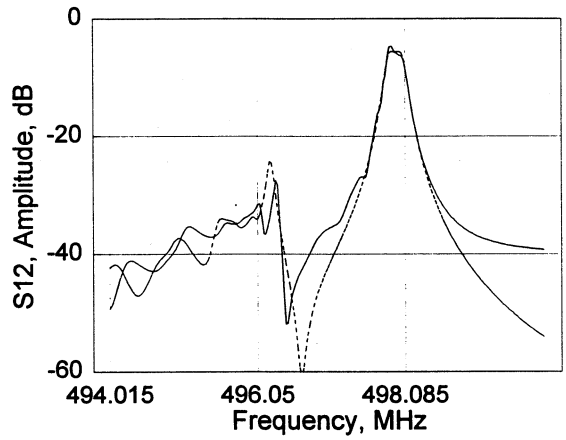


Fig .1 Measured (solid line) and calculated (dashed line) responses of sample 11, metal thickness 2000A.

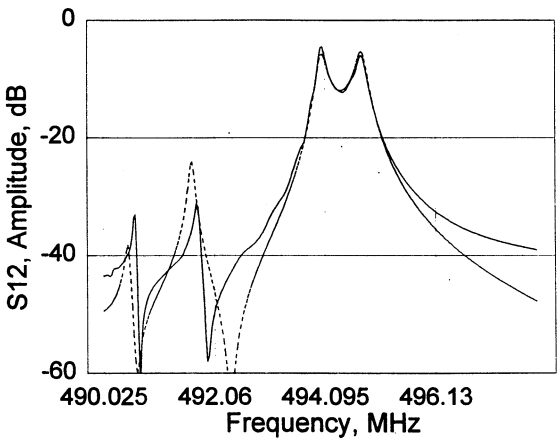


Fig.2. Measured (solid line) and calculated (dashed line) responses of sample 15, metal thickness 2500A.

An example of a response in a wide frequency range is given in Fig.3. As an illustration examples of responses calculated with corrections made for the attenuation and electromechanical coupling coefficients in the upper frequency range are shown in Fig.4 and 5.

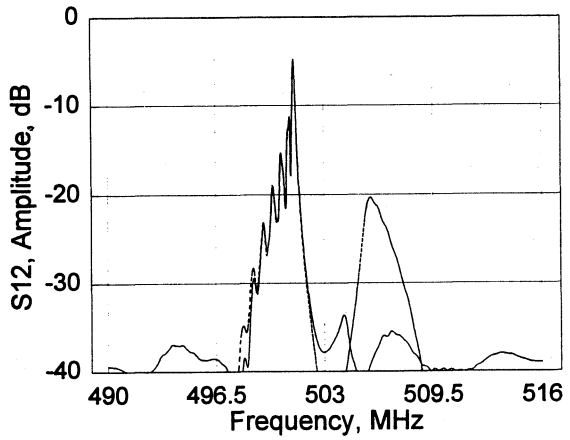


Fig.3. Measured (solid line) and calculated (dashed line) responses of sample #18, metal thickness 1500A.

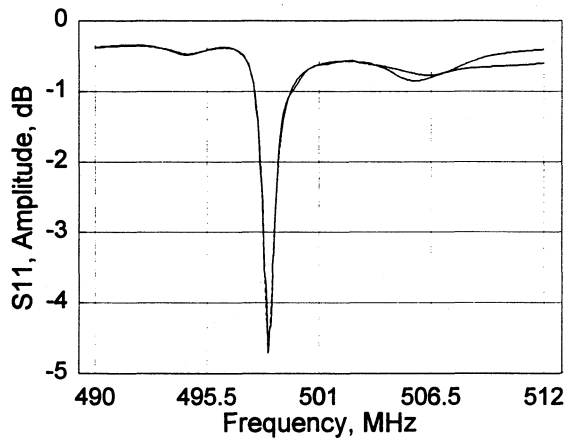


Fig.4. Measured (solid line) and calculated (dashed line) responses of sample 1, metal thickness 2000A.

To avoid the complexity related to STW↔SSBW interactions only the portion of a frequency response below the cut-off frequency was used in the parameter extraction procedure.

The parameter extraction results are summarized in Fig.6-8, where experimental points correspond to quantities extracted from measurements (all designs, metal thickness 1000A, 1500A, 2000A, 2500A, 3000A). As it was expected, the reflection coefficient and the electromechanical coupling coefficient increase and the velocity decreases with increasing metal thickness.

Data for the attenuation coefficient are rather scattered. In a number of papers it was pointed out that the position of the resonance must be limited to the lower part of the stopband, otherwise the insertion loss of the resonator will be high because of scattering energy into the substrate when the resonance is in the higher half of the stopband. The experimental results correlate well with this suggestion while the model makes no prediction about it. As a rule, the attenuation is high when the resonant frequency is placed in the higher half of the stopband of the reflector.

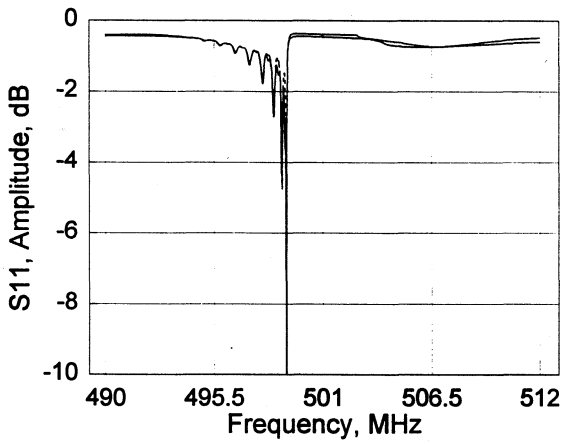


Fig.5. Measured (solid line) and calculated (dashed line) responses of sample 9, metal thickness 2000A

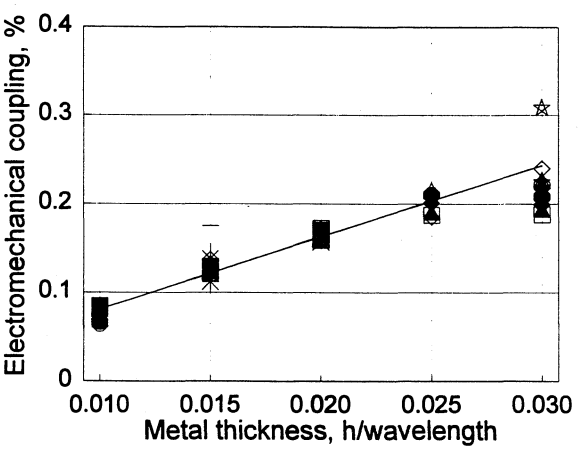


Fig. 6. Electromechanical coupling coefficient versus metal thickness.

To relate COM parameters to metal thickness it was assumed that the reflection coefficient, the electromechanical coupling and the velocity can be expressed in terms of metal thickness as follows:

$$\kappa = \kappa_0 \sin(\pi\Theta)\left(\frac{h}{\lambda}\right)^2 \quad (6)$$

$$k^2 = k_0\Theta\left(\frac{h}{\lambda}\right) \quad (7)$$

$$V = V_s(1 - \Delta_V \kappa) \quad (8)$$

where Θ - metallization ratio.

κ_0 ,k_0 , V_s and Δ_V were estimated by least square fit of COM parameters obtained from measurements for different values of metal thickness:

$k_0 = 0.165$; $\kappa_0 = 109.6$; $V_s = 5098.4$; $\Delta_V = 0.055$

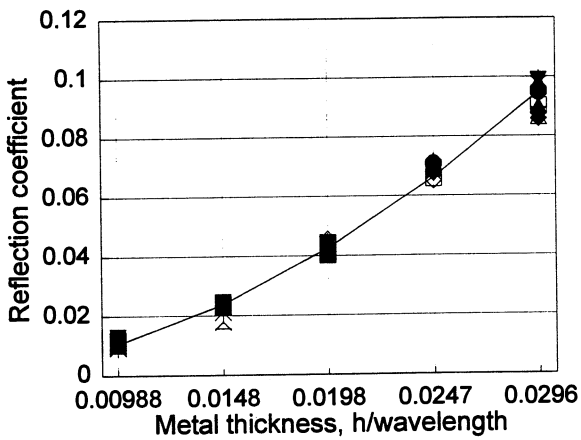


Fig. 7. Reflection coefficient versus metal thickness.

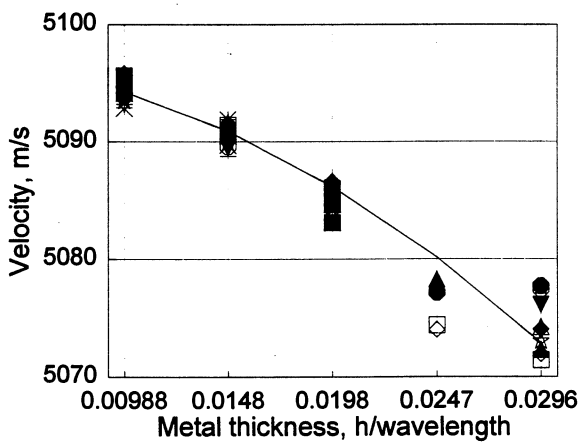


Fig. 8. Velocity versus metal thickness

When the COM parameters evaluated according to (6) - (8) are used in the frequency response calculations the agreement between the theoretical and experimental frequency responses is about the same as with COM parameters obtained for a particular device.

6. CONCLUSIONS

COM parameters are obtained from measured data for 500 MHz one-port and two-port resonators with different design parameters and different values of metal thickness. Based on the extracted quantities the reflection coefficient, the electromechanical coupling coefficient and the STW velocity are expressed in terms of the metal thickness. In the frequency range below the cut-off frequency the theoretical frequency responses obtained with these calculated COM parameters are in a good agreement with experimental data for most devices.

Above the passband the agreement between theory and experiment is rather poor. The reason is that the frequency dependencies of the attenuation and electromechanical coefficient are not taken into account properly. This suggests that to improve the model, a way has to be found to implement correct frequency dependencies of the attenuation and electromechanical coupling coefficient.

ACKNOWLEDGEMENTS

This work was partially supported by the U.S. Department of Commerce, grant 8850-95-8A38 under the Special American Business Intership Program.

REFERENCES

1. B.P.Abbott, K.Hashimoto, Ultrasonics Symposium, 1995, pp. 222-229.

2. V.P.Plessky, Ultrasononics Symposium, 1993, pp. 195-200.

3. D.F.Thomson, B.A.Auld, Ultrasonics Symposium, 1986, pp. 261-266.

4. E.Bigler, B.A.Auld, E.Ritz, E.Sang, Proc. 45th Ann. Freq. Cont. Symp., 1991, pp. 222-229.

5. B.I. Boyanov, K.D.Djordjev, V.L. Strashilov, I.D.Avramov, Ultrasonics Simposium, 1995, pp.317-320

6. K.Hashimoto, M.Yamaguchi, Ultrasonics Symposium, 1994, pp. 253-258.

7. C.A.Flory, R.L.Baer, Ultrasonics Symposium, 1987, pp. 313-318.

THEORETICAL INVESTIGATIONS TO DESIGN NEW TEMPERATURE SENSORS

T.Messaoudi

Laboratoire de Chronométrie, Electronique et Piézoelectricité (LCEP)
Ecole Nationale Supérieure de Mécanique et des Microtechniques
26 Chemin de l'Epitaphe
25030 BESANÇON CEDEX - FRANCE
E-mail : tijani@ens2m.fr **Fax** 0381885714

1. ABSTRACT

In this paper an attempt is made to investigate theoretically the thermal sensitivity and the trapped energy of new bulk acoustic waves (BAW) micro-resonators operating essentially in thickness modes. As a result, frequency temperature coefficients as well as mechanical coupling contour plots are presented.
In conclusion these investigations allow us to select some cuts satisfying the trapped energy condition for which a good thermal sensitivity is obtained.

Keywords : Temperature Sensor, Frequency-Temperature Curves, BAW, Effective Elastic Constants, Trapped Energy, Micromachining, Thickness Modes.

2. INTRODUCTION

Temperature sensors based on the variation of resonance frequency with temperature such as quartz plate resonators or tuning forks can be developed. Up to now applications are limited to the doubly rotated cut (LC) and to singly rotated cuts (θ=40° and θ=0°) (Ref. 1). Among these vibrating structures only tuning forks are usually fabricated by the so-called micromachining process. This fabrication process can be also conveniently applied to vibrating plates.

Attempts are made to develop new temperature micro-sensors which use frequency output and have good accuracy. In order to have good metrological performances namely frequency stability and high quality factor as for conventional resonators we can use :

1. suspension bridges to reduce undesirable effects (e.g. axis force, ...) due to fixation process of the vibrating structure.
2. trapped energy by mesa or hole in the middle of quartz plate.

Suspension bridges and mesa can be fabricated by a chemical micromachining process.
However, here we publish only the various theoretical results concerning the thermal sensitivity and the trapped energy of new BAW micro-resonators in doubly rotated cuts.
To design a new thermal sensor, we need to know precisely the temperature sensitivity of the BAW vibrating plate, that is to say, to determine the first, the second and the third order temperature coefficients. Calculations are performed in the special case of an endless perfectly planar plate which vibrates in thickness modes.
Here we have to point out that doubly rotated cuts certainly offer various possibilities. Moreover, before designing a vibrating structure, we need to evaluate its optimum dimensions by considering essentially the trapped energy principle.

3. THERMAL SENSITIVITY OF BAW VIBRATING PLATE IN DOUBLY ROTATED CUTS AND OPERATING IN THICKNESS MODES

We adopt a method which involves the Piola - Kirchhoff stress tensor (Refs. 2, 3). To derive quartz effective elastic constants for a given temperature, we use conventional Taylor developments. As a result, when we express the resonance frequency, only this tensor depends on temperature. Contour plots of first, second and third order temperature coefficients of frequency versus the angles of cut (φ, θ) can be then numerically fitted.
Let a propagation of an elastic wave along a thickness of the quartz plate in doubly rotated cuts, which is submitted to a static deformation due to homogeneous temperature change (figure1).

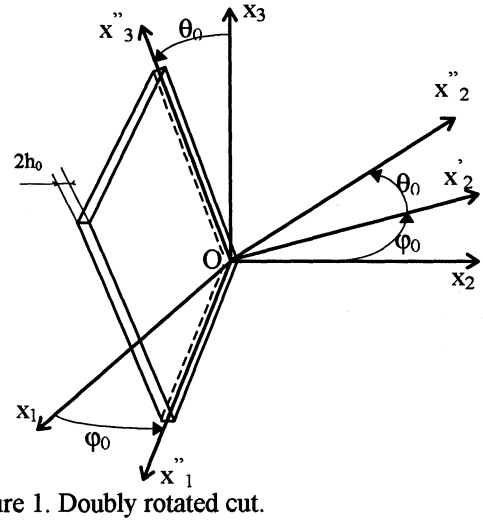


Figure 1. Doubly rotated cut.

Direction of propagation is made along the (ox_2") axis of the mobile coordinates system (ox_1"x_2"x_3") defined in reference state of plate. We distinguish three states described in (Refs. 2, 3, 6) : a reference state (at fixed reference temperature T_0), an intermediate state (at a current temperature T and prior to vibration) and a final state (resulting from the combinaison of the adiabatic vibration and the intermediate state). $2h_0$ is the thickness of this plate and ρ_0 its mass density at reference configuration. The elastic wave is characterized by the following dynamic fields (Ref. 7) :

$$(1)\qquad \begin{cases} u_\alpha = u_\alpha^\mu \exp[j(\omega t - x_2''/v^\mu)] \\ \tilde{\Phi}_2 = \tilde{\Phi}_2^\mu \exp[j(\omega t - x_2''/v^\mu)] \end{cases}$$

where v^μ is the velocity of the wave at instant t, related to the reference configuration. The subscript μ identifies symbolically the mode of vibration (Ref. 7).

Let $\tilde{L}'$ be the Piola-Kirchhoff stress tensor (Ref. 4) and $\tilde{D}'$ be the material electric displacement vector in the coordinates system associated to the plate (into the reference configuration). Thus we can write (Ref. 6) :

$$(2)\qquad \begin{cases} \tilde{L}_{lj} = (G_{l\gamma m\varepsilon} \cdot u_{\varepsilon,m} - R_{m\gamma l} \cdot \tilde{\Phi}_m) \cdot \delta_{j\gamma} \\ \tilde{D}_l = N_{lm} \cdot \tilde{\Phi}_m + R_{l\alpha n} \cdot u_{\alpha,n} \end{cases}$$

in which $G_{l\gamma m\varepsilon}$, $R_{l\alpha n}$, and N_{lm} define respectively the effective elastic constants, effective piezoelectric constants and effective permittivities. Since these tensors depend only on the coordinate (x_2'') the stress equation of motion and charge equation reduce to :

$$(3)\qquad \begin{cases} \tilde{L}'_{2j,2} = \rho_0 \cdot \ddot{u}'_\alpha \cdot \delta_{j\alpha} \\ \tilde{D}'_{2,2} = 0 \end{cases}$$

Substituting (2) into (3), we obtain (4) :

$$\begin{cases} G'_{2\alpha 2\beta} \cdot u'_{\beta,22} + R'_{2\alpha 2} \cdot \tilde{\Phi}'_{,22} = \rho_0 \cdot \ddot{u}'_\alpha & (4.1) \\ R'_{2\alpha 2} \cdot u'_{\alpha,22} - N'_{22} \cdot \tilde{\Phi}'_{,22} = 0 & (4.2) \end{cases}$$

The subscript (′) indicates that these quantities are evaluated in the mobile coordinates system. Substituting (4.2) into (4.1) and using (1), we obtain the following Christoffel linear system (5) :

$$\begin{bmatrix} G'_{29} + \frac{(R'_{29})^2}{N'_{22}} & G'_{29} + \frac{R'_{22}R'_{29}}{N'_{22}} & G'_{49} + \frac{R'_{27}R'_{29}}{N'_{22}} \\ G'_{29} + \frac{R'_{22}R'_{29}}{N'_{22}} & G'_{29} + \frac{(R'_{22})^2}{N'_{22}} & G'_{24} + \frac{R'_{22}R'_{27}}{N'_{22}} \\ G'_{49} + \frac{R'_{27}R'_{29}}{N'_{22}} & G'_{24} + \frac{R'_{22}R'_{27}}{N'_{22}} & G'_{44} + \frac{(R'_{27})^2}{N'_{22}} \end{bmatrix} \cdot \begin{bmatrix} u_1^\mu \\ u_2^\mu \\ u_3^\mu \end{bmatrix} = \rho_0 (v^\mu)^2 \begin{bmatrix} u_1^\mu \\ u_2^\mu \\ u_3^\mu \end{bmatrix}$$

Then, using the usual transformation rules for tensors and the rule used to compress indexes into a matrix form (G and R coefficients have lower symmetry than usual coefficients) it can be shown (Ref. 8) that the relations between the (G, R, N) and (G', R', N') constants can be written :

$$(6)\qquad \begin{cases} [G'] = [M] \cdot [G] \cdot [M]^T \\ [R'] = [\Omega] \cdot [R] \cdot [M]^T \\ [N'] = [\Omega] \cdot [N] \cdot [\Omega]^T \end{cases}$$

where M is the Bond's matrix (Ref. 9).

According to system (5) and to relations (6), the velocity field of the wave is only governed by the thermal dependence of the G, R, and N effective constants. Moreover, in the following study we neglect temperature variations of R and N constants and we use Taylor developments to express the G elastic constants in the vicinity of the reference temperature T_0 :

$$(7)\qquad \begin{cases} G_{l\gamma m\varepsilon}(T) = G_{l\gamma m\varepsilon}(T_0)[1 + \sum_1^N T^n \cdot G_{l\gamma m\varepsilon} \cdot (T - T_0)^n] \\ T^n G_{l\gamma m\varepsilon} = \frac{1}{G_{l\gamma m\varepsilon}(T_0) \cdot n!} \cdot \frac{d^n G_{l\gamma m\varepsilon}}{dT^n} \end{cases}$$

where $T^n G_{l\gamma m\varepsilon}$ is the n order effective elastic coefficient, which are known up to the third order (Ref. 6).

In the case of the electroded plate, let respectively h_e and ρ_e be the thickness and the mass of electrodes. The resonance frequency can be approximately deduced from the equation of dispersion. The latter result is obtained by satisfying the classical boundary conditions on the major surfaces (at $\pm h_0$) (Ref. 8) :

$$(8)\quad (2\Pi f)^2 = \frac{n^2\Pi^2}{4h_0^2} \cdot (v^\lambda)^2 \cdot (1 - 8\frac{K^\lambda}{n^2\Pi^2} - 2\frac{\rho_e h_e}{\rho_0 h_0})$$

where (9) $K^\lambda = \frac{R'_{2\alpha 2} \cdot u_\alpha^\lambda}{(\bar{G}^\lambda \cdot N'_{22})^{1/2}}$ is the electromechanical coupling factor, n the partial ranks, $\bar{G}^\lambda$ the eigenvalue of the linear system (5) and the u_α^λ normalizing eigenvectors solution of the same system (5).

4. DOUBLY ROTATED QUARTZ TRAPPED ENERGY RESONATOR WITH RECTANGULAR ELECTRODED MASS-LOADING

There are different BAW quartz resonators classified according to their modes of vibration (flexion, shearing, elongation, torsion, ...) and to the orientation of their geometric sides with respect to the crystallographic axes. For our application, the vibrating structure is a disk-shaped quartz plate crossed by electric field along the thickness. In fact, lateral dimensions of the plate are not endless and vibration not vanishes at the edge of the vibrating structure. Consequently, fixations can absorb the acoustic wave and induce a decrease in the quality factor. To avoid this, we try to confine the vibration in the middle of the contoured plate. In practice, different technologies are used (flat plate with mass-loading, either plano-convex or bi-convex plate) (Refs. 4, 5) to trap energy resonators.

Theoretical analysis of these resonators is not easy to do in the general doubly rotated case (Ref. 5), indeed :

1. three waves can be coupled,
2. eigenvectors are oriented independently of usual axes in the plane of the plate,
3. all mixed derivatives appear in the differential equation of motion.

In order to obtain analytical solutions for velocities of the three waves and then to evaluate frequencies in the case of the plane resonator operating essentially in thickness modes, we adopt the model established by H. F. Tiersten and D. S. Stevens. This model concerns a doubly rotated quartz resonator where energy is trapped by means of rectangular electrodes (Ref. 5) whereas in the present work we replace the rectangular electrodes by a circular or a rectangular electroded mesa in the middle of the plate (figure. 2) to obtain a mass loading effect.

Let us study an electroded mass-loading disk-shaped resonator whose thickness is very thin compared with lateral dimensions (figure 2).

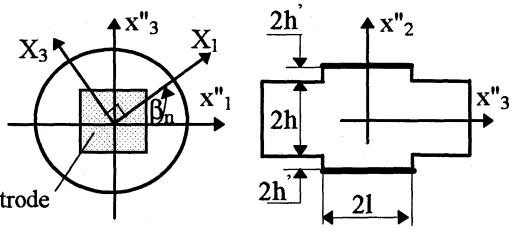


Figure 2. Electroded mass-loading of the trapped energy resonator.

In the general case of doubly rotated cuts the three-dimensional asymptotic model of Stevens and Tiersten (Ref. 5) predicts that the inhomogeneous differential equations for electroded and unelectroded region added to the edge continuity conditions govern the nth odd harmonic family of modes. To remove the mixed derivative term of these differential equations, the authors (Ref. 5) introduce an appropriate planar transformation by using the β_n rotation of the independent coordinates $x"_1$ and $x"_3$ (figure. 2). The β_n rotation depends on the mode family - slow shear C, thickness extensional A, and fast shear B - . The resulting transformed equation with separable variables is thus written as :

$$(10)\quad \left| \begin{array}{l} M_n' \dfrac{\partial^2 \tilde{u}_1^n}{\partial X_1^2} + P_n' \dfrac{\partial^2 \tilde{u}_1^n}{\partial X_3^2} - \dfrac{n^2 \Pi^2 \tilde{c}^{(1)}}{4h^2} \tilde{u}_1^n - \rho \ddot{\tilde{u}}_1^n \\ = \rho \omega^2 (-1)^{(n-1)/2} \dfrac{e_{26}}{c^{(1)} n^2 \Pi^2} 4 v e^{iwt} \end{array} \right.$$

in the new coordinates system (X_1, X_2, X_3).

Here, n is the overtone number, ρ the mass density. M_n' , P_n' are dispersion constants that only depend on the $\hat{c}_{mnrs}, \hat{e}_{mrs}$ and ε_{mn} elastic, piezoelectric and dielectric material constants respectively in the original coordinates system. Otherwise, they also depend for a given overtone number on the orientation of the quartz plate.

The $\tilde{c}^{(1)}$ quantity is the effective stiffness for a thickness vibration which constitutes a solution of the following Christoffel equation :

$$(11)\quad \left| \begin{array}{l} (\bar{\hat{c}}_{2nr2} - \tilde{c}^{(i)} \delta_{nr}) A_r^{(i)} = 0, \quad i = 1,2,3 \\ \bar{\hat{c}}_{2nr2} = \hat{c}_{2nr2} + \hat{e}_{22n} \hat{e}_{22r} / \varepsilon_{22} \end{array} \right.$$

where $A_r^{(i)}$ are the orthogonal eigenvectors.

In the present case of rectangular electroded mass-loading resonators (figure. 2) working in the (10-20) MHz range, the in-plane distribution of the mechanical amplitude $\tilde{u}_1^n$ for a given overtone n and a mode family corresponds to analytical solutions of the equation (10).

5. NUMERICAL RESULTS OF NEW CUTS TO DESIGN MICROMACHINING TEMPERATURE SENSORS

In the expression (8) of frequency, only the v^λ velocity of modes family is governed by the temperature change. So, in the vicinity of reference temperature T_0 the variation of resonance frequency versus temperature can be expressed as a polynomial function :

$$(12)\quad \frac{\Delta f^\lambda}{f_0} = a^\lambda (T - T_0) + b^\lambda (T - T_0)^2 + c^\lambda (T - T_0)^3$$

involving changes in temperature ($T-T_0$) for given λ-mode and overtone n.

In the above equation a^λ, b^λ and c^λ are respectively the first (TCF), second (TCS) and third (TCT) orders temperature coefficients for frequency resonance. In the doubly rotated cuts the three different families are classified with respect to their propagation speed (A-mode , B-mode and C-mode). Thus, at all there are nine temperature coefficients of frequency for a given n partial (which we call (aTCF, aTCS, aTCT) for A-mode, (bTCF, bTCS, bTCT) for B-mode and (cTCF, cTCS, cTCT) for C-mode).

As frequency, the electromechanical coupling is also governed by the thermal variations because of the temperature dependence of the u_α^λ eigenvectors and of the $\bar{G}^\lambda$ eigenvalues in accord with relation (9) (variations with temperature of effective constants $R'_{2\alpha 2}$ and N'_{22} are neglected (section. 3)). Since λ characterizes the mode family of propagation, there are three electromechanical coupling factors related to A-mode, B-mode and C-mode which are called K^a, K^b and K^c respectively.

We elaborate a computer program in C-language in order to determine the order temperature coefficients of frequency and the electromechanical coupling factor for a given (φ, θ) crystalline orientation of the plate. In the flow chart we firstly evaluate the G effective constants at temperature T from the G constants at the fixed reference temperature (25C) by relation (7). Secondly we return the material constants G(T), R(25C) and N(25C) in the mobile axes of the plate by relation (6). In a further stage, we solve the Christoffel system (5) to obtain generally three velocities for propagating wave which are classified by mode family and then to determine frequencies and electromechanical coupling. We repeat this sequence at different temperatures in the (0C-100C) range. Finally we plot either velocities or frequencies versus temperature as cubic curves. The λ-TCF, λ-TCS and λ-TCT values can be determined by means of a

least square method. Taking into account the symmetry axes of the 32 class contour plots of the nine temperature coefficients and of the three coupling factors versus the angles of cut are then drawn in (0°, 60°) φ-range and (-90°, 90°) θ-range.

A second computer program allows us to evaluate the nine values for the M'_n , P'_n dispersion constants and for the β_n orientation of the mesa for a given n overtone and a given cut (section. 4).

Among the nine temperature coefficients of frequency, only three of them are significant, namely the a-TCF, b-TCF and c-TCF.

For use as a temperature sensor, a resonator with a first order TCF larger than TCS (second order) and TCT (third order) is required. It is also desirable that the K^{λ} electromechanical coupling must be as higher as possible. In order to obtain trapped energy resonator, values for the M'_n and for the P'_n dispersion constants must be positive (mass-loading by mesa) or negative (mass-loading by hole) (Ref. 5) and preferably between 20 GPa and 200 GPa.

Let us be interested now to the A-mode family, the contour plots of TCF versus the angles of cut (φ, θ) shown in figure 3a indicate that the a-TCF coefficient is greater than 60ppm/°C for the angle θ ranging from 20° to 90° and the angle φ in the range [30°, 60°]. Turning our attention to the K^a electromechanical coupling factor in figure 3b we observe that K^a remains greater than 40 when angles φ and θ belong respectively to the [30°, 45°] and [0°, 40°] ranges. Thus, it seems interesting to choose the orientation (φ,θ) close to (37°,40°). Looking now to the values for M'_n and P'_n listed in table. 1 (for 1, 3 and 5 overtone numbers) this choice seems convenient. Finally, frequency-temperature curves in the (0C, 100C) range are given for this cut and for the overtone n=3 in figure 3c.

Additional results for the c-TCF and K^c are given in table 1 and figures 4a and 4b. For the particular doubly rotated cut close to φ=40° and θ=-50° the frequency temperature curve drawn in figure 4c exibits a general behavior which does not depart very much from the behavior depicted for the A-mode and for the cut close to φ=37° and θ=40°. Nevertheless, the temperature sensitivity is twice greater for the last cut.

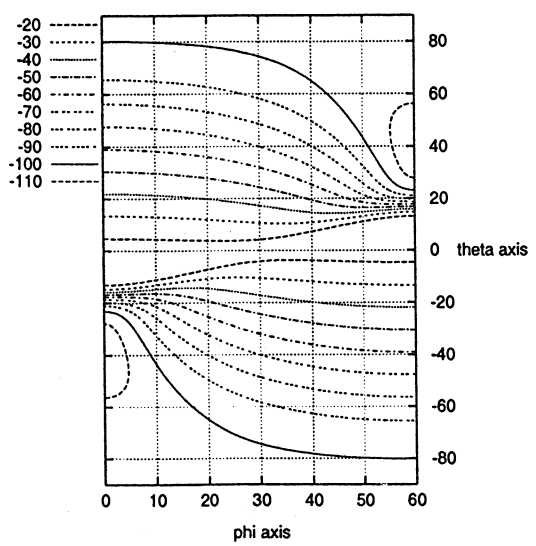


Figure 3a. Contour plots of the first frequency coefficient for the A-mode family.

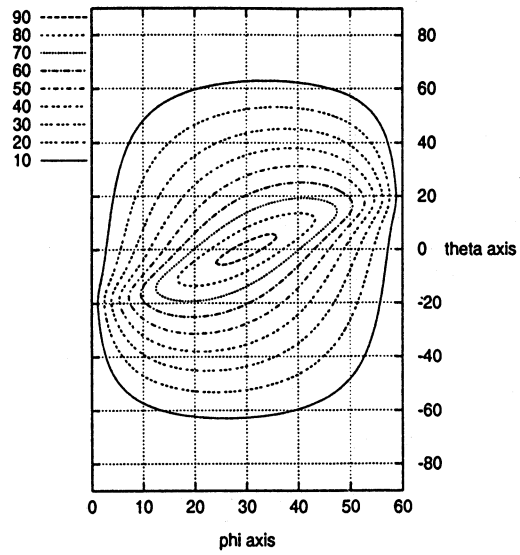


Figure 3b. Contour plots of the electromechanical coupling factor for the A-mode family.

Cut	Mode,overtone	M'_n	P'_n	β_n
φ=37° θ=40°	A,1	-26.2	-34.2	14.2°
	A,3	140,4	43.1	34.0°
	A,5	68.8	54.1	39.3°
φ=40° θ=-50°	C,1	42.1	136.4	42.8°
	C,3	140,4	94.2	38.9°
	C,5	68.8	70.5	-26.1°

Table 1. Dispersion constants and orientation of the mesa.

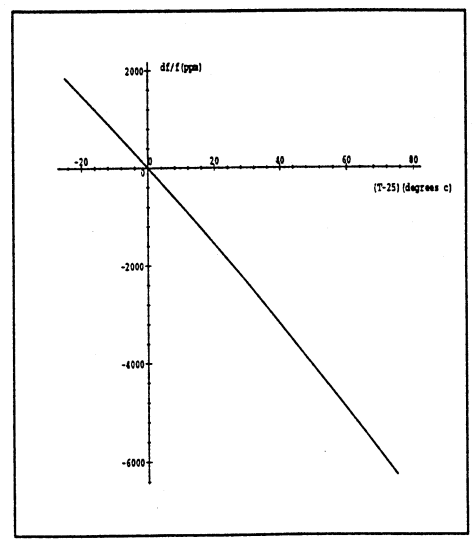


Figure 3c. The frequency-temperature curve for the overtone n=3. The case of a doubly rotated cut (φ=37° and θ=40°) with $2h_0$=500μm.

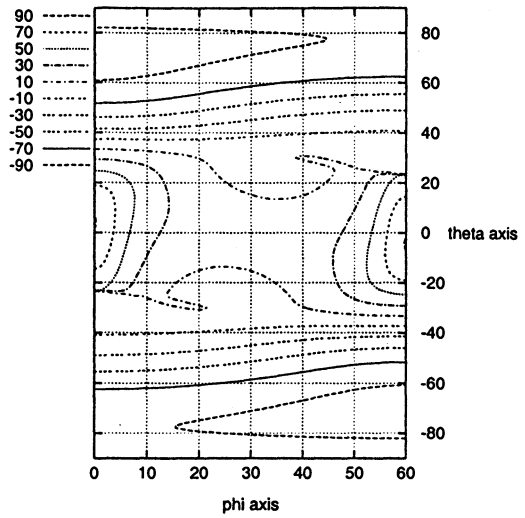


Figure 4a. Contour plots of the first frequency coefficient for the C-mode family.

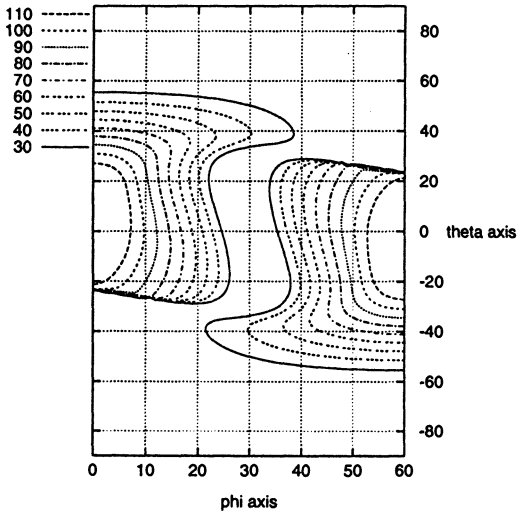


Figure 4b. Contour plots of the electromechanical coupling factor for the A-mode family.

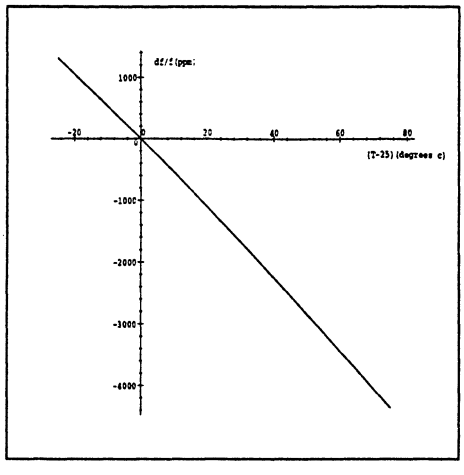


Figure 4c. The frequency-temperature curve for the overtone n=3. The case of a doubly rotated cut (φ=40° and θ=-50°) with $2h_0$=500μm.

6. CONCLUSION

In this paper we have chosen to present theoretical results for a given cut and for a given thickness mode of vibration. But care must be taken that the present analysis has revealed that several new doubly rotated cuts operating in different thickness modes can constitute the sensing element of temperature sensors working in the (0C, 100C) temperature range. The further step of this study consists in a systematic investigation of possibilities of fabricating the resonant structure by a wet micromachining process.
The anisotropy of the chemical etching which induces undesirable effects such as for quartz crystal, sharp underetch or formation of limiting facets slightly disoriented with respect to the reference cut will certainly lead to a restriction of the number of possible cuts (Ref. 10). This experimental work will be reported in a future paper.

Acknowledgements

I wish to thank B. Dulmet, R. Bourquin, C. R. Tellier and T. Leblois of Laboratoire de Chronométrie, Electronique et Piézoelectricité (LCEP. Besançon) for their helpful advice and discussions during this work.

7. REFERENCES

(1) T. Ueda, F. Kohsaka, T. Lino, D. Yamazaki, "Temperature Sensor Using Quartz Tuning Fork Resonator", 40th Annual Frequency Control Symposium, 1986, pp.224-228.
(2) B. Dulmet, R. Bourquin, "Méthodes de détermination des courbes fréquence-température d'un résonateur à quartz vibrant en mode d'épaisseur", Revue de Physique Appliquée, 18(1983), pp619-624.
(3) R. Bourquin, B. Dulmet, "Méthode rapide de calcul des courbes fréquence-température des résonateurs à quartz - Application au barreau vibrant en flexion", C. I. C. 1984, Session II, Communication 2.
(4) C. J. Wilson, "Vibration modes of AT-cut convex quartz resonators", J. Phys. D7, 2449, 1974.
(5) D. S. Stevens and H. F. Tiersten, "An analysis of doubly rotated quartz resonators utilising essentially thickness modes with transverse variation, J. Acoust. Soc. Am, Vol 79 (1986), pp. 1811-1826.
(6) B. Dulmet, R. Bourquin, "Application of Lagrangian Effective Material Constants to the Study of the Thermal Behavior of SAW Propagation in Piezoelectric Crystals", Proc. IEEE Ultrasonics Symposium, pp 331- 334 (1994).
(7) E, Dieulesaint et D, Royer, "Ondes élastiques dans les solides", Masson et Cie, Paris, (1974).
(8) B.Dulmet, "Contribution à l'étude des vibrations à énergie piegée des résonateurs piézoélectriques à ondes de volume" Thesis, No. 243, Besançon, France, 1991.
(9) W. Bond, "The Mathematics of Properties of Crystals", BSTJ, 22, pp 1-72, (1943).
(10) T. Messaoudi, T, Leblois and C. R. Tellier, "Chemical Etching of Y-Rotated Quartz Plates Orientation Effects on 2D and 3D Shapes", Proc. of the 10th EUROSENSORS, Leuven, Belgium, 1996.

CORRECTION OF QUARTZ RESONATORS ANGLE OF CUT. GEOMETRIC MODEL AND EQUATIONS FOR SC CUT.

Georgios PENTOVELIS, Caroline BOUDY
C.Q.E. TEKELEC TEMEX
rue Robert Keller
10150 Pont Sainte Marie, FRANCE

1. ABSTRACT.

The quartz resonator angle of cut can be corrected by mechanical, optical or some other similar process. When the AT cut is corrected, the Y' axis is moved on the X level (rotation around the X axis). When a double rotation cut is used the Y'' axis is moved on an other level. In this paper we present the main equations to determine the Y'' axis movement in the SC cut case. These equations can be used to correct simultaneously the θ and ϕ angles of SC cut quartz blanks by an automatic process.

2. INTRODUCTION.

The cut of a quartz blank determines the frequency-temperature behaviour of the resonator. Angles of cut are choosen in accordance with the mounting and the application of the resonator.

The tight angle tolerance for an SC cut resonator are such that it is not possible to saw to the required precision, we therefore need to angle correct. The AT cut angle tolerances can be achieved directly from the saw.

Where special tight angle orientation is required for the AT-cut, correction is effected by a rotation of the Y' face around the X axis for δθ. In practice this rotation is made by a manual process.

The same mechanical sawing process is used in both the AT and SC cut. The SC cut correction is operated for both θ and ϕ modification, this compounds the problem making correction doublery difficult over the AT. It is therefore very important to improve this process by automation.

3. GEOMETRIC MODEL.

In the case of an AT cut quartz blank when θ angle changes the vector of the Y' face is moved on the YZ level for δθ . The surface Y' rotates around the X axis of the crystal.

3.1. θ change of SC cut blanks (figure 1).

In the case of a SC cut, θ angle change is similar to that of AT cut crystal. The vector of the Y'' face is moved on the Y''Z level for δθ. The face Y'' rotates around the X'' axis of crystal.

Angle ϕ of the SC cut is not changed during θ change operation.

3.2. ϕ change of SC cut blank (figure 2).

When ϕ angle of SC cut quartz blank changes the surface vector Y'' makes a rotation around the Z axis for δϕ. The face Y'' rotates around the Z axis of the crystal. θ angle is not changed during ϕ change operation.

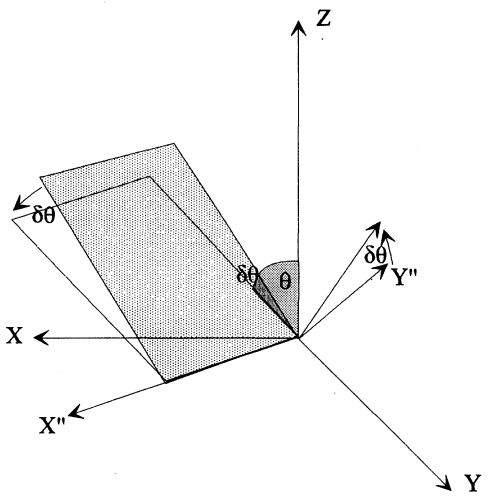


figure 1.

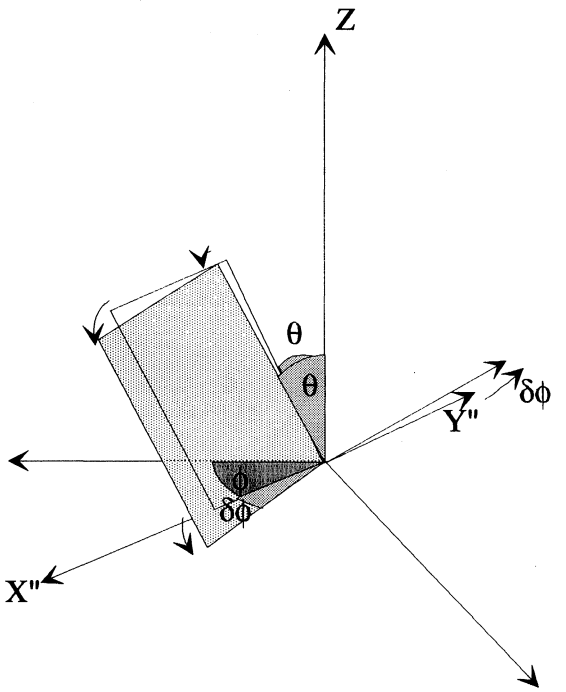


figure 2.

On the figure 3 we show a SC cut quartz blank with its axis X'', Y'', Z'' and Z. After the surface orientation modification the new surface vector is called Y'''.

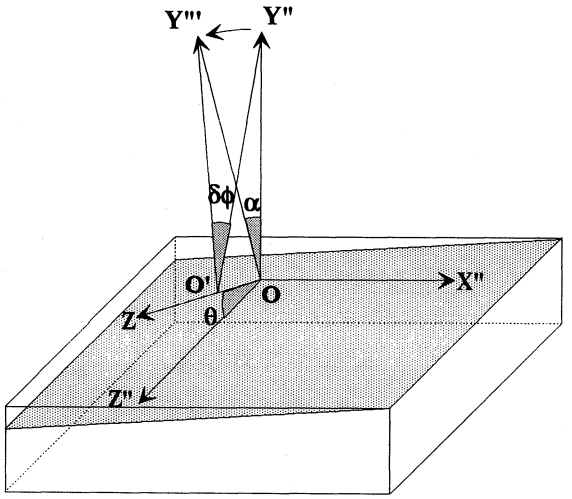


figure 3.

Usually the angle ϕ change is of the order of some minutes of angle. So we can admit that Y'' axis rotation around Z axis for δϕ angle is equal to a rotation around Z'' axis for α.

$$\widehat{O'Y'''O} = \widehat{O'Y''O} = \theta$$

$$\sin\left(\frac{\alpha}{2}\right) \approx \cos(\theta) \cdot \sin\left(\frac{\delta\phi}{2}\right)$$

$$\left.\begin{aligned} \alpha &\approx \cos(\theta) \cdot \delta\phi \\ \cos(\theta) &\approx 0.8 \end{aligned}\right\} \Rightarrow$$

$$\alpha \approx 0.8 \cdot \delta\phi$$

When then ϕ angle of SC cut crystal blank changes, the vector of the Y'' face is moved on the X''Y'' level for $\cos(\theta).\delta\phi$ angle. Surface Y'' change like a rotation around the Z'' axis of crystal.

3.3. θ et ϕ simultaneous change of SC cut blank.

θ et ϕ angle changes of SC cut are carried out in two movements: firstly a rotation around the X'' axis for δθ angle and secondly a rotation around the Z'' axis for δϕ angle. In figure 4 we see that the surface vector is moved from the BA direction to BC direction (θ change) and finaly to BD direction (ϕ change).

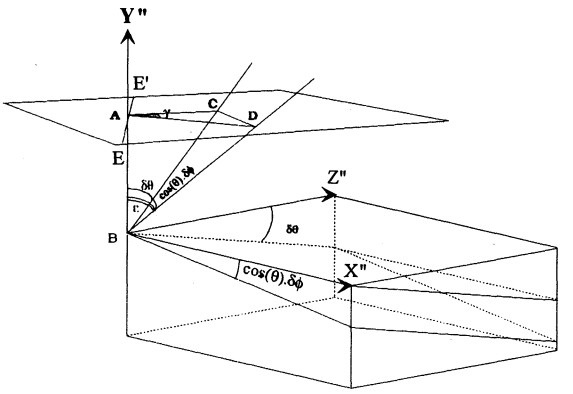


figure 4.

$$\tan(\gamma) = \frac{\tan(\cos(\theta).\delta\phi)}{\sin(\delta\theta)}$$

$$\tan(\gamma) \approx \frac{\cos(\theta).\delta\phi}{\delta\theta} \qquad (1)$$

$$\cos(\varepsilon) = \cos(\cos(\theta).\delta\phi).\cos(\delta\theta) \qquad (2)$$

The result of the double rotation is a simple rotation around the EE' axis for ε.

The EE' rotation axis is on the old surface of the blank and γ angle from the X'' axis.

4. APPLICATION.

Equations (1) and (2) are necessary to correct θ and ϕ angle of SC cut blank. From these we work out the ε and γ angles of blank. Each blank is placed on a rectification table according to γ angle. Surface plan is sloped down for ε.

On the figures 5a, 5b, 5c we show three differents cases of correction of the SC cut angles: only θ is corrected (figure 5a), only ϕ is corrected (figure 5b), and finaly θ and ϕ are corrected (figure 5c). In the three cases ε angle can be the same and only γ angle is different.

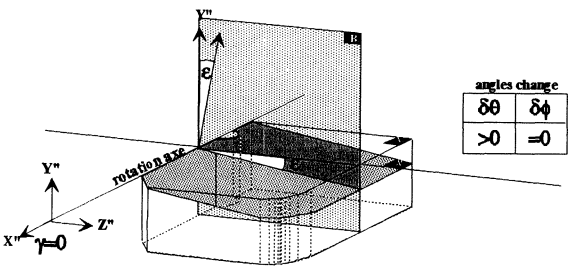


figure 5a.

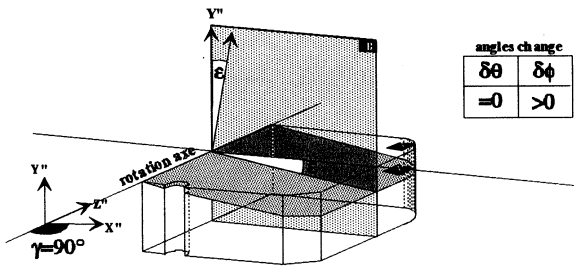


figure 5b.

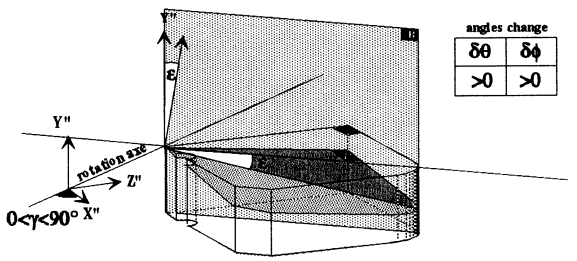


figure 5c.

Example:

CASE	δθ (sec)	δφ (sec)	ε (sec)	γ (deg)
a	60	0	60	0
b	0	73	60	90
c	46	47	60	40

The cut of the three blanks can be corrected simultaneously because only their position (γ angle) on the rectification table is different.

5. CONCLUSION

The double rectification of the SC cut blank's angles can be reduced to a simple operation. The number of the correction operations are reduced and more blanks can be batched together. So automatisation of this process is possible.

In our process SC cut blanks are sorted according to the θ and φ angle and additionally they are batched together according to the ε angle.

6. ANNOTATIONS.

φ First rotation angle of SC cut
θ Second rotation angle of SC cut
δφ φ required correction
δθ θ required correction
α Surface level slope for only φ correction
ε Surface level slope for θ and φ correction
γ Angle between rotation axis and X'' axis for θ and φ correction
X,Y,Z Crystallographic axis of quartz crystal
X'' Modified crystallographic axis of SC cut quartz blank.
Y'' Modified crystallographic axis and surface vector of SC cut quartz blank.

7. KEY WORDS.

Quartz, resonators, SC cut, angle correction, θ angle, φ angle

MEASUREMENT METHOD FOR HIGH FREQUENCY QUARTZ CRYSTALS

Eberhard Seydel

R&D Quartz Crystals
KVG QUARTZ CRYSTAL TECHNOLOGY GmbH
A VECTRON INTERNATIONAL COMPANY
D-74922 Neckarbischofsheim , P.O.Box 61

ABSTRACT

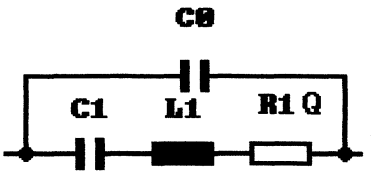
For high frequency quartz crystals the resonance of the holder coincides with the quartz resonance frequency. This requires new measurement steps different to the IEC-444-5.
The holder inductance is added to the equivalent circuit of the quartz crystal.
The new measurement method first determines the value of the inductance and the static capacitance of the complete system. In a second step it calculates the equivalent electrical data only for the quartz crystal.

1. INTRODUCTION

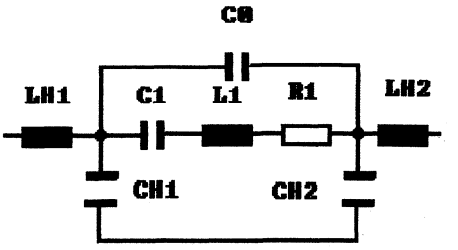
The measurement of the equivalent data of quartz crystals in the frequency range over 250 MHz is very difficult. The electrical properties of the mounting system influence the measured values. The resonance of the holder-system with the static capacitance is normally in the range of 400 to 1000 MHz. It is useful to find a complex equivalent circuit-diagram for the full system. This allows to calculate the real electrical data for the quartz crystal and for the holder-system. This requires new measurement and transformation steps different from these described in the IEC-444-5.

2. EQUIVALENT CIRCUIT

The reduced equivalent circuit for the quartz crystal (fig.1) was extended by the holder inductance (fig.2) . The properties of this equivalent circuit are near the real properties of a high frequency quartz crystal.



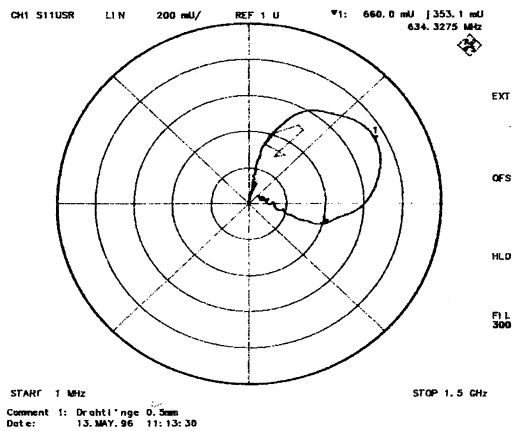
Equivalent circuit blank fig.1



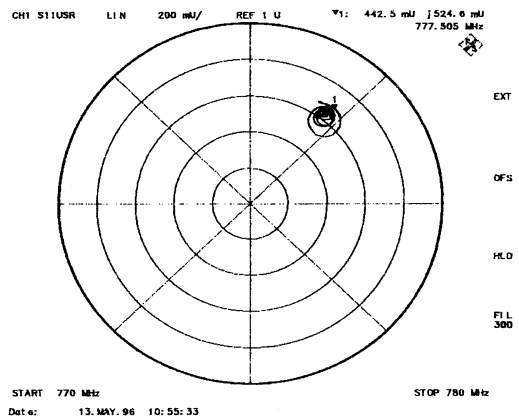
Equivalent circuit blank with enclosure fig.2

3. MEASUREMENT

The measurements were made with a network-analyser ZVR from Rohde & Schwarz. The figures 3 and 4 show the behaviour of a 777 MHz quartz crystal at 5th overtone in a standard PI-network. A sweep over the frequency range from 1 MHz to 1.5 GHz is shown in figure 3 and the sweep over the quartz resonance at 777 MHz in figure 4.



Frequency sweep 1..1500 MHz **fig. 3**

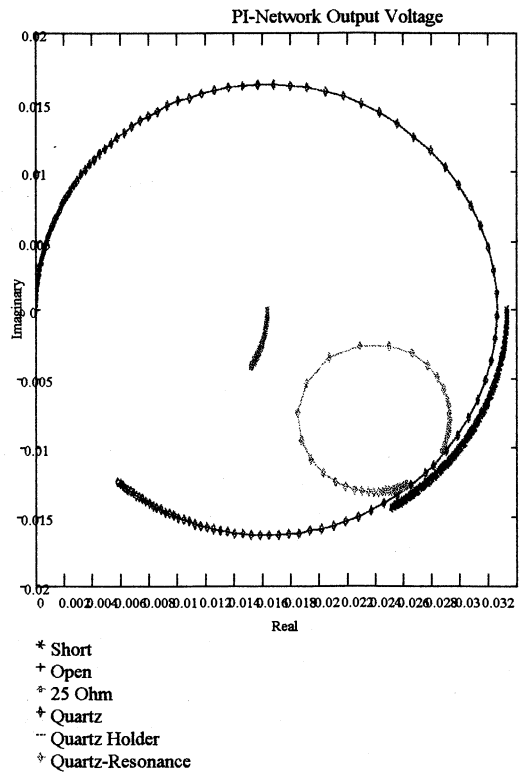


Frequency sweep 770..780 MHz **fig. 4**

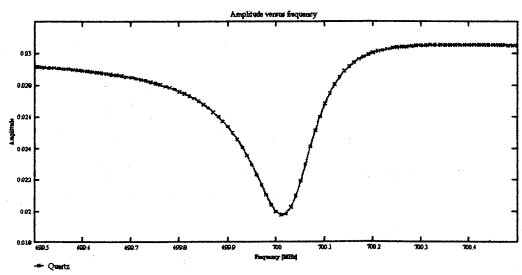
The resonance circle of the quartz crystal rotates around a point, which is described by the values of static capacitance, holder inductance and resonance frequency. The simple linear transformation according to IEC-444-5 with C0 measured near the resonance frequency gives unacceptable values.

4. TRANSFORMATION

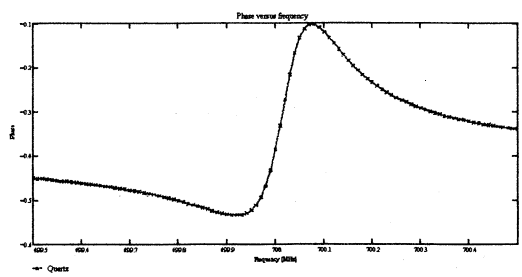
For a measurement of high frequency quartz crystals it is necessary to determine the static capacitance C0 and the holder inductance LH. The measured values before the transformation are shown in figure 5 and after transformation in figure 8. Figure 6 and 7 show the behaviour of amplitude and phase for the high frequency quartz crystals. The resonance frequency is near the point of lowest magnitude.
The first transformation of all measured values is given by the transmission behaviour of the PI-network and these routines for calibration and measurements of the PI-network are included in the IEC 444-5.



Real and imaginary **fig. 5**
Before transformation



Amplitude **fig. 6**
Before transformation

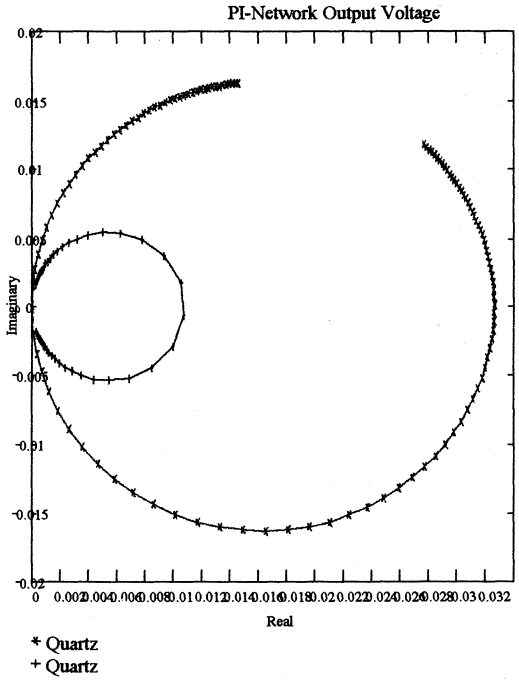


Phase **fig. 7**
Before transformation

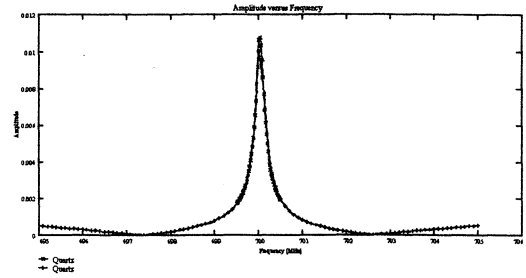
The measurement of the static capacitance is real made at 12.5 MHz , out of the range of the holder resonance.
The resonance of C0 and LH gives a circle in the real-imaginary-diagram. The coordinates of the central point and of this circle [c,i*d] and its radius r can be determined by measuring at least three points [x,i*y] of this curve (1).

$$r^2 = (x-c)^2+(y-d)^2 \qquad (1)$$

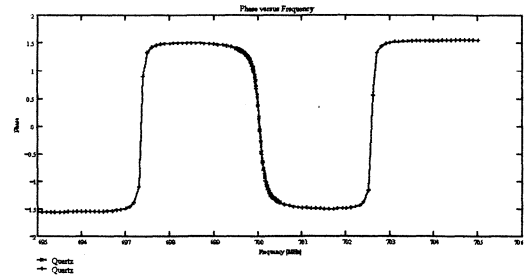
The error of these measurements can be reduced by the method of minimising the mean square root deviation.
The value of the resistance of electrode film and holder wires is near zero, therefore it can be neglected.
With these measured values the value of LH can be calculated.
The coordinates of the central point [c,i*d] of the C0-LH-circle determine the rotation centre. The point of quartz resonance frequency on this curve determines the rotation angle for the transformation.



Real and imaginary **fig. 8**
After transformation



Amplitude **fig. 9**
After transformation



Phase **fig. 10**
After transformation

Figure 9 and 10 show, that the behaviour in magnitude and phase of a high frequency quartz crystal after transformation is the same as of a low frequency crystal.
During a measurement process each measured point must be transformed. After that the normal routines of the IEC-444-5 can be used.

5. CONCLUSIONS

High frequency quartz crystals can be measured with more accuracy by using this transformation, and also quartz crystals with high C0-values. The values for the equivalent circuit R1, C1, C0, and additionally the value of LH , will be provided for the customer.
C0 includes in this case the value of CH.

LITERATURE

/1/ IEC-444-5

QUARTZ-CRYSTAL FREQUENCY STANDARD WITH A LONG-TERM AGING COMPENSATION

Yu.S.Shmaly, A.Ph.Kurochka* , A.V.Marienko**

"Sichron" Center, 4 Skripnika Str., Kharkiv, 310057, Ukraine,
tel/fax: 380 572 45 09 55; e-mail : yush@sichr.kharkov.ua
** St.,Co "IRVA", 12/14 Radisheva Str., Kyiv, Ukraine

This paper discusses an application of the modulational method of a quartz-crystal frequency control [1,2] for decrease of long-term aging instabilities in quartz-crystal standards. The discription of these ones, in which the compensation of long-term frequency instabilities is realized with the use of the method, is given. The method is based on a use of the properties of an unharmonic resonances of quartz-crystal units.

The results of the experimental studies and correlation-regression analysis of thickness-shear eigenfrequency aging special feature of AT-cut vacuum quartz-crystal units of PK-187 (Phonon) type, 5 MHz are given. The probability properties of a long-term frequency behaviour of AT-cut quartz-crystals unharmonic resonanses modes h_{511}, h_{513} had been studied during two years. It has been found the long-term frequency behaviour of correlation R_{ij} and regression K_{ij} matrixes of f_{511} , f_{513} frequencies for ten quartz units.

With the use of the method and with taking into account of R_{ij} and K_{ij} the mathematical model of the system of a long-term frequency instabilities compensation has been studied. The results are given of investigations of three units of the quartz frequency standards of Ч1-88 type, which are realized with the use of the method. Also it is given the results of studies of the mathematical model of these standards with the test results and comparisons. The principle standards characteristics without and with the use of the compensation circuits and the experimental functions of frequency instabilities in various operating regimes of the standards are given.

It's shown that without the use of the method the frequencies had the aging rate of $(1...2)\cdot10^{-10}$/day and drift of $(3...5)\cdot10^{-8}$/year. With the use of the compensation regime the frequencies had the aging rate of $(1...5)\cdot10^{-11}$/month and drift less than $(1...2)\cdot10^{-10}$/year. With this the day's Allan variance with the temperature of $(27\pm0.5)^{\circ}C$ was less than $2\cdot10^{-10}$ in regime of static compensation and will be less than $5\cdot10^{-11}$ (the studies are continued) in regime of dynamic one. The short-term Allan variance of the standards frequencies is $(1...3)\cdot10^{-12}$/sec.

The calibration of the standards carry out with the use of the SICHRON LELCAL software for MS DOS.

References :
[1]. Yu.S.Shmaly. Proc.of 49th IEEE IFCS, 1995, pp.579-589.
[2]. Yu.S.Shmaly,V.N.Romanko,A.V.Marienko. Proc.of 7th Int.Congress METROLOGIE-95, 1995, pp.279-283.

EFFECTS CAUSED BY NONLINEAR VIBRATIONS OF RESONATORS

A.G. Smagin
Cristal Co., 105425, Moscow, 9th Park Street, 47-1-29, Russia

Nonlinear effects have a great influence on the behaviour of the main radiophisical parameters and characteristics of piezoelectric systems. With great amplitudes of mechanical vibrations the resonance curve $A(\omega)$ is deformed by maintaining maximum displacing at $k > 0$ to the region of higher frequency (yxl/+35^0,f=1 $ MHz) and at $k < 0$ to the region of lower frequency (yxl/-48^0, f=1 MHz). Further amplitude of forcing strength increase causes an abrupt nonlinearity of resonance characteristic: at a certain frequency ω_c the amplitude decreases in a jump up to the minimum value at the curve $A(\omega)$. With cyclic frequency variations the hysteresis phenomenon originates with alternative transitions. Based on the nonlinear equation

$$I^2 [Q^{-2} + (y - kI^2)^2] = \varphi^2 / R^2 Q^2 \qquad (1)$$

we define frequencies limiting instability region, from the condition $dI/dy = \infty$

$$y_1 = y_2 = y = 2kI^2 = \begin{cases} 2Q^{-1}, & k > 0, \\ -2Q^{-1}, & k < 0 \end{cases} \qquad (2)$$

$$k^2 I^4 - Q^{-2} = 0, \qquad I^2 = (kQ)^{-1}$$

The critical value of driving voltage φ_c, at which the break of vibration amplitude occurs, we find by substitution of the expressions I^2 and y in eq. 1: $\varphi_c = 2R^2/kQ = 0{,}85$ V, which agrees with the experimental value of 0,82 V. The critical value of frequency change corresponding to vibration break, also agrees with the theory and is $y = 1{,}27 \cdot 10^{-6}$. Investigations of the above-mentioned resonators under nonlinear conditions made by X-ray topography method had shown that at amplitude of vibration break in all the difractograms field a special structure appears - a great number of localized mosaics. Appearance of a mosaic structure (and within hysteresis region of submosaic one) can be explained by abrupt amplitude change of higher harmonics of mechanical vibrations, which is characteristic for anharmonism region only.

MINIATURIZATION OF MONOLITHIC PIEZOELECTRIC FILTERS OPERATING OVER THE FREQUENCY RANGE FROM 3 MHz TO 200 MHz

I.A. Tislenko, I.M. Larionov, V.A. Mostiaev
Scientific Research Institute Phonon
44, Krasnobogatyrskaya st., Moscow, 105023, Russia

New design-technological approach to mass production of monolithic piezoelectric filters with increased vibration resistance and minimum volume-weight parameters is described. These filters are enclosed in flat metal packages, in which strip-type piezoelectric elements are mounted by means of a metal frame. They have low price and high parameter repeatability. Their frequency characteristics are listed and design features are given. New ratios of relative width of crystal strip-type elements are established and for the first time the ratios for lithium tantalate and langasite crystal elements are derived. Development and realization results of wide-band monolithic crystal filters (MCF) without use of added inductors with the pass band up to 0.5% are presented. Frequency characteristics and design of fabricated 4th-order MCFs operating within the frequency range from 30 MHz to 200 MHz, realized in a single package with the size of 12 x 12 x 5 mm with the use of two strip-type piezoelectric elements having mesa of a circular or rectangular form, ensuring guaranteed attenuation of 80 dB and the absence of unwanted pass bands are also given. Special interest represents the design solution of the 8th-order MCF in the same package with the location of two electrode pairs at a single mesa of each piezoelectric element. Frequency characteristics of filters designed in such a way, are also shown.

THE QUARTZ-CRYSTAL OSCILLATOR WITH OUTPUT SIGNAL'S CONTROLLED PHASE

A.S.Vasilenko, Yu..S.Shmaly

Kharkiv Military University &
"Sichron" Center, 4 Skripnika Str.
Kharkov, 310057, UKRAINE
tel/fax: 380572 450 755
e-mail: yush @ sichr. Kharkov.ua

The quartz-crystal oscillator circuit is performed on the base of a quartz-crystal resonator with additional electrodes (XRAE), the use of which allows to combine the functions of generation, frequency stabilisation and control over the output signal's phase within one device, is discussed. The dependence of both a slope and a stability of the characteristics from parameters of the quartz resonator which is essential for control of the output signal's phase, was found.

The frequency of stable oscillations ω_0 is located within the limits of a quartz-crystal resonator resonance gap. Due to the interactions between the perpendicular exciting fields (between the main electrodes) and the parallel ones (between the main and additional electrodes) in XRAE the phase ratio of the oscillations on the load resistor appears to be sensitive to the complete resistance changes of the external T-shape RC circuit. This property is used for mechanical or electronic control of the output signal's phase by means of change of R1 resistance; this function can be performed by any active element, for example, field transistor.

The investigations showed that the quartz-crystal oscillator had high slope level of control characteristics as well as high stability of output oscillations amplitude in a regime of phase control, in particular, the slope of static modulation characteristic S_{fm} reached 20 000 degree/v within the phase range $\Delta\varphi = \pm 85^0$.

ELECTRIC SELF-SUSTAINED OSCILLATIONS IN LAYERED STRUCTURES

Vladimir A.Vyun

Institute of Semiconductor Physics
Acad. Lavrentjev Avenue 13, Novosibirsk 630090, RUSSIA
(E-mail: vyun@ispht.nsk.su; FAX: 007-3832-357578)

The paper presents studying results on oscillations of acoustoelectric, photoelectric and thermoelectric voltage in layered structures with piezoelectric, semiconductor or pyroelectric layers some aspects of them were presented in the first papers [1, 2]. This phenomenon exists due to besides nonlinear properties of the layered structures and a feedback there are relaxation properties of electronic processes on semiconductor surface or thermal processes in the structure. Up to date this effect has not been studied although it can be easy realized practically in any layered structures. It appears to have considerable promise.

Experimental results with a quasi-harmonic mode of the self-sustained oscillations are at the center of attention. Presented theory taking into consideration nonlinearities, feedback, and relaxation processes on surface state charge in semiconductor or thermal relaxation explains experiments. We demonstrate that the results obtained give principal methods to improve sensitivity and precision because frequency of the self-sustained oscillation ω, usually in up to 1 MHz frequency band, is expressed in terms of the charge/discharge relaxation time τ of the structure and the value τ to be determined directly.

This connection between ω and τ can successfully be used in physical and chemical sensors with frequency output and may be a basis of nondestructive methods of investigation of semiconductor and environment parameters. It is the great interest because of in the precise sensors and investigation methods we use self-organization processes in which the frequency of electrical signal can be easy measured with high precision and thus it gives high sensitivity. The sensitivity to temperature, illumination, humidity, gas impurities, and other environment parameters is demonstrated.

References

[1]. V.A.Vyun, "Self-sustained oscillation of acoustoelectric effects," *Proc. 1992 IEEE Ultrasonics Symp.*, Tucson, USA, pp.987-989.

[2]. V.A.Vyun, "Self-sustained acoustoelectric and photoelectric oscillations in systems with relaxation," *Proc. 1993 Intern. Frequency Control Symp.*, Salt Lake City, USA, pp.509-513.

THE MASS-LOADING EFFECTS ON THE PARAMETERS OF THE PLAN-PARALLEL AND CONTOURED SC-CUT QUARTZ RESONATORS

I. Mateescu, J. Zelenka *, K. Weiss **, G.Pop

National Institute of Materials Physics, Bucharest-Magurele, Romania
* Technical University of Liberec, Liberec, Czech Republic
** Tele & Radio Research Institute, Warszawa, Poland

ABSTRACT

New experimental results related to the harmonic dependence of the characteristics of the plan-parallel AT-cut, SC-cut and plan-convex SC-cut resonators on various electrode parameters are presented. On the basis of the analysis of the trapped-energy resonators vibrating in the coupled thickness-shear and thickness-twist the dependencies of motional inductance on electrode thickness and diameter were computed. The experimental results prove the strong influence of the effects associated to the deposition of thin electrodes on the behaviour of plan-parallel AT and SC-cut resonators. The analysis of experimental data for plan-convex SC-cut resonators shows that the influence of the contouring on the trapping is higher than the influence of the electrode. The dependence of the frequency spectrum on electrode parameters is presented.

1. INTRODUCTION

The exact transmission-line analogs for a single thickness mode in an infinite plan-parallel plate piezoelectric resonator (Ref. 1), is useful in analysing the structures where multiple harmonics of the resonator are excited.
In order to account for the non-uniform distribution of motion found in finite plane electroded plates, a correction of the mass-loading and coupling coefficient was realised for AT-cut resonators (Ref. 2), using the Tiersten's analysis (Ref. 3) of trapped-energy resonators vibrating in coupled thickness-shear and thickness-twist modes in the vicinity of the fundamental and odd overtone thickness-shear frequency.
Based on a previous study (Ref. 4) of the trapped-energy SC-cut resonators, Zelenka has performed in (Ref. 5) a theoretical computation of effective mass-loading, effective coupling coefficient and motional capacitance. A comparison was made between theoretical and experimental results. The calculated values give a good description of the range of the experimental ones in all cases. A good agreement was obtained especially on fundamental frequency and third harmonic, for thin electrodes and for bigger electrode diameter. The different behaviour of the parameters of SC-cut resonators with thick and small electrodes working on high harmonics might be due to the effects associated to the electrode deposition.
The experimental results of the comparative study realised in (Ref. 6) show that stress at interface electrode-substrate and inertial effects have a strong influence on the behaviour of the harmonic dependence on the characteristics of the AT-cut and SC-cut resonators.
Thus, the non uniform distribution of motion could be ascribed to the coupling of thickness-shear with thickness-twist modes, and to the effects at interface electrode-piezoelectric substrate too.
Based on previous results (Ref. 5, 6) in this paper we have measured and computed the motional inductance values on fundamental frequency and overtones of plan-parallel AT-cut and SC-cut for various electrode parameters, in order to determine the optimum electrode thickness and diameter corresponding to minimum variation of inductance with harmonic order.
The previous analysis of Tiersten, Stevens and Smythe (Ref. 7, 8, 9) for plan-parallel AT and SC-cut quartz resonators was extended to contoured ones, where the plate thickness was introduced as a slowly varying function. The resonator, driven into coupled thickness-shear and thickness-twist vibrations by the application of a driving voltage across the electrodes, was analysed for fundamental and anharmonic overtones of the fundamental and each harmonic overtone of thickness mode. In the contoured resonators the influence of the electrode on the trapping is less than that of the contouring and causes the mode to be highly trapped in the vicinity of the centre of the plate.
Our present study was extended to the plan-convex SC-cut resonators of 10 MHz working on third harmonic. The harmonic dependence of the motional inductance and the content of anharmonic resonances for various electrode thickness and diameter are performed. The analysis of the experimental results shows that the evolution of the motional inductance with harmonic order can be explained by a higher influence of the contouring on the trapping than the influence of the electrodes. The frequency spectrum of contoured SC-cut resonators is strong influenced by the electrode parameters.

2. INDUCTANCE BEHAVIOUR OF PLAN-PARALLEL AT AND SC-CUT RESONATORS

The behaviour of inductance, the most significant parameter of electrical equivalent circuit of resonators vibrating in thickness shear mode, is analysed. The motional inductance of AT and SC-cut resonators changes with harmonic order despite the lack of the usually analytical relationship due to the non-uniform distribution of motion.
The plan-parallel polished AT-cut (θ=35°17') and SC-cut (YXwlt 21.93 / 33.93 / 76.4) quartz plates with 14 mm diameter and ≈5 MHz fundamental frequency were used for experiments. The CrAu electrodes with 100, 300 and 500 nm thickness and 4.6, 6 and 7 mm diameters were deposited by thermal evaporation in vacuum. The resonance frequencies of

the fundamental, 3rd, 5th and 7th harmonics for every mass-loading and electrode diameter were measured using an HP 4194A Impedance/ Gain Phase Analyser.

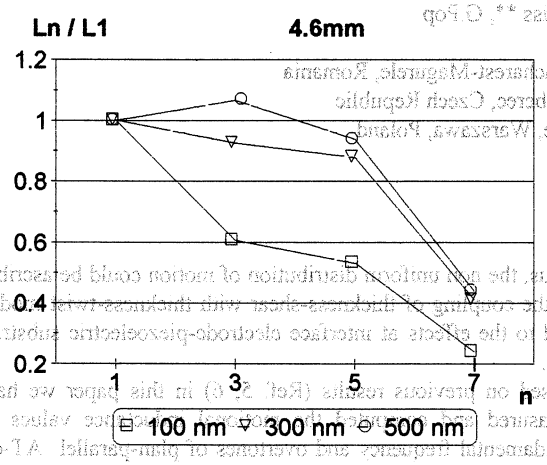


Fig. 1 Harmonic dependence of normalized inductance of AT-cut resonators with 4.6 mm electrode diameter

The increase of the electrode diameter from 4.6 mm to 7 mm for AT-cut resonators determines the change of sense of harmonic variation of motional inductance, while for SC-cut resonators this change was not observed (Figure 1, 2, 3, 4). Because the SC-cut is stress-compensated, the stress effect is very low and the inertial effect prevails. In AT-cut resonators the stress and inertial effect are present and with the change of electrode diameter one of them becomes more important. It is possible to associate the decrease of the inductance to inertial effects and the increase of the inductance to stress effects.

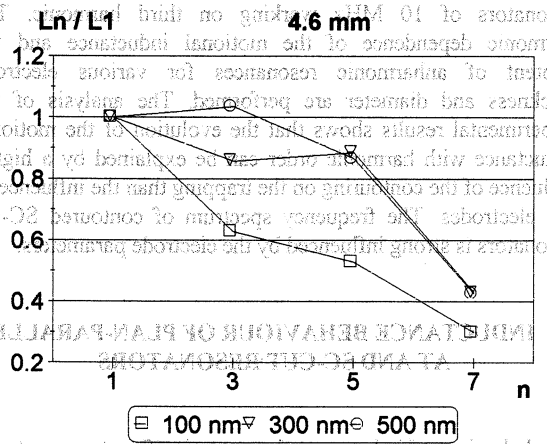


Fig. 2 Harmonic dependence of normalized inductance of SC-cut resonators with 4.6 mm electrode diameter

In the case of AT-cut resonators with small electrode diameter (4.6 mm), the inertial effects prevail over the stress effects and the behaviour of AT-cut and SC-cut are similar (Figures 1, 2). The increase of electrode thickness decreases the inductance values for AT -cut and SC-cut resonators with small electrode diameter.

The effects of internal stresses grow with electrode diameter (7 mm) and become more important than the inertial effects, thus the inductance of AT-cut resonators increases with harmonic order (Figure 3).

Figure 3 shows a slight decrease of inductance between 5th and 7th harmonics for AT-cut resonators with 500 nm electrode thickness because the inertial effect is greater than the stress effect.

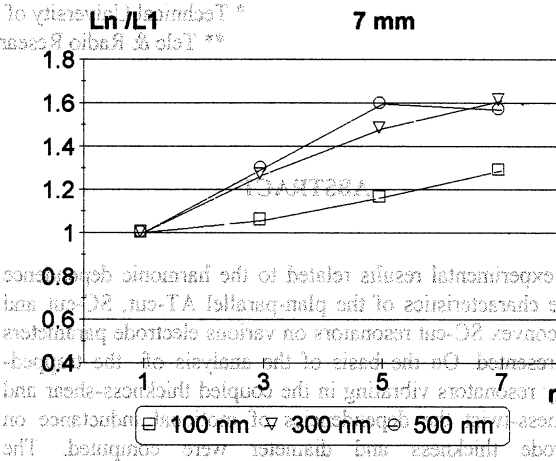


Fig. 3 Harmonic dependence of normalized inductance of AT-cut resonators with 7 mm electrode diameter

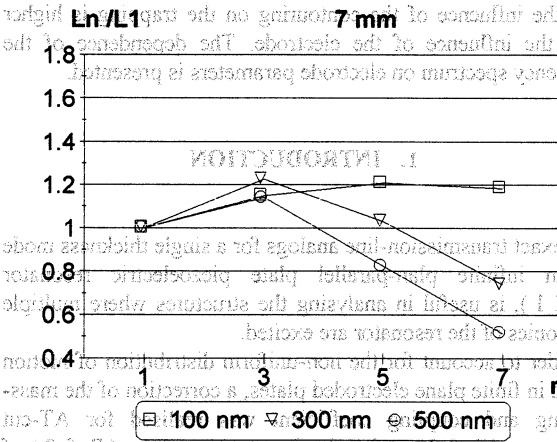


Fig. 4 Harmonic dependence of normalized inductance of SC-cut resonators with 7 mm electrode diameter

For fundamental and third harmonic and thin electrode (100 nm) the inductance of SC-cut resonators increases slightly with harmonic order due to the stress effect. This behaviour is in a good agreement with theory based on analysis of the trapped-energy resonators vibrating in coupled thickness shear and thickness twist mode (Ref. 5). For high frequencies (5th and 7th overtones) and 300, 500 nm electrode thickness, the inductance of SC-cut resonators decreases due to the inertial effects.

The theoretical calculation of the motional inductance for various electrode diameters were performed using the analysis of AT-cut and SC-cut trapped-energy resonators in vicinity of a coupled thickness-shear and thickness-twist resonances based on Tiersten's theory (Ref. 3).

The experimental and calculated values of inductance for AT and SC-cut resonators for fundamental frequency and 3rd, 5th,

7th harmonics, 100 nm Au electrode thickness and 4.6, 6, 7 mm electrode diameters are presented in Figure 5 and 6. In these figures, for 6 mm electrode diameter, there was observed that the inductance variation for all frequency range is minimum. In this case the additional effects (inertial and stress) reciprocally compensate. For the same electrode parameters the best agreement between theoretical and experimental results are obtained.

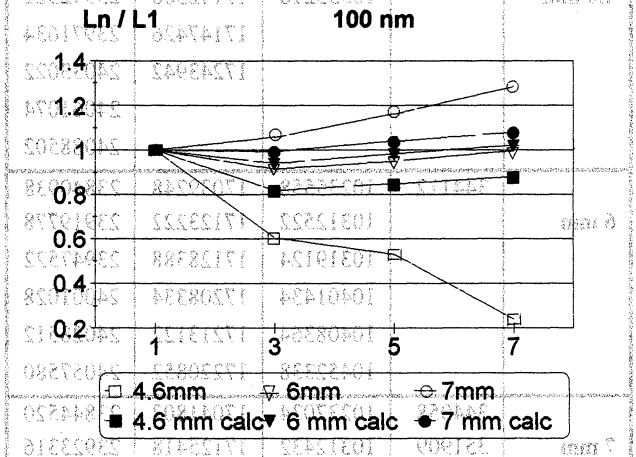


Fig. 5 Harmonic dependence of normalized inductance of AT-cut resonators with 4.6, 6, 7 mm electrode diameters

The discrepancies between theoretical and experimental values of the inductance for small and big electrode diameters could be explained by the stress effect for 7 mm diameter and by the inertial effect for 4.6 mm diameter.

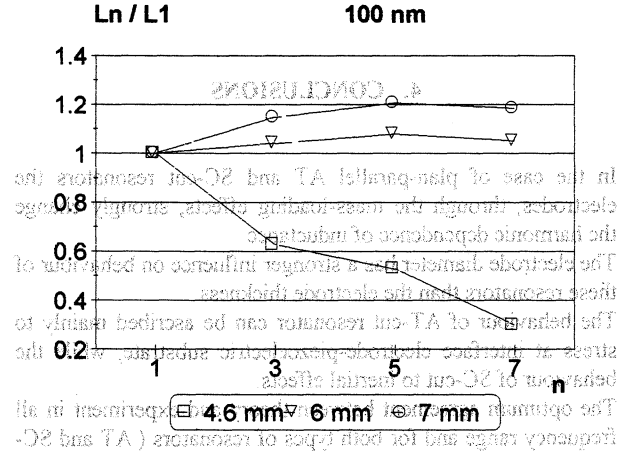


Fig. 6 Harmonic dependence of normalized inductance of SC-cut resonators with 4.6, 6 and 7 mm electrode diameters

3. INDUCTANCE BEHAVIOUR AND ANHARMONIC SPECTRUM OF PLAN-CONVEX SC-CUT RESONATORS

The study performed in this paper on SC-cut plan-parallel resonators was extended to the plan-convex SC-cut resonators with the orientation $\theta=22°10'$ and $\phi=34°07'$, 14 mm plate diameter, 750 mm curvature radius, working on 10 MHz (third harmonic).

The resonance frequencies of the fundamental, 3rd, 5th and 7th harmonics for every mass-loading and electrode diameter were measured using an HP 4194A Impedance/ Gain Phase Analyser. The anharmonic frequency spectrum at fundamental, 3rd, 5th and 7th harmonic modes were measured too.

The behaviour of the plan-convex SC-cut resonators was analysed for the same electrode diameters of Au electrodes with 75 nm, 125 nm, 175 nm and 225 nm thicknesses.

In Figures 7-10 are presented the harmonic dependencies of motional inductance for 4.6, 6 and 7 mm electrode diameters and for each electrode thickness.

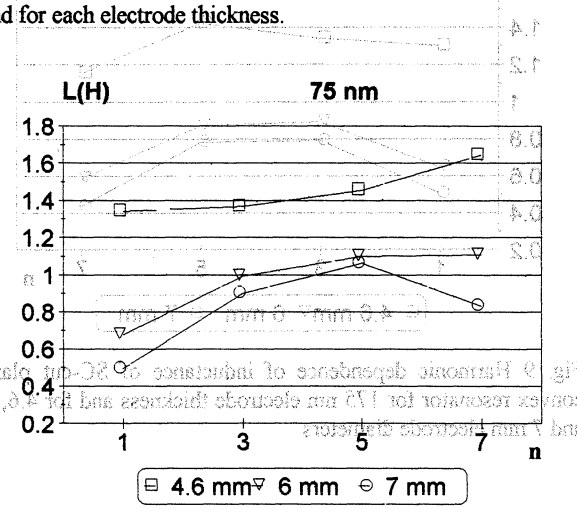


Fig. 7 Harmonic dependence of inductance of SC-cut plan-convex resonator for 75 nm electrode thickness and for 4.6, 6 and 7 mm electrode diameters

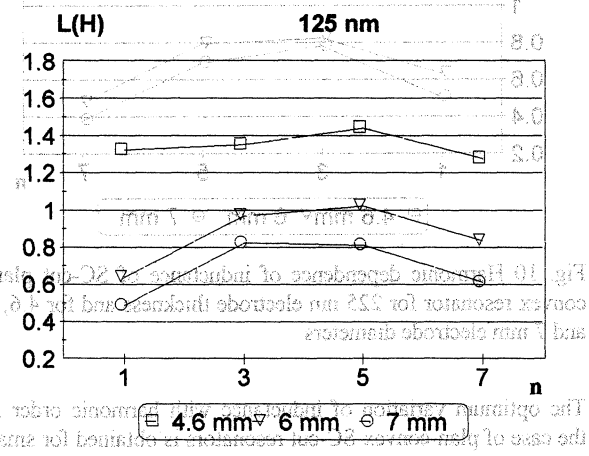


Fig. 8 Harmonic dependence of inductance of SC-cut plan-convex resonator for 125 nm electrode thickness and for 4.6, 6 and 7 mm electrode diameters

By analysing these figures we can see that for small electrode diameter (4.6 mm) the motional inductance is almost constant for fundamental, 3rd, 5th and a slowly variation at 7th harmonic depending on electrode thickness. In the case of 6 and 7 mm electrode diameters the harmonic dependence of inductance presents a slight decrease with electrode thickness.

The change of electrode diameter produces a greater variation of inductance than the change of electrode thickness.
Anyway, the influence of mass-loading on behaviour of harmonic dependence of motional inductance is smaller than in the case of the plan-parallel SC-cut resonators due to the larger variation of plate thickness comparing with the discontinuity between the electroded and unelectroded regions.

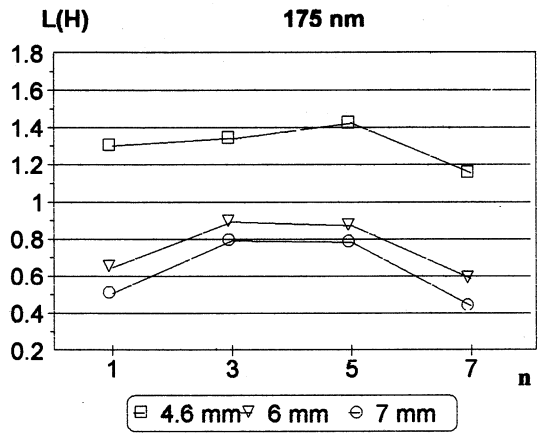


Fig. 9 Harmonic dependence of inductance of SC-cut plan-convex resonator for 175 nm electrode thickness and for 4.6, 6 and 7 mm electrode diameters

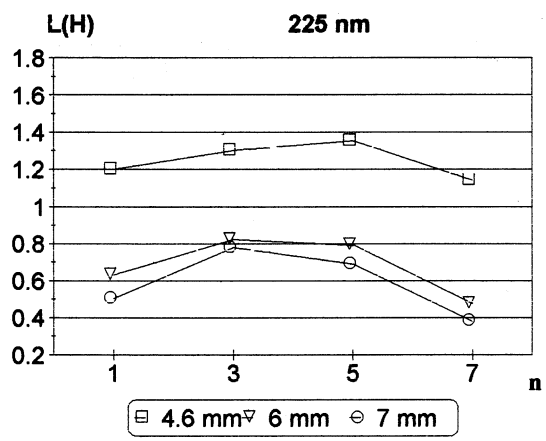


Fig. 10 Harmonic dependence of inductance of SC-cut plan-convex resonator for 225 nm electrode thickness and for 4.6, 6 and 7 mm electrode diameters

Thc optimum variation of inductancc with harmonic ordcr in thc case of plan-convex SC-cut resonators is obtained for small electrode diameter and thickness (4.6 mm, respectively 100 nm).
The explanation of this behaviour comes from the modes in plan-convex resonators being highly trapped in the vicinity of the centre of the plate.
The influence of the electrode parameters on the content of anharmonic modcs was cvidcnccd by mcasuring thc harmonic and anharmonic resonances of thc plan-convex SC-cut resonators.
The Table 1 presents the measured frequency spectrum of SC-cut plan-convex resonators with 125 nm electrode thickness and 4.6, 6 and 7 mm electrode diameters. With the increase of electrode diameter more anharmonic resonances are trapped and they come near the harmonic resonances.

Electrode diameter	f(Hz) for n=1	f(Hz) for n=3	f(Hz) for n=5	f(Hz) for n=7
4.6 mm	344603	10245178	17054312	23861698
		10333218	17142386	23942522
			17147426	23971634
			17243942	24033022
				24061074
				24088502
6 mm	344112	10235558	17039248	23840938
		10312522	17123222	23919778
		10319124	17128388	23947522
		10401434	17208334	24001028
		10408364	17213122	24028512
		10452338	17220852	24057580
7 mm	344058	10237034	17041802	23844520
	351909	10312432	17125418	23923316
		10318946	17130350	23950742
		10392032	17207664	24003768
		10396342	17212800	24030876
		10405258	17219656	24058176

Table 1 Harmonic and anharmonic resonances of SC-cut plan-convex resonators for electrode thickness 125 nm and electrode diameters 4.6, 6 and 7mm

4. CONCLUSIONS

In the case of plan-parallel AT and SC-cut resonators the electrodes, through the mass-loading effects, strongly change thc harmonic dcpcndcncc of inductancc.
The electrode diameter has a stronger influence on behaviour of these resonators than the electrode thickness.
The behaviour of AT-cut resonator can be ascribed mainly to stress at interface electrode-piezoelectric substrate, while the behaviour of SC-cut to inertial effects.
The optimum agreement between theory and experiment in all frequency range and for both types of resonators (AT and SC-cut) is obtaincd for thin clcctrodc (around 100 nm) and for medium electrode diameter (around 6 mm).
For these electrode parameters the inductance has a minimum variation with harmonic order.
The mass-loading effects for plan-convex SC-cut resonators are diminished because the influence of the contouring on the trapping is much greater than the influence of the electrodes.
The harmonic dependence of inductance changes more with clcctrodc diamctcr than with clcctrodc thickncss, but lcss than in the case of plan-parallel resonators.
The minimum variation of inductance with harmonic order is obtained for thin electrode (around 100 nm) and small electrode diameter (around 4.6 mm).

The measured anharmonic frequency spectrum contains more resonances for plan-convex resonators than for plan-parallel ones and there are present especially at high harmonics (5th and 7th) and for large diameters.
The electrode parameters can change the behaviour of motional inductance on harmonics for plan-parallel resonators vibrating in thickness-shear mode, while this influence is strongly diminished in the case of plan-convex resonators.

5. ACKNOWLEDGEMENTS

The authors wish to thank Dr. M. Popescu for his helpful comments and suggestions.
The authors express their gratitude to Dr. Suchanek for the AT and SC-cut samples.
For their help in the measurements and numerical calculation by computer the authors would like to thank to C. Ghita and C. Bora. The authors wish to acknowledge the assistance of E. Jelea and P. Farcasanu for the preparation of the resonators.

6. REFERENCES

1. A. Ballato, 1972, "Transmission-line analogs for stacked piezoelectric crystal devices", Proc.of the 26th ASFC, pp.86-91

2. J.Kosinski, A.Ballato, I.Mateescu, I.V.Mateescu, 1994, "Inclusion of non-uniform distribution of motion effects in the transmission-line analog of the piezoelectric plate resonator : theory and experiment", in Proc. of the IEEE SFC, pp. 229-235

3. H.F. Tiersten, 1976, "Analysis of trapped-energy resonators operating in overtones of coupled thickness-shear and thickness twist", J. of the Acoust. Soc. of Amer., vol. 59, pp. 879-888

4. J. Zelenka, 1996, "Influence of electrode mass-loading on the electrical equivalent circuit of the trapped-energy AT-cut quartz resonators", Proc. of the 10th EFTF, pp. 234-237

5. I. Mateescu, J. Zelenka, I. V. Mateescu, 1996, "The contribution to the non-uniform distribution of motion effects obtained for the SC-cut quartz plates", Proc. of 50th IEEE SFC, pp. 405-411

6. I.Mateescu, I.V.Mateescu, G.Pop, 1996, "Comparison of mass-loading effects on the characteristics of AT and SC-cut resonators", Proc. of 9th PIEZO '96, in print

7. H.F. Tiersten, R.C. Smythe, 1979, "An analysis of contoured crystal resonators operating in overtones of coupled thickness-shear and thickness-twist", J. of the Acoust. Soc. of Amer., vol. 65, nr. 6, pp. 1455-1460

8. H.F. Tiersten, D.S. Stevens, 1982, "An analysis of contoured SC-cut quartz resonators", Proc. of the 36th AFCS, pp. 37-45

9. D.S.Stevens, H.F.Tiersten, June 1986, "An analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variation", J. of the Acoust. Soc. of Amer., vol. 79, nr. 6, pp. 1811- 1826

Miniature SAW duplexer for ISM mobile phones

V. P. Plessky, S. N. Kondratiev

Micronas Semiconductor SA, CH-2022 Bevaix, Switzerland

ABSTRACT

Two ways of reducing coupling coefficient for 36-$LiTaO_3$ substrates are discussed: 1) transducer with dummy electrodes and 2) SiO_2 sublayer between the electrodes and the substrate. Theoretical coupling of modes (COM) model is used to find conditions for twofold decrease of coupling. Experiments are made with test resonator-type structures for extraction of COM parameters and the obtained data used for the SAW duplexer design.

1. Introduction

For many applications SAW filters are demanded with a relative passband close to 1%. The main difficulty to design such filters is the absence of a suitable material with $K^2 \approx 2\%$. For ladder type Impedance Element (IE) Filters the passband is primarily determined by the frequency gap $\delta f/f$ between the resonance and anti-resonance (R-a-R) of the used resonator. In many cases long transducer (including a few reflectors) is used as the resonator, and the R-a-R relative frequency difference is proportional to the $\Delta V/V$ parameter: $\frac{\delta f}{f} \approx 0.9 \frac{\Delta V}{V}$ (for single electrodes) [1] . It is clear that 64-$LiNbO_3$ or 36-$LiTaO_3$ leaky wave substrates frequently used for low loss filter designs have a too strong piezoelectric coupling, while quartz , lithium tetraborate and 112-$LiTaO_3$ are too weak. For a SAW ladder structure there is a very limited possibility to use tuning ("expanding") inductances. Therefore, a possible solution is to decrease artificially the coupling efficiency of 36-$LiTaO_3$. From the coupling of mode analysis (COM) it follows that $\frac{\delta f}{f} \propto \frac{\alpha^2}{C_{st}}$ [1] , where α is COM coupling parameter, C_{st} is the static capacitance. One can see that there are two ways to decrease R-a-R gap: either to increase static capacitance C_{st} or to decrease the coupling efficiency α of the electrode system. The static capacitance can be increased relatively easily by incorporating an additional interdigital capacitor in parallel to each of IEs used. The electrode pitch and orientation must be chosen so that undesired SAW generation in this capacitor structure will be minimized. The disadvantage of this approach is the larger chip area required, potential parasitic responses available from the additional transducers and the increased sensitivity to the additional capacitance magnitude variations.

Another alternative is to decrease coupling, where there are several possibilities: transducers with not all electrodes active, waveguide (aperture) weighted transducers, phase weighted transducers. For a transducer with dummy fingers the R-a-R frequency gaps is reduced proportionally to the percentage of active electrodes. For example, if we flip the polarity of every 4th electrode (...+ - - - + - - - + ...) we reduce the R-a-R gap approximately by half, and we can use this structure to design IEFs with 1% to 1.5 % relative passband. The subharmonic response could be a potential problem for certain applications, but with the advantage of increased power handling capability.

One more evident possibility is to separate electrodes from the substrate by some passive dielectric layer. This technique has been used for many years [2] and it is known that due to the large difference in dielectric constants the coupling strength decays very rapidly with the increase of the sublayer thickness. Because of this strong dependency this approach was rejected recently [3] as a method to regulate coupling when a relatively small decrease (10%) is demanded. We will show that the method works for a 50% decrease.

The paper has the following structure: In Section 2 we will give simple theoretical arguments for both the dummy fingers transducer approach and also the SiO_2 sublayer technique. Then in Section 3, experimental data extracted from the test structure measurements will be presented. Also briefly discussed is the application of this technique for ISM duplexer design.

2. Theoretical considerations

2.1 Transducer with dummy electrodes

If we decrease the number of active (hot) fingers in a long transducer resonator structure replacing them with dummy electrodes, we decrease coupling coefficient α and capacitance roughly proportionally to the density of

hot fingers. But one can see from the formula $\frac{\delta f}{f} \propto \frac{\alpha^2}{C_{st}}$ that in this case the resonance anti-resonance frequency gap will decrease. We have simulated this effect using the coupling of mode (COM) model for transducers and gratings. The structure we considered included a few parallel connected equivalent transducers separated by reflectors (all identical), the pitch of the structures being the same with all structures placed synchronously, thus observing complete periodicity of electrodes with no additional spacers used. The structures were cascaded using a group of one transducer and two reflectors on each side of it and then repeating the cascading procedure (see figure 1) .

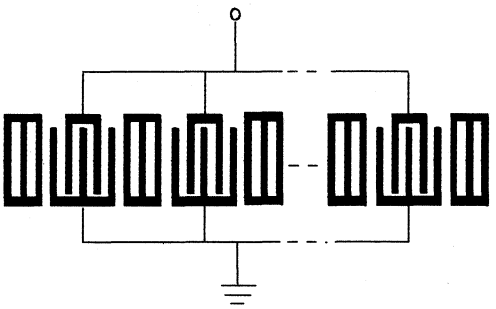


Fig. 1: Example of structure used for simulation of R-a-R gap dependency on dummy electrodes.

Obviously , the coupling efficiency can be changed in a wide range. If the transducer and the grating occupy equal part of the structure the R-a-R frequency gap is reduced approximately by half. One can expect the capacitance per wavelength to be reduced proportionally to the density of hot electrodes, but the capacitance of the busbars must be taken into account with this contribution becoming relatively more important for the transducer with reduced number of hot fingers.

2.2 SiO_2 sublayer

The SiO_2 sublayer not only decreases the coupling efficiency of the leaky SAW on 36-$LiTaO_3$ but also changes the wave velocity and the reflectivity of *Al* fingers. The velocity increases initially with a corresponding increase of the silicon dioxide layer thickness, due to the effect of short-circuiting of SAW electric fields becoming weaker (the $\frac{\Delta V}{V}$ effect tends to disappear). But with a further increase of the layer thickness the mass loading effect starts to dominate and the velocity of the wave decreases again. It means that there is a range of thicknesses wherein the SAW velocity is relatively insensitive to the sublayer thickness variations. This effect was recently observed by F. S. Hickernell and co-authors [4]. The values of thicknesses corresponding to the maximum velocity is in the 2% - 4.5% range. The decrease of coupling due to the dielectric layer can be easily calculated using the effective permittivity concept [2]. The analysis shows that if we want to decrease the R-a-R relative frequency distance to about 1%, we will need the thickness of SiO_2 close to 2% (Fig. 2).

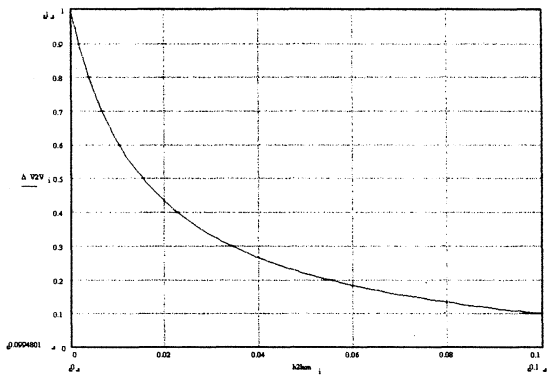


Fig. 2: Decrease of $\frac{\Delta V}{V}$ parameter as a function of SiO_2 sublayer thickness.

Therefore, there are a few positive points using a SiO_2 sublayer: the needed layer is reasonably thin, the leaky SAW velocity and reflectivity from fingers seems to be relatively insensitive to the layer thickness, the SAW attenuation was found to be minimal at 2% thickness [4] and , finally the SiO_2 layers are known to decrease temperature coefficients of delay (TCD) of Lithium Tantalate.

3. Experiments

In order to investigate experimentally the SAW propagation parameters we used the technique based on careful analysis of the admittance curve of a single port synchronous resonator structure previously described in [5].

Figures 3a and 3b show the admittance curves for normal single electrode transducer test structure, (which included 240 fingers in transducer and 50 fingers in short-circuited reflectors on each side of it), and for similar structure in which every 4th electrode in the transducer was connected to the "wrong" busbar. to the "wrong" busbar. One can see that the resonance/anti-resonance distance is decreased from roughly 25 MHz to about 10 MHz, which corresponds to the targeted value.

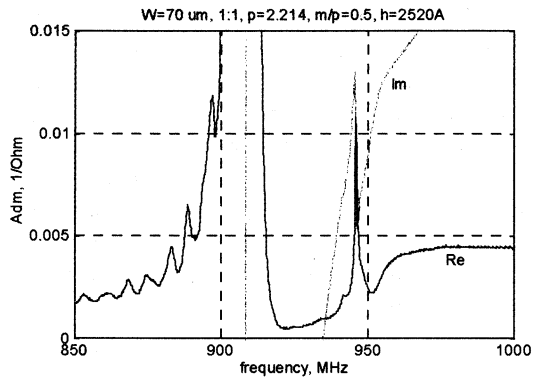


Fig. 3 a : The admittance curves for normal (1:1) transducer test structure.

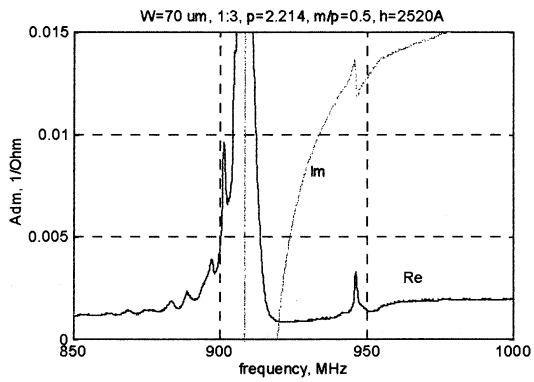


Fig. 3 b : Admittance curves for the (1:3) test structure with 1 hot and 3 ground electrodes in a period .

As to the leaky SAW velocity and reflectivity of electrodes -- the magnitudes are quite similar.
In the case of SiO_2 sublayer (Fig.4) the reduction of coupling was decreased to about a 10 MHz R-a-R gap .

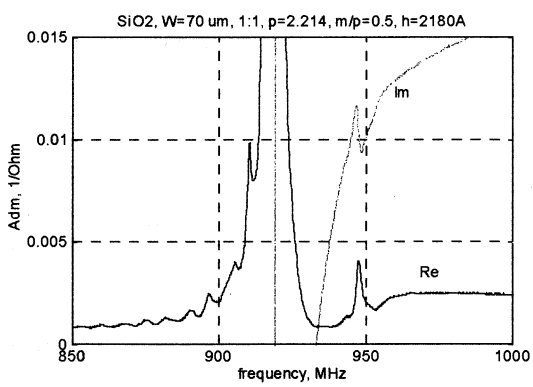


Fig . 4 : The same structure as in Fig.4a, but with 800 Å SiO_2 sublayer.

This is accompanied by an increase of the wave velocity and decrease of the electrode reflectivity. The latter effect can be clearly seen from reduced width of the stopband.

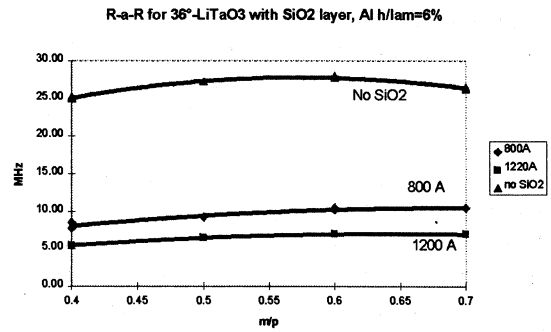


Fig. 5 : Decrease of the resonance/anti-resonance frequency gap.

Both approaches have been used for the development of SAW filters for mobile communications with about 0.5% to 1% passband. A current example is a SAW duplexer for ISM phone. As previously discussed , there is no material exactly suitable for this specification. Figure 6 shows the results obtained with the "thinned" hot electrodes transducer approach. The ladder IE filter scheme was chosen to provide low loss in the passband, steep skirts and high rejection levels.

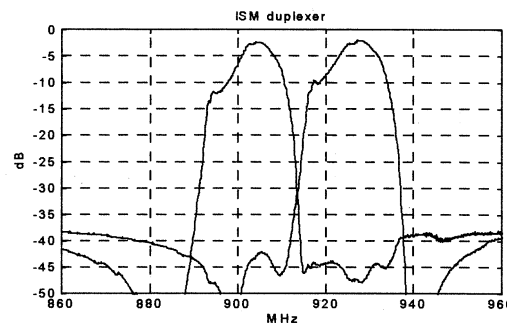


Fig. 6 : ISM duplexer performances, "thinned" hot electrodes on 36-$LiTaO_3$.

The period of the electrode structure included 4 electrodes, connected in 1:3 sequence to the corresponding busbars. In a sense, we work on 2nd harmonic of this transducer. Figure 7 shows frequency response in a wide frequency rande .One can observe the subharmonic responses at about 1/2 and 3/2 central frequency (in reality higher at low frequency and lower at high frequency than main response) due to fundamental mode operation and dispersion . Fortunately, subharmonic response amplitudes exactly at 1/2 and 3/2 central frequency are about -45 dB . It is

very important for portable telephone performances. Another consequence of using 4 electrodes per period , which might have been expected, would be an increased attenuation due to the bulk radiation and scattering of the slow shear and fast shear bulk waves. The scattering of the slow shear waves always exists for leaky waves, but is very weak for 36° cut of $LiTaO_3$, while the scattering of the fast shear waves normally appears only at frequencies higher than that of Bragg stopband. This effect is clearly seen in Figures 3 and 4 as an increase of real part of admittance (at approximately 950 MHz and higher) corresponding to the additional losses. Working at the 2nd harmonic must add a few more scattered waves, including fast shear bulk waves, which propagate approximately perpendicular to the substrate surface. This effect is almost independent of frequency in the relatively narrow range of interest. From Fig. 3b one can conclude that this additional attenuation (if any) has a relatively weak effect, i.e. not introducing dramatic changes into impedance (admittance) characteristics. Another positive feature of this 1:3 structure is that the stopband remains as wide as for normal 1:1 grating. The perturbations of the admittance curve (Fig. 3b) at the upper edge of the stopband in this case are not harmful for the Impedance Element Filters performances. For the structure in Fig. 3 the anti-resonance frequency point is situated roughly in the middle of the stopband. Also, irregularities of admittance at the right edge of the stopband will not influence the passband performance of the devices designed using such elements.

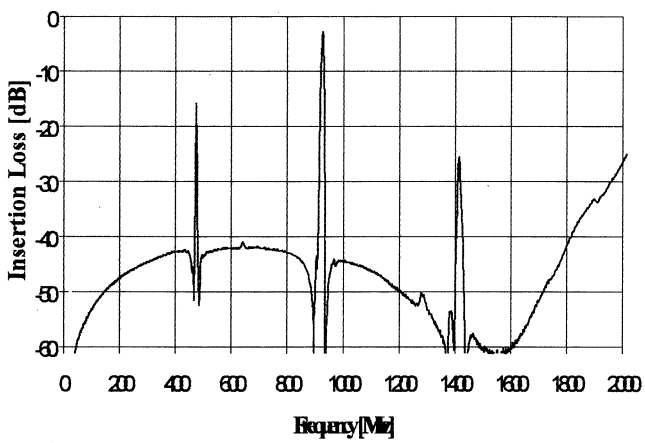


Fig. 7 : ISM duplexer performance , a wide frequency range.

4. Conclusions

For the Impedance Element (IE) SAW filters based on the use of long transducer structures as resonators, both of ladder and balanced bridge type, the optimal passband is determined by the coupling coefficient of the substrate used. We have demonstrated the possibility to modify (decrease) the coupling coefficient of 36°-$LiTaO_3$ in a wide range using one of the two following techniques: in the first approach some sections of hot electrodes transducer can be made dummy by alternate busbar connections. If it is done with every second hot electrode we get a structure for which every one electrode connected to the first busbar is followed by three electrodes connected to the second busbar (1:3 structure). For this structure the coupling coefficient α is about 2 times smaller than for "normal" 1:1 long transducer [1]. Another method is to introduce a passive SiO_2 layer between the electrodes and the substrate. The layer decreases the strength of the electric fields penetrating into the piezoelectric substrate and thus decreases coupling. It was found that 2% thick layer can decrease the coupling at about 50%. A practically important feature is that the leaky SAW velocity is relatively insensitive to the layer thickness for thicknesses ranging from 2.5% to 4% .

Acknowledgment

The authors would like to thank P. Muralt and M.A. Dubois for the preparation of SiO_2 layers and C. Lambert for the some calculations .

REFERENCES

[1] V. P. Plessky, "SAW Impedance Elements", IEEE trans. UFFC, vol. 42. No. 5, pp.870-875.

[2] D. P. Morgan, 'Surface-Wave devices for signal Processing'. Elsevier, 1985, 1991

[3] K. Hashimoto, M. Ueda, O. Kawachi, H. Ohmori, O. Ikata, H. Uchishiba, T. Nishihara, and Y. Satoh , "Development of ladder type SAW filter with high shape factor", IEEE Ultrasonics Symp. Proc. pp. 113-116, 1995

[4] F. S. Hickernell, H. D. Knuth, R. C. Dablemont, and T. S. Hickernell, "The Surface Wave Propagation Characteristics of 64° Y-X $LiNbO_3$ and 36° Y-X $LiTaO_3$ substrates with Thin-Film SiO_2", IEEE Ultrasonics Symp. Proc., pp. 345-348, 1995.

[5] C. S. Hartmann, and V. P. Plessky, "Experimental measurements of propagation, attenuation, reflection and scattering of leaky waves in Al gratings on 41°, 52° and 64°-$LiNbO_3$", IEEE Ultrasonics Symp. Proc., pp. 1247-1250, 1993.

POSTERS - T/F TECHNIQUES AND ATOMIC STANDARDS

A NOVEL CONCEPT FOR THE DIRECT INTERCOMPARISON OF THE ALLAN VARIANCE OF FREQUENCY STANDARDS AT THE 10^{-18} LEVEL WITH APPLICATION TO METROLOGY, GLOBAL COMMUNICATION AND NAVIGATION

[1] Farhat ABBAS, [2] David W. ALLAN, [3] Neil ASHBY, [4] Clifford C. HODGE, [5] Jens HOLLAND, [6] Dieter JEDAMZIK, [7] Conway D. LANGHAM, [8] Lutfollah MALEKI and [9] Thomas TROWLES

[1,4,7] Centre for Quantum Metrology, National Physical Laboratory, Teddington, Middlesex, United Kingdom TW11 0LW
[2] Allan's Time Interval Metrology Enterprise, Fountain Green, UT 84632-0066, USA
[3] Department of Physics, Campus Box 390, University of Colorado, Boulder, CO 80309-0390, USA
[5,6,9] GEC-Marconi Limited, Elstree Way, Borehamwood, Hertfordshire, United Kingdom WD6 9RX
[8] Frequency Standards Laboratory, NASA Jet Propulsion Laboratory, Pasadena, CA 91109-8099, USA

PAPER SUMMARY

Over the past ten years, the advent of new solid-state materials has produced opportunities for new laser injection wavelengths in the visible and near-infra-red, allowing the construction of compact systems with reduced sensitivity to environmental perturbations. Today, there exists a new generation of atomic frequency standards in the microwave, such as the laser-pumped space rubidium and caesium atomic clocks developed by Neuchâtel {1}, and the realisation of the caesium fountain {2} and the mercury linear ion-trap {3}. In respect, there have also been significant improvements to the active-cavity hydrogen maser {4} and the passive cryogenic dielectric resonator {5}. Some of the applications for the new millennium include global synchronisation in the benign environment of space for time-transfer and time-dissemination, with developments as diverse as the next-generation of Global Navigation Satellite Systems {6}, proposals for a collaborative space-based clock mission on-board the International Space Station platform (Alpha) {7, 8}, and optical coherent communications {9}. Yet, an optical frequency standard has the potential for even greater accuracy. Here, the major advantage of frequency standards operating in the optical rather than the microwave range is provided by their higher frequencies: the influence of level shifts is reduced, and the fractional resolution can be increased by several orders of magnitude. And in metrology, an optical frequency standard of an ultra-stable known frequency could also provide a unified standard of length, time and frequency.

Recent analysis of the performance of the mercury linear ion-trap appears to indicate its potential for an absolute fractional frequency stability in the 10^{-18} range for the ultra-violet (194.2 nm) pumping transition, and for interrelating this to the 40.5 GHz hyperfine microwave transition. We show that such a capability would permit the realisation of Doppler variations of the Earth's geopotential at the 1 centimetre level. There still remains the requirement to independently intercompare the mercury linear ion-trap with other standards of comparable stability, by linking the optical to the microwave region at the highest levels of precision. For traceability, this measurement should be undertaken with just one frequency standard. Fortunately, cryogenic interferometers may hold the solution: the key to this unprecedented optical frequency stability attributable to the Neodymium-Yttrium Aluminium Garnate (Nd:YAG) laser, which has the lowest free-running linewidth of just 3 kHz for an operating frequency of 242 THz (1064 nm wavelength) {9}. The idea of stabilising a laser to a cryogenic interferometer has been detailed in 1995 {10}. Already, for all-silica cryogenic interferometers {11}, a fractional stability of 3×10^{-14} has been realised, with a relative lock instability for two lasers of 5×10^{-16}. By using sapphire, which has a thermal expansion coefficient of more than two orders of magnitude lower than silica at liquid helium temperatures, there is the potential for correspondingly improved fractional frequency stability at the 10^{-18} level. We develop these ideas further in this paper to propose the notion of a Nd:YAG laser stabilised to a cryogenic superconducting-coated sapphire SuperMirror resonator, which would allow the fractional frequency stability in the optical to be realised at the 10^{-18} level, and permit the simultaneous intercomparison from the same frequency standard against its microwave resonance at the 10^{-16} level, for averaging times up to 10,000 seconds. This offers the prospect of the cryogenic superconducting-coated sapphire SuperMirror resonator being used as the local oscillator in support of the mercury linear ion-trap (or caesium fountain) frequency standard.

We detail the factors affecting the fractional frequency stability for an all-sapphire SuperMirror interferometer: effects that influence such as thermal expansion, permittivity, transition metal impurity levels, and so on, and go on to show that maintaining the interferometer at the lambda point of helium (2.3 kelvin) has distinct advantages over liquid helium temperature (4.2 kelvin) operation. We also provide a detailed analysis of the options available for the cryogenic SuperMirror resonator geometry, considering the superconducting cavity technology of the 1970's {12}, through to the superconducting-coated dielectric resonator technology of the 1980's {13}, and on to the superconducting-loaded sapphire dielectric resonator technology of the 1990's {14}. Some preliminary results are given towards the construction of the NPL superconducting-coated sapphire SuperMirror resonator, with particular regard to the sputter deposition technology being developed in collaboration with GEC-Marconi, and towards the demonstration of the cavity-stabilised helium-neon laser used for intercomparison against the NPL iodine-stabilised HeNe laser (the wavelength standard). In conclusion, we believe that a fractional frequency stability of 10^{-18} appears to be achievable, and that this will pave the way for exciting developments in Time and Frequency metrology, and for space applications for global communication and navigation.

Time Frequency Generation from LORAN and GPS systems

MORS Company
M. BOUTTEVILLE, CAILLIEZ

MORS is a french company recognised as a specialist in electronic and computer systems manufactured for the professional market.

In particular, MORS has developed radionavigation systems and equipments in Time-Frequency domain.

The subject of this show is to introduce our activities in these fields.

This show presents two main activities currently in progress :

- On board Time and Frequency Generation System for Navy (defense, experimentation ...).
- Equipment delivering UTC Time and Frequency references elaborated from a LORAN Signal.

The interest of the first subject lies in the capability to generate a constant quality 1 pps top even during GPS synchronization loss.

The second subject is a Time-Frequency application based on LORAN-C signals reception.

The principle is to take advantage of the LORAN wave high stability (10^{-12} at short term) to generate reference frequencies. The delivered 1 pps top is based on UTCB time (NELS network) or UTC USNO time and its accuracy is better than 100 ns.

TIME TRANSFER VIA METEOR BURST RADIO LINK

L.A. Epictetov, R.R. Merzakreev, V.V. Sidorov,
A.V. Logashin and L.V. Vladimirov

Kazan State University, Russia

The results of few time transfer experiments performed in Russia in the recent years are presented in the paper. In these experiments, the KAMA-7 meteor communication and synchronization equipment developed by Kazan State University was used. Operation of the equipment is based on two-way transfer of time signals through the meteor radio link. It provides precision time transfers and data communications simultaneously. The phase measurement technique and periodic automatic calibration of equipment phase delays on a few carrier frequencies are used in the instrument to minimize the noise floor and instrumental errors of time and time interval transfer.

In standard mode of operation, maximal difference of carrier frequencies of about 0.5 MHz is used. This mode allows the equipment to transfer time with RMS error of about 15-20 ns and time intervals with error of less than 0.5 ns. Using the precision measurements of time intervals, we can average time measurements on intervals from 15 minutes up to few hours and decrease the error of time transfer down to 5-2 ns. The precision mode, using additional carrier with frequency shift up to 10 MHz, allows the equipment to transfer time with error less than 0.5 ns.

The performances of meteor synchronization equipment were verified by experiments performed on different radio links in 1992, 1993 and 1995 years. The last two experiments was done to demonstrate a scope for using of meteor burst radio for data communications and time transfer between the ground transmitting stations of Chayka (Russian LF navigation system), and between stations of Chayka and Loran C. The Russian-Norwegian experiment conducted in January 1995 between the Chayka station in Tumanny (Russia) and Loran C station in Bø (Norway) has shown that meteor radio links may be used both as inexpensive data communication links and as supplementary means for precision synchronization in joint Chayka / Loran C chains.

New Results in Theory Relativity Effects in Satellite Sinchronisation Systems

E. Erjomin
Science Centre of Metrology (Military Standards) -
Kharkov - Ukraine

The theoretical bases of Time and Frequency Standard Sinchronization with regard to influence gravitation-relative effects and really existing time and frequency accuracy of synchronized standards were considered.

Analytic depending of the function of scale difference U(t) on gravitation-relative effects and metrology characteristics of frequency standard was got. Presence of intresting physical phenomenon: time moment of getting extremums of the time functions (type U(t)) and time-frequency (type dU(t)/dt) measurement are not correlate: $t_{min} \neq t_o$ is proved.

At this meaning of corresponding time interval $\Delta t = t_{min} - t_o$ is the function parameters of moving satellite (velocity - V_c, altitude of the trajectory - R_o), that define gravitation-relative characteristics of process synchronization and also is the function by metrology characteristics of syncronization standard : primery time scale difference U_o and relative accuracy of frequency $\Delta_o f$. Approximate correlation connecting interval Δt, parameters of moving and $\Delta_o f$ presents as :

$$t_{min} - t_o = \frac{\Delta_o f * c * R_o}{V_c^2 * (\Delta_o f - 1)} + GR \ ,$$

were c - velocity of light; GR - gravitation-relative correction.

At this appear potential possibilities high accuracy definition the time and frequency errors synchronized standards with the help of "non-traditional" methods on short-time interval measurement of U(t).

COMPARISON OF CESIUM, H-MASER AND GPS-BASED TIMESCALES

Eskelinen P., Matola M., Nuotio J. and Salpala A.
Kotka Institute of Technology, Paaskysentie 1, 48220 Kotka, Finland

1. ABSTRACT

The experiment of time and frequency transfer has been continued at the research laboratory of Kotka Institute of Technology (KIT) to find out the achievable uncertainty with different high quality frequency standards available to normal end users. The 1 pps and 10 MHz outputs of the GPS-receiver at KIT have shown a relative mutual peak difference of 20-30 ns with a 5 hour integration constant. For most of the time, both Cs1 and the H-maser agree better than 2.8 $*10^{-13}$ and the 1 pps outputs differ less than 50 ns. A 50-100 ns staircase-effect and 10-20 day fluctuations of more than 100 ns, coming possibly from the GPS-system, have been recorded. Medium term anomalies with time constants between 200 and 4000 s have been observed with the two units. Rubidium clocks are presently controlled to an uncertainty less than $2*10^{-11}$ and the national frequency service, delivered through the TV-chain, suffers from an inherent frequency offset, which exceeds $7*10^{-12}$.

Keywords: timescales, GPS clocks, cesium standards

2. INTRODUCTION

The KIT laboratory has upgraded its clock system, which was described earlier in (Ref.1) and further demonstrated in 1996 EFTF (Ref.2), by adding a new HP 58503A GPS-receiver and through the very kind donation of an Oscilloquartz primary cesium clock by Swiss Telecom. The clock data of a commercial H-maser owned by the Radioastronomical research station of Helsinki University of Technology (HUT) has been available for comparison through a joint venture, which was presented at URSI international conference in 1996 (Ref.3). Some fluctuations of the official national frequency have been monitored simultaneously for academic interest. An exchange of GPS clock data has been going on with the Swiss Telecom since June 1996.

Slight re-configuration of the overall KIT clock system has been necessary in order to better adopt the Cs-unit into the clock ensemble. Normal GPS-receivers are still used instead of the special, highly priced common-view hardware (Ref. 4) or geodetic equipment (Ref.5) and the integration time has been kept at 5 hours. The number of phase comparators has been increased and most of the controlling software for the rubidium units has been modified. As a background for the tests, studies like (Ref.6), which suggests the capabilities of GPS to exceed that of stand-alone primary standars, and (Ref. 7) for H-maser performance have been used.

3. STATISTICAL VIEWS ON SHORT TERM STABILITY

The short term stability of three clocks was compared: the Oscilloquartz Cesium 3200, HP 58503 A GPS and Efratom FRS rubidium standard. Both a digital phase comparator and a double balanced RF mixer were used together with a fast 9-digit DVM (HP 3458) as shown in Fig. 1. All results were viewed against the same relative scale although the original digital signals were 5 MHz TTL and analog ones attenuated 10 MHz sine wave. A suitable drive level was experimentally found to be around 200 - 300 mV. Before each mixer measurement the phase bias was adjusted as near 90 degrees as possible to obtain maximum linearity and sensitivity.

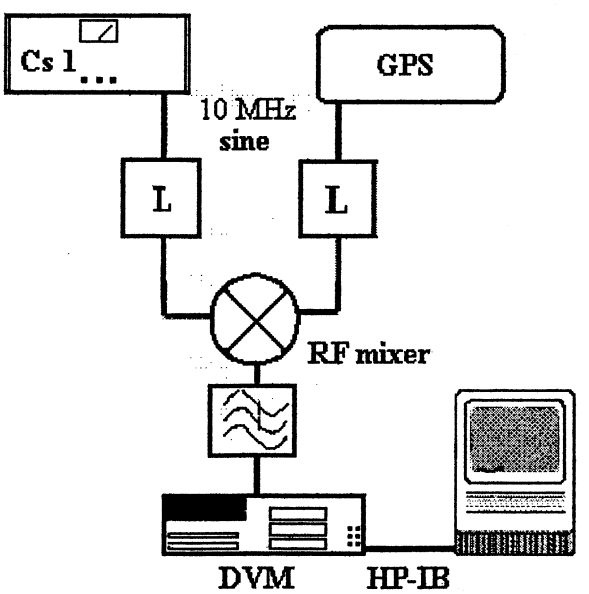


Fig.1. The test system used for short and medium term measurements utilizes a fast 8 ½ digit DVM and a phase comparator, which can be an RF mixer.

A total of 5000 samples were gathered during a 100 second interval and then plotted as a histogram. To reach the required resolution and the necessary speed all samples were stored in DVM memory an transferred later as a complete block to the computer. One typical result is shown in Fig. 2 for Rb1 and in Fig. 2 for GPS. Confidence tests were performed with the same set-up but connecting the two comparator or mixer inputs together in a conventional fashion. Test system noise level is below 10 ps. Measurements were repeated numerous times with all signals in order to reveal any possible medium term parameters which would have altered the short term result. Visual judgement shows a distinct instability of the GPS chain.

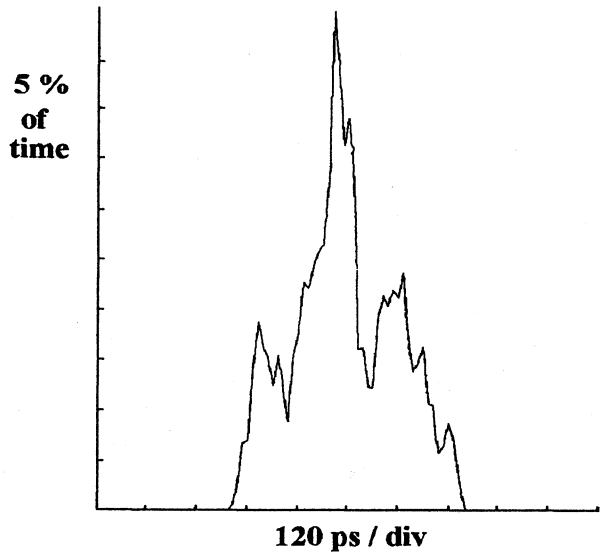


Fig.2 A time histogram of the rubidium oscillator indicates good short term behaviour (clear maximum) within 100 s.

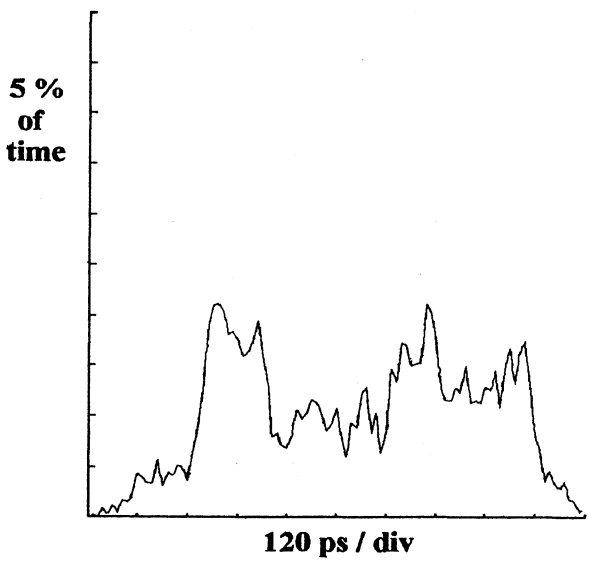


Fig.3. The GPS result in histogram form is distorted due to noise and no dominating "average" value can be seen.

4. SPECTRUM ANALYSIS OF MEDIUM TERM BEHAVIOUR

The medium term stability of the Cs1-standard, Rb1, Rb2 and that of the HP-GPS is studied separately with e.g. a spectrum analyzer combined to an ARB generator. The general scheme of the data collection system is similar to that used for short term tests. The main target was to find out possible regular fluctuations. Time constants between 200 and 4000 s have been observed.

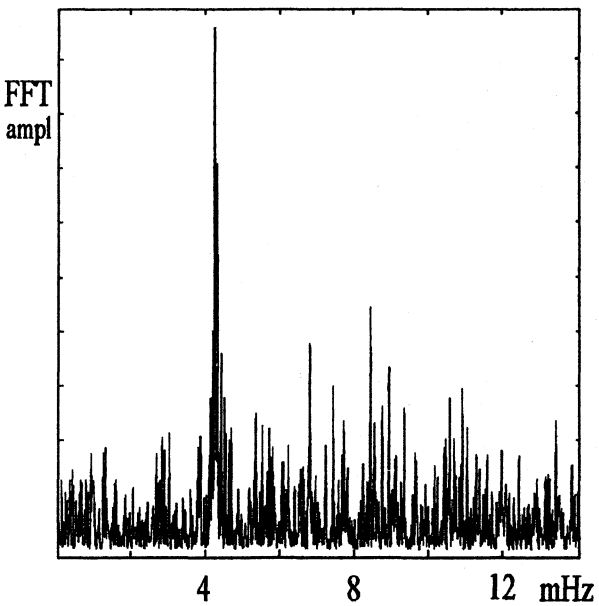


Fig. 4. A numerical FFT shows unwanted modulation of the Rb1 output at about 4.5 mHz in this medium term recording of 100 000 s.

Data was initially gathered to an HP9000 computer and then fed through IEEE-488 bus to an arbitrary waveform generator. FFT analysis was first tried with a digital oscilloscope fitted with the respective option (HP54600), but the dynamic range was not sufficient. An Advantest spectrum analyzer gave in this case better results but was not capable of a required combination of sweep time and resolution bandwidth (We hardly can work against physics !).

The best choice turned out to be a visual inspection in time domain and a experimental subtraction of very low frequency sinewave components, not within the capabilities of either numerical FFT or test equipment performance, in the HP9000 Basic code. Some twenty sinewave components were found necessary for optimization. Pure mathematical processing was done with Matlab. However, feeding the test data into different software solutions and platforms was initially by no means a straight forward action. Two plots of this analysis are shown in Fig. 4 and Fig. 5.

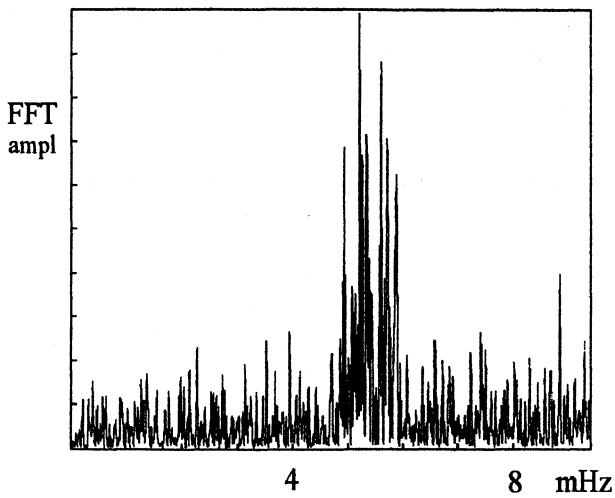


Fig. 5. Rb2 output is different from Rb1 despite of same manufacturer, as the dominating FFT result is around 5.5 mHz. Total recording time is 100 000 s.

5. LONG TERM OBSERVATIONS

The results of the first test sessions of summer and autum 1996 are, to some extent confusing. The 1 pps and 10 MHz outputs of the GPS-receiver at KIT have shown a mutual peak difference of 20-30 ns with a 5 hour integration constant. The HUT H-maser has unfortunately lost its internal lock a couple of times. For most of the time, both Cs1 and the HUT H-maser agree better than 2.8 $*10^{-13}$ and the 1 pps outputs differ less than 50 ns, when the linear term is subtracted. The comparison between Cs1 and GPS timing shows a staircase-effect (amplitude 50-100 ns), which at the time of writing, is supposed to be induced by the receiver's internal algorithm and slow fluctuations (duration 10-20 days) found by visual correlation of H-maser and Cs results, with an amplitude of more than 100 ns, coming very obviously from the GPS-system, not the receiver.

The HUT GPS-receiver (Magnavox) has more noise in its 1 pps output, as can be seen in Fig.6, although this plot is already composed of averages over several tens of individual samples. For the convenience of the reader, following Fig. 7 shows the simultaneous recording of KIT Cs1 against HP-GPS. Finally, the correlation result is drawn in Fig. 8. It is noteworthy, that the decreasing trend of timing error is further compensated starting around MJD 50278. From the end user's point of view, the two GPS receivers and the space segment do not provide a reliable long term reference for e.g. prediction of atomic clock behaviour. Without access to multiple atomic standards, one would easily suppose the drift to be induced by the local clock.

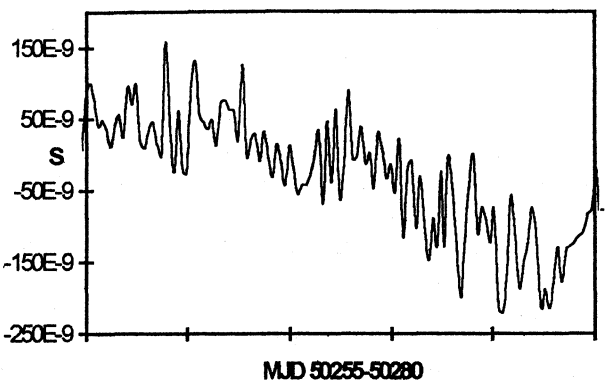


Fig.6. The HUT H-maser recording against Magnavox GPS shows much noise. However, the trend is clearly visible.

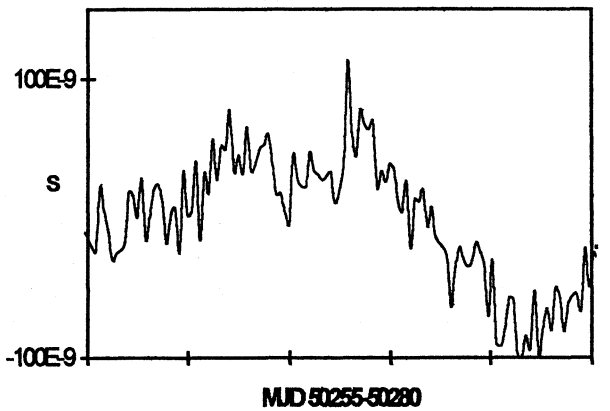


Fig.7. Here we see the KIT Cs1 measured against HP-GPS during the same test period as in Fig.6. The vertical scale is different from Fig. 6.

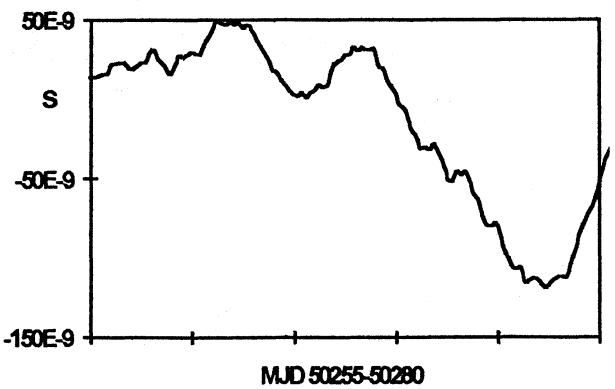


Fig. 8. A long term measurement of H-maser, Cs1 and GPS timing reveals a drift of more than 100 ns in the GPS system, occuring within 10 days.

The national frequency service, delivered through the TV-chain, suffers from an inherent frequency offset, which typically exceeds $7*10^{-12}$. Frequent switching spikes and unknown propagation delay changes, which are mandatory due to availability reasons, make the proper use of this source in the clock system very tedious.

6. OPTIMIZING THE RUBIDIUM CONTROL ALGORITM

The control algorithm for Rb2 was derived from a data set gathered during a half year test period. The basic Rb control voltage was initially kept at a fixed value (e.g. 4.195 V). A hysteresis of about 12 mV (converted from frequency domain to control voltage) was seen, when the voltage was adjusted back and forth in 25 mV steps (ten times the converter resolution) and aging and temperature related phenomena were excluded. Below this control parameter limit the respective changes in Rb frequency can be classified as noise. A control voltage step of 25 mV with a respective 3 $*10^{-11}$ frequency change was chosen. Basically, the shortest observed time interval, within which this limit is exceeded with an uncontrolled oscillator, was found to be 5 hours. The algorithm was tested and found to be correct by several successive tests. A typical result is shown below in Fig. 9.

The integartion time for the mutual comparison of GPS and Cs1 can not be made longer than the maximum allowable Rb2 control delay. Due to the relatively poor performance of the rubidium oscillator, a noticeable improvement, actually by one decade, is achieved even though some uncertainty exists in the GPS sensor output.

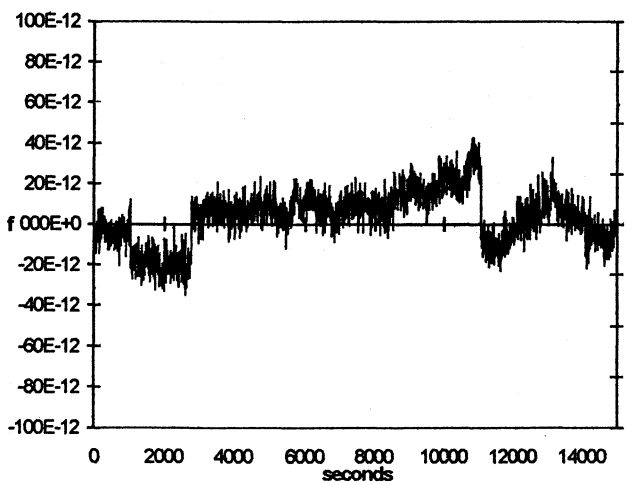


Fig.9. Optimized, direct Rb2 output frequency control is possible with a 3* 10^{-11} resolution due to the unavoidable hysteresis of the basic oscillator. Vertical axis shows frequency deviation.

7. CONCLUDING REMARKS

The owner of several high quality frequency standards, if being not one of the official time and frequency laboratories, can be a lot confused due to the apparent differences in timing information provided by the units. Particularly, the frequently and loud emphasized top performance of the GPS receivers and the total system does unfortunatly not reach all end users. Both short and medium term noise, but perhaps most, the observed day or week long anomalies hamper the rock solid use of these non-autonomous clocks.

If financial reasons or operational arrangements do not prevent, the most reliable present-day solution seems to be based on multiple ground-based atomic clocks. Performance exceeding by 3-10 times that of a commercial GPS receiver can be achieved with a 20 year old cesium as demonstrated in this paper. It is well possible, that similar long term stability can be obtained with intelligently controlled rubidium oscillators or even crystal units.

8. REFERENCES

[1] Eskelinen P. et al,"A low-cost ovenized crystal oscillator", 9th EFTF, Besancon, 1995

[2] Eskelinen P. et al,"A computer controlled multi-reference clock system", 10th EFTF, Brighton 1996

[3] Eskelinen P. et al,"GPS-based comparison of cesium and H-maser timescales",URSI/IEEE XXI Convention on Radio Science, Helsinki, 1996

[4] Davis J. et al,"Comparison of two-way time transfer and GPS common-view time transfer during the Intelsat field trial", 10th EFTF, Brighton, 1996

[5] Baeriswyl P. et al,"GPS time transfer using geodetic receivers on a short baseline", 10th EFTF, Brighton, 1996

[6] Allan D. et al, "Improved time and frequency transfer opportunities via GPS", 9th EFTF, Besancon, 1995

[7] Weiss M. and Walls F., "Preliminary evaluation of time scales based on hydrogen masers", 10th EFTF, Brighton 1996

TIME TRANSFER AND FREQUENCY SYNCHRONIZATION METHOD FOR SDH NETWORKS

Mieczysław Jessa, Andrzej Dobrogowski

Institute of Electronics and Telecommunications
Poznań University of Technology
ul.Piotrowo 3a, 60-965 Poznań
tel.: +48 61 782 631, +48 61 782 293, fax: +48 61 782 572,
e-mail: mjessa@et.put.poznan.pl, dobrog@et.put.poznan.pl

ABSTRACT

In the paper we present two methods for time transfer in the SDH networks[1]. One of these methods is an adaptation of the double ended control method used for frequency synchronization of oscillators. The second method applies radar techniques for time transfer between many nodes. The time transfer process we use for frequency synchronization of the nodal clocks in telecommunication network. We show that both processes i.e., time transfer and frequency synchronization could have completely independent control circuits. It is relatively simple if we generate two reference impulses, time reference impulse and frequency reference impulse.

Key words: time transfer, synchronization, two-way methods, echo method.

1. INTRODUCTION

The trend to near-globe telecommunication network based on SDH system forces the new approach to the network synchronization problems. The rapid increase of customer demands, multimedia services and new techniques for data transport, e.g. ATM, require to reconsider the approach to network synchronization (Ref. 12). We can't further tolerate long chains of synchronized clocks in wide-range B-ISDN networks due to the wander and jitter accumulation. The other reason is the network and customer's requirements for time signals. The common usage of GPS/GLONASS systems for time transfer and frequency synchronization contains a political risk due to the GPS/GLONASS control by specific departments of two countries. In this paper we try to develop ideas presented by Kihara and Imaoka and others (Refs. 3,5,6,7,11) and propose simple and, as we think cheaper, alternatives for the usage of satellite systems in time transfer and frequency synchronization in wide-range telecommunication networks.

[1] This work was financially supported by Scientific Research Committee (Project No TB-44-485/97-DS)

2. TWO-WAY TIME TRANSFER AND FREQUENCY SYNCHRONIZATION IN DIGITAL TELECOMMUNICATION NETWORKS

Double-ended control considered for the early frequency synchronization systems (Ref. 10) could be used for time transfer in telecommunication networks. In this case, it is simply two-way time transfer method considered for time distribution in many mediums (Refs. 1,2,4,8,10,11,13). We can distinguish two basic solutions of double-ended control applied for time transfer between geographically separated clocks. In the first one we measure two phase differences (time intervals). The first difference we measure in the local node. It is a difference between the phase of the 1 pps reference impulse R transmitted, e.g., in SOH STM-N frame header from a master node to the local node (slave) and locally generated 1 Hz reference impulse R_f. The second phase difference is measured in the master node between incoming R_f reference impulse transmitted, e.g., in SOH header from the local node to the master node and the 1 pps reference impulse R. Then, we calculate the difference of the both differences.

In the second solution of two-way time transfer we use 'returnable timing' (Refs. 9, 14) of the locally generated 1 Hz reference impulses R_f. As previously, we measure two phase differences. Both differences are measured in the local node. The first difference is a difference between the phase of the 1 pps reference impulse R transmitted from a master node to the local node (slave) and locally generated 1 Hz reference impulse R_f. The second phase difference is measured between incoming impulse R and the reference impulse R_f transmitted to the master clock (master node) and than back, to the slave clock (local node). To distinguish between the impulse generated in the local node and the returning impulse the last one we will mark as R_{ff} impulse. In the slave node we calculate the difference between both differences. It is a measure of time scales shift in geographically separated nodes.

In both methods the final result depends on the difference between signal propagation times from master node to slave node and from slave node to master node. If this difference is small enough we can use the result to control the frequency of the local oscillator and to create local 1 pps impulse R_t which position in time agrees with position of impulse R. Impulse R_t can be used next as reference impulse R for time and frequency synchronization in the next nodes. In Figure 1 we see the basic scheme for two-way time transfer between two nodes without returnable transmission of impulse R_f. The master clock generates impulses R which mark out the characteristic instants for time scale of this node. The leading edge of a single impulse R starts the time interval meter which measures the time between the occurrence of this edge and the beginning of time slot devoted for transmission of impulse R from master node to the slave node. In Figure 1 the SDH network is used for impulse R transportation from master clock to the slave clock. We change a single bit in a single byte, let us say in byte T1 proposed by Kihara and Imaoka (Refs. 7,11), from '0' to '1'. After this, the control circuit reads the result D_m of the delay between the impulse R generation and the beginning of the time slot in which impulse R should be transmitted from master node to slave node. Next, the control circuit puts the result just behind the time information (year, month, day, hour, minute, second) and transmits them in the same 64 kbits/s communication channel as reference impulse R.

The second meter located in the master node measures the time interval TI_m between the instant of receiving of impulse R_f from slave node and the instant of generation of impulse R. Impulse R_f is transmitted in the same way as impulse R but in the opposite direction. The instant of impulse R_f reception in the master node is delayed by the signal propagation time D_{s-m} from slave node to the master node and time D_s which passed between impulse R_f generation and its 'putting' into the proper time slot in the SOH header. The value of TI_m is sent in byte T2 (next byte in SOH header) to the local node. In the local node we measure time interval TI_s between the instant of impulse R receiving and the instant of impulse R_f generation. The reception of impulse R is delayed by the signal propagation time D_{m-s} from master node to slave node and the value of delay D_m between impulse R generation and the beginning of the time slot of the nearest byte T1 in STM-N frame.

After a single measurement cycle (it may lasts up to 2 seconds) we get in the slave node the following results:

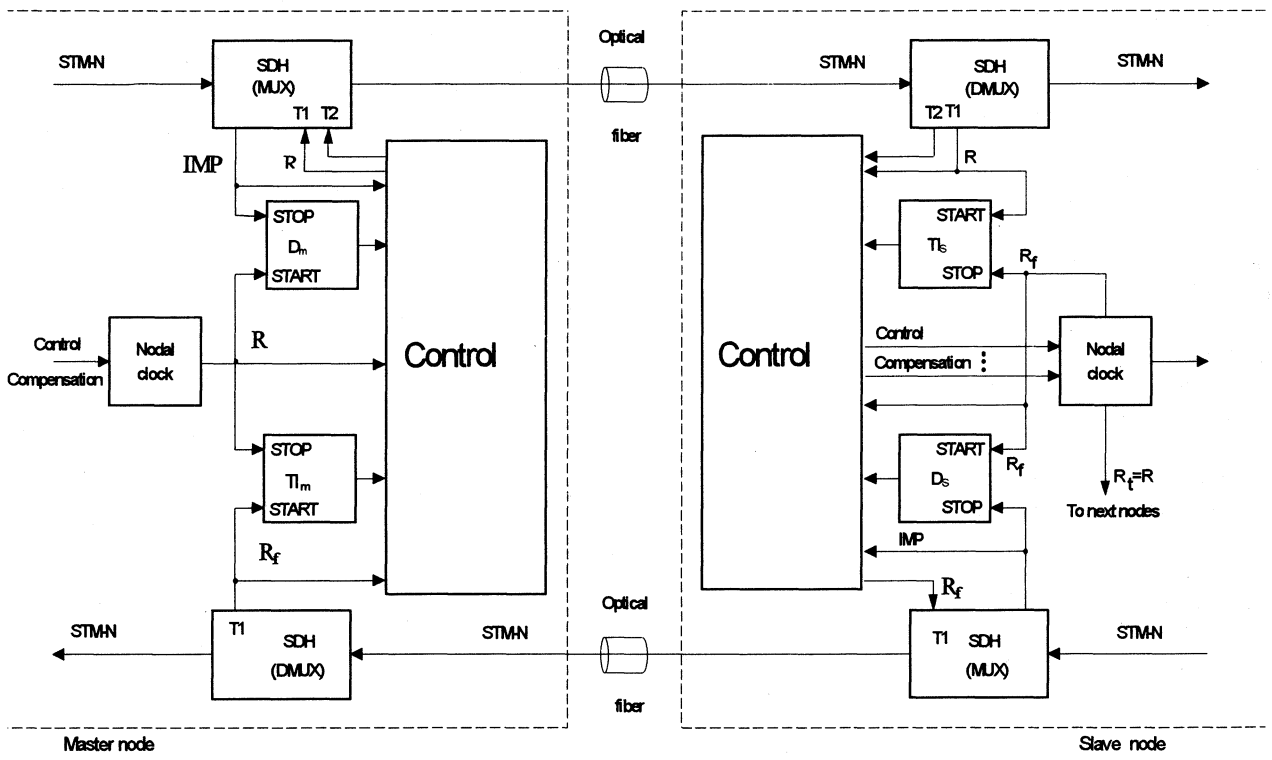


Fig.1. Two-way time transfer in the SDH network; IMP - the beginning of the time slot devoted for impulse R (R_f) transportation in SOH header.

$$TI_s(n) = [T_m(n) - D_m(n) - D_{m-s}(n)] - T_s(n) \quad (2.1)$$
$$TI_m(n) = [T_s(n) - D_s(n) - D_{s-m}(n)] - T_m(n) \quad (2.2)$$

where n=1,2,.. .Quantities $T_m(n)$ *and* $T_s(n)$ denote respectively the phase of impulse R and the phase of impulse R_f. Results (2.1) and (2.1) are obtained with 1 Hz nominal frequency and with maximal delay equal 2 seconds comparing to the instant of impulse R generation in the master node. Let us notice that the term 'phase difference' may be alternatively regarded here as 'time difference' between leading edges of 1 second impulses.

If we assume that time information which has been sent from master clock to slave clock enables the preliminary time scale synchronization with 1 s accuracy then the difference

$$\Delta(n) = T_m(n) - T_s(n) \quad (2.3)$$

is the measure for time scales difference in geographically separated clocks. We have assumed in the above formulas that the frequency of both clocks are close. That is, the frequency difference between these clocks doesn't cause a significant error during a single step of measurement. For example, the relative frequency error equal to 10^{-9} will cause 2 ns error after 2 seconds of observation. We have also assumed the change of signal propagation time during observation period equal up to 2 seconds is negligible too.

After substitution (2.3) into (2.1) and into (2.2) and after several operations we obtain

$$\Delta(n) = \frac{1}{2} \cdot \{[TI_s(n) - TI_m(n)] + [D_m(n) - D_s(n)] + [D_{m-s}(n) - D_{s-m}(n)]\} \quad (2.4)$$

For two-way time transfer and returnable transmission of impulse R_f we have

$$\Delta(n) = \frac{1}{2} \cdot \{[TI_s(n) + TI_{sf}(n)] + [2D_m(n) - D_s(n) - D_{mf}(n)] + [D_{m-s}(n) - D_{s-m}(n)]\} \quad (2.5)$$

where TI_{sf} denotes the time difference between impulses R and R_{ff} reception in the slave node. It is measured in slave node. Quantity D_{mf} is the delay between impulse R_f reception in the master node and the instant of its returnable transmission to the slave node in SOH header. As we can see the time difference between clocks in both cases depends on the unknown value for difference of propagation time of impulse R from master node to the slave node and the propagation time of impulse R_f from slave node to the master node. If we assume that $D_{m-s}(n)=D_{s-m}(n)$, n=1,2,... then we can use the error signal $\Delta(n)$ for time scale correction and frequency synchronization of the local clock.

The reference time impulse R_t is used as time reference impulse for the local node and as reference impulse R for the next nodes, slave relative to the local node. In Figure 2 we show a simple circuit for impulses R_t and R_f generation.

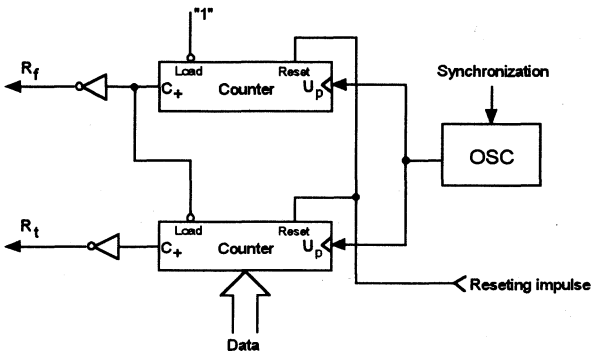


Fig.2. The generation of impulses R_f and R_t.

The circuit shown in Fig. 2 consists of two identical binary counters. They are reseted at the beginning of time transfer and frequency synchronization process. Both counters count the impulses of the reference oscillator OSC synchronized by the local clock.

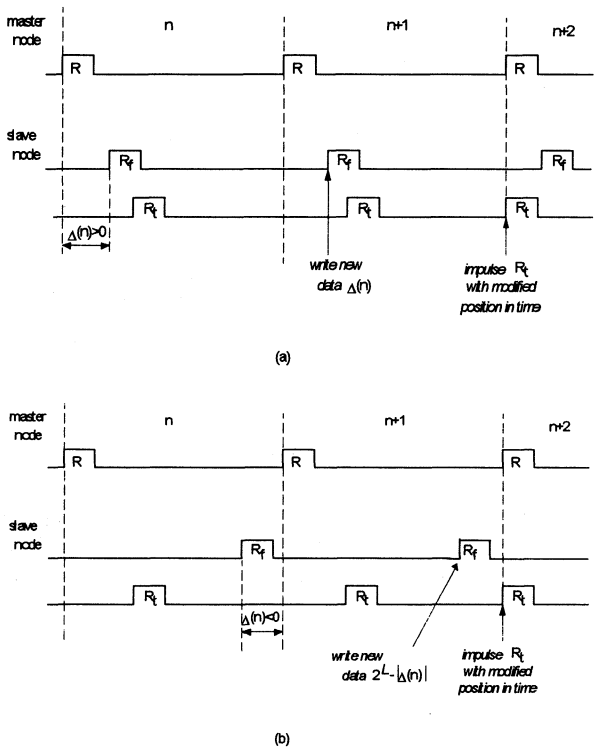


Fig.3. Time relations between reference impulses.

The upper counter counts modulo 2^L where L is the length of the counter. The lower counter is a programmable counter. It is written by new data at the beginning of each impulse R_f. The value for data depends on the sign of time error $\Delta(n)$. If $\Delta(n)$ is positive then we write to the lower counter the value $\Delta(n)$ (Fig. .3a). If $\Delta(n)$ is negative then we write number $2^L - |\Delta(n)|$ (Fig. 3b). Let us notice that the

correction of impulse R_t position can be done not earlier than 2 seconds after the last calculation of signal $\Delta(n)$ value.

If we compare both solutions of the two-way time transfer method we can notice that double-ended control without returnable transmission of impulse R_f seems to be simpler. It requires less number of measuring circuits and less data needed to be transmitted from master node to the slave node. The basic advantage of 'returnable timing' introduced for frequency synchronization of oscillators - the lack of data transmission between two nodes, disappears in time transfer procedure in SDH network.

3. ECHO TIME TRANSFER METHOD

The significant disadvantage of previously discussed solutions for time transfer in telecommunication networks is their local character. The time distribution process is carried out in a chain. So, time scale correction error is accumulated and may achieve significant values at the and of the chain. The current ITU-T recommendations allow chaining up to 60 nodes in one synchronization branch from a primary source.

The other obstacle in achieving a high accuracy time transfer in the whole SDH network is the difference in propagation times between the neighbour nodes. It can be caused by the usage of different optical links or cables. Additionally, the reconfiguration processes of the SDH network may lead to the quite different routs for reference impulses transmitted between neighbour nodes. We may of course design the fixed trails for time transfer but this solution seems to be not optimal.

If we decide ourselves on fixed trails for time transfer we may use echo time transfer method. The block diagram for this method is shown in Figure 4.

The slave node sends to the master node impulse R_f, e.g. as a light impulse. In the master node this impulse is reflected and transmitted back to the slave node in the same optical window and in the same fiber. In the master node we measure the time interval TI_m between the instant of impulse R generation and the instant of impulse R_f reception. The result of this measurement, together with the information about date and time is sent to the slave node by any data transmission network. In the slave node we measure the time interval between the instant of impulse R_f generation and its return after the reflection in the master node as impulse R_{ff}. In the slave node we have the following results

$$TI_m(n) = T_m(n) - [T_s(n) - D_{s-m}(n)] \quad (3.1)$$

$$TI_s(n) = D_{s-m}(n) + D_{m-s}(n) \quad (3.2)$$

As impulse R_f is transmitted in both directions in the same optical window and in the same fiber we can assume that $D_{s-m}(n)=D_{m-s}(n)$. After several transformations we get

$$\Delta(n) = TI_m(n) + \frac{TI_s(n)}{2} \quad (3.3)$$

The necessary condition for the usage of this method is the full bidirectionality of the medium we used for impulse R_f transmission. It means we can't use in practice any unidirectional circuits like regenerator. We estimate that the accuracy of time transfer equal to 1 ns can be achieved for distances up to 400-500 km. The basic limitation for better results is the dispersion phenomena and the current state of technology of optical transmitters and receivers. Let us notice that the echo time transfer method we can use for instantaneous time transfer from one node to many nodes when these nodes are ordered respect to the distance (along the optical fiber) to the master node. We link, e.g. K slave nodes, and the master node with the same optical fiber. The master node sends a light impulse (impulse R) to the all linked nodes. It is reflected in the successive slave nodes and goes back to the master node. In the master node we measure K time intervals $TI_m(i)$, i=1,2,..,K, between the instant of impulse R sending and its returning from the *i*th slave node. In the slave node we measure the time interval $TI_s(i)$, i=1,2,..,K between impulse R reception

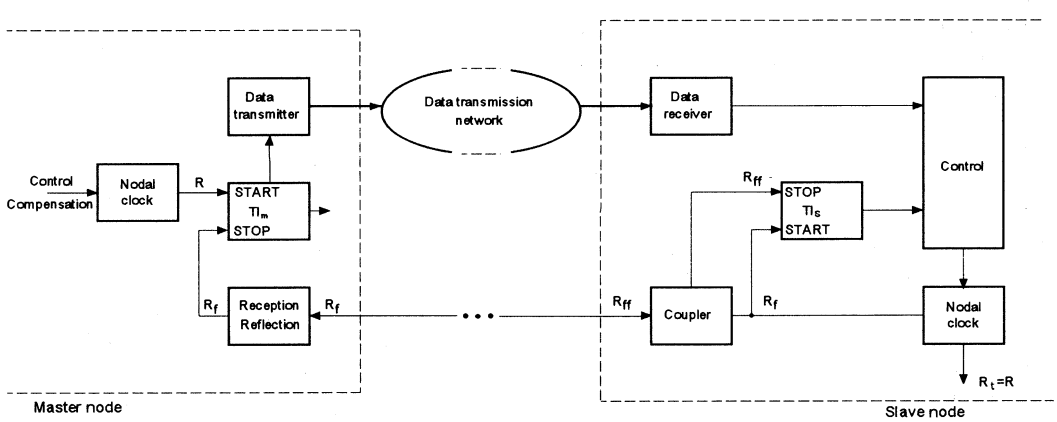


Fig.4. Echo time transfer method between neighbour nodes.

in this node and the instant of impulse R_f generation. The result of measurement obtained in the master node is sent together with time information to the proper slave node by means of arbitrary data transmission system. In the *i*th slave node we have the following results

$$TI_m(n,i) = D_{s-m}(n,i) + D_{m-s}(n,i) \quad (3.4)$$

$$TI_s(n,i) = \left[T_m(n,i) - D_{m-s}(n,i)\right] - T_s(n,i) \quad (3.5)$$

where i=1,2,..,K. Assuming that $D_{s-m}(n,i)=D_{m-s}(n,i)$ we can write

$$\Delta(n,i) = TI_s(n,i) + \frac{TI_m(n,i)}{2} \quad (3.6)$$

where $\Delta(n,i) = T_m(n,i) - T_s(n,i)$ is the time error between reference impulse R and the reference impulse R_f of the *i*th slave node.

We carry out the time correction and the frequency synchronization in the same way as described in section 2. Notice, that frequency synchronization of oscillators can be done in the existing synchronization circuits. The only change is the substitution of the error signal in the control circuits by the new signal $\Delta(n,i)$.

In Fig. 5 we see the potential scheme for the future synchronization network based on time transfer process which uses the echo method. The existing G.812 clocks are supplied with the proper control circuits devoted for time transfer. In this scheme a master clock G.812 is connected with K-1 slave nodal clocks and with one G.812 slave clock by the same optical fiber. Slave clock G.812 is a master clock for the next G.812 clock from the lower synchronization layer and for the next group of nodal clocks, e.g. SEC clocks of the SDH system.

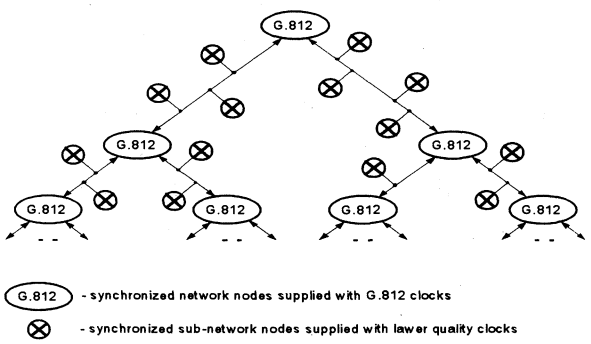


Fig.5. Network synchronization structure based on echo time transfer method.

The time error signal is accumulated in a relatively short chain of G.812 clocks. The time distribution process can be protected against failures in the same way as in the existing synchronization networks. The only thing needed is an optical fiber devoted only for time transfer process between different clocks. This limitation seems to be much more less important than in the past. The traffic increase in telecommunication networks can be served nowadays by the TDM and WDM multiplication of channels in the existing optical fibers.

4. CONCLUSIONS

In this paper two methods for time transfer in telecommunication networks have been suggested. The echo time transfer method used for a set of nodes ordered respect to the distance to the master clock seems to be the most adequate method for future applications in B-ISDN networks requiring high quality timing and time signals. We think that described solutions could be an interesting, high-quality and low-cost alternative to the usage of GPS/GLONASS system for supplying time and frequency for telecommunication networks.

5. REFERENCES

[1] Hetzel P., Polewka T.: Time from European Time Keeping Laboratories Via Telephone Modem, *Proc. of the 10th. EFTF*, pp. 540-547, 1996.

[2] Jackson D. Douglas.: A telephone-based time dissemination system, Proc. 18th Precise Time and Time Interval Meeting, pp. 541-547, Dec. 1986.

[3] Jeffers S.R. et al.: Two-Way Time Transfer Through SDH and SONET Systems, *Proc. of the 10th. EFTF*, pp. 461-465, 1996

[4] Kartaschoff P. Synchronization in Digital Communications Networks, Proceedings of the IEEE, Special Issue on Time and Frequency, pp. 1019-1028, July 1991.

[5] Kihara M., Imaoka A.: Timing and Time Signal Distribution in Digital Communications Network, *Proc. of the 6th. EFTF*, pp. 489-492, 1992.

[6] Kihara M., Imaoka A.: Accurate Network Synchronization for Telecommunication Networks, *Proc. of the 7th. EFTF*, pp. 83-88, 1993.

[7] Kihara M., Imaoka A.: System Configuration for Standardizing SDH-based Time and Frequency Method, *Proc. of the 10th. EFTF*, pp. 465-470, 1996.

[8]Kirchner D.: Two-Way Time Transfer Via Communication Satellites, Proceedings of the IEEE, Special Issue on Time and Frequency, pp. 983-990, July 1991.

[9]Lindsay W.C., Kantak A.V., Dobrogowski A.: Network Synchronization by Means of a Returnable Timing System, IEEE Trans. on Commun., vol. COM-26, pp. 892-896, June 1978.

[10]Miller M.R.:Some feasibility studies of synchronised oscillator systems for PCM telephone networks, Proc. IEE, vol. 116, pp. 1135-1143, July 1969.

[11]NTT: *Time Information Distribution,* SG 13, ITU-T Contribution, February 1996.

[12]Shankar B.: *Synchronizing Service and Customer Demand*, Telecommunications, July 1996.

[13]Sullivan D.B., Levine J.: Time Generation and Distribution, Proceedings of the IEEE, Special Issue on Time and Frequency, pp. 906-914, July 1991.

[14]Dobrogowski A: Synchronization of the Clocks of Digital Communication Network by Means of Conditionally Double-ended Control, No 136, Poznan University of Technology, 1982, Poznan (in Polish).

TV NETWORK SIGNAL AS A TEMPORARY SOURCE FOR NAVIGATION

K. Kalliomäki, J.Mannermaa

Univ. of Oulu, Dept. of Electrical Eng., FIN-90570 Oulu, Finland

Most defense forces would like to have their operational systems as independent as possible. One of the most important prerequisite to this is the self-controlled navigation. In times of peace many different ways to navigate, such as OMEGA, LORAN-C / NELS, GPS and GLONASS are available, but non of they is controlled by an ordinary military user. One possibility to built one's own navigation system is to utilize the local TV network.

TV programs are distributed to all over the country. Using a special TV receiver and good antenna, the same program can be recorded from three different fixed stations.

The relative signal delays from different stations are measured and corrected by the known link delays. Then the position can be calculated. The navigator must know the positions of TV transmitter antennas but those can be stored in computer memory. To start with, the navigator has to know their rough position with 50 to 100 km inaccuracy.

In Finland, TV synchronization pulses are locked to a rubidium standard and separate TV programs are mutually synchronized to avoid loss of synchronization when changing the channels. These make the construction of the navigation device simpler. One problem is that there is more than just one possible link route to a certain point, which makes the delay ambiguous. This problem, however, can be solved.

We have monitored the signals of atomic standards, GPS, LORAN-C and the Finnish TV network at Oulu University and VTT/AUT in Espoo for more than ten years and so have gathered a lot data which is utilized in this study. The used measurement system includes an ordinary TV receiver, a rubidium standard, GPS, multiplexer, counter and a PC.

The accuracy of the navigation seems to be about 30 m, i.e. the same as with GPS.

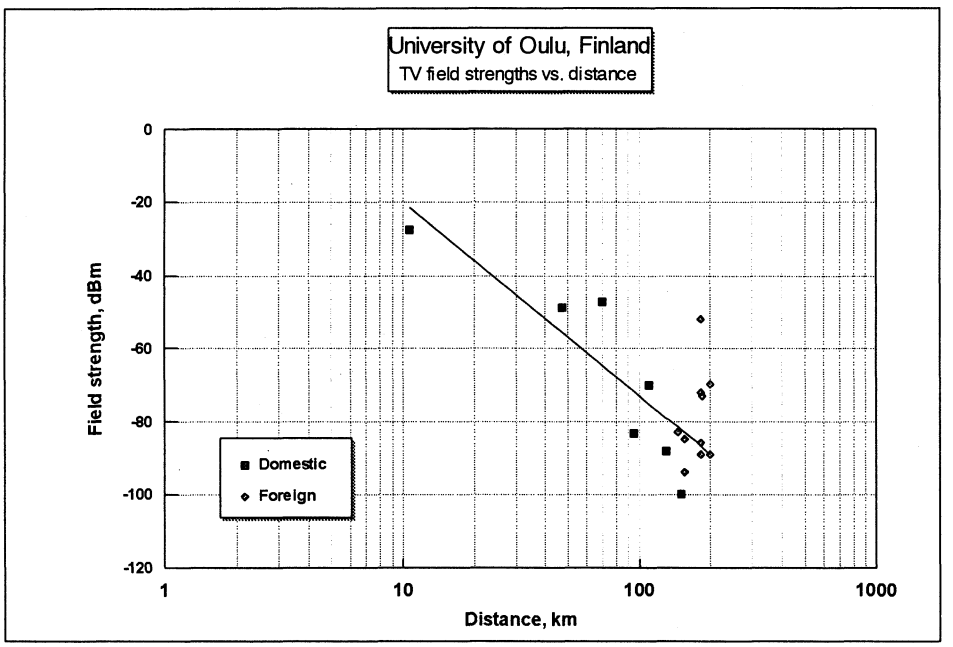
Navigation using TV-synchronization pulses

Navigation utilizing radio signals requires three or four mutually synchronized waves to calculate 2 -dimensional (2-D) or even 3-D position. Standard TV-system offers those signals without any cost.. Of course, the accuracy will be better if more than 3-signals are available. 3-D navigation is hardly possible because TV stations are practically on the same surface as user. In Finland TV-signals are synchronized to a master rubidium clock. Therefore simultaneous reception of available stations is not needed, only one multiplexed receiver is enough. The delays between different signals from TV-transmitters to user are calculated. Coarse position is obtained from frame synchronization (20 ms, 6000 km unambiguousness) pulses and fine position from the line (64 us, 19.2 km). In principle, the whole video signal can be used, too and the correlation functions between the signals are calculated to find out the delays.

Field Strength as a function of distance

Typical TV-transmitter power (ERP) is about 30 kW and the corresponding range around a station is about 50..100 km.(horizon). The signal (150...250 MHz) is strongly damped behind the visual horizon (about 60..90 km) but the navigation range can be reached up to 200 km due to narrow bandwidth reception. Field strength lower than vision threshold is enough. Following picture indicates empirical measuring results of field strengths of some domestic and foreign TV stations as a function of distance.

Fig.1. Field strengths of Finnish and foreign TV stations as function of distance.

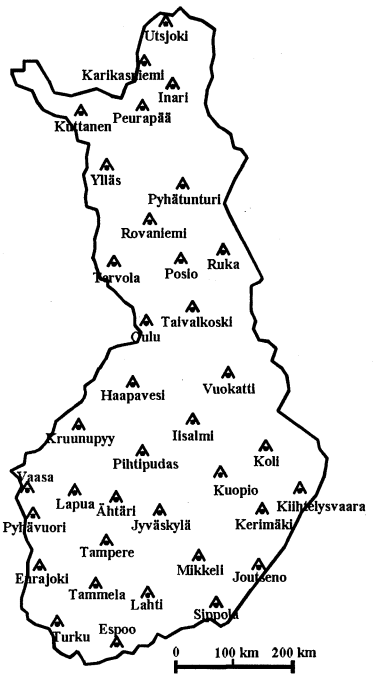


An empirical model fitted to measured values is as follows:

Eq. 1 $P_A = 33 - 53 \cdot \log (s/km)$ dBm

where s is distance and P_A power in antenna terminals.
This model is not accurate but helps to find the navigation range. Low values beyond 100 km makes the construction of a special receiver inevitable

Availability of TV transmitters for navigation purposes.



The map of the positions of the Finnish TV-stations indicates that the average distances between suitable TV stations are from 80 to 200 km in Finland. The worst case situation for navigation exits when the navigation receiver is near one transmitter, because the distances to other two necessary transmitters is at the maximum. Therefore availability analysis is not difficult, random analysis is not necessary. The situation seems to be quite good in southern Finland up to Arctic circle but in Lapland there'll arise problems.

Antennas for navigation receiver
Unfortunately standard TV-antennas are quite large and have to be situated above terrain obstacles, trees and houses etc.. In addition sharp directional characteristics in vertical plane is not wanted in navigation but a vertically stratified antenna could be used. Therefore a special antenna has to develop. Low gain (6..10 dB), erectile and rotary antenna may be one solution.

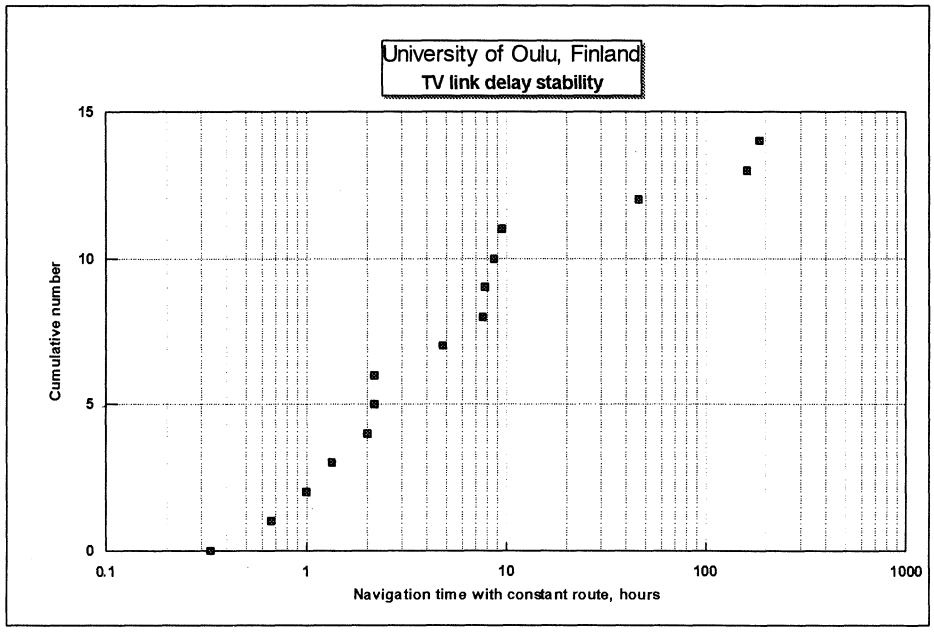
Phase stability of TV-transmitters.
TV line and frame pulses are locked to rubidium master oscillator which means that timing is not critical, the phase is stable as a function of time. The first problem is varying link connections. There are many routes to transfer TV-pictures to distant places (TX). Following table shows link delays and distribution observed at Oulu 600 km from TV-studio (Helsinki) during one month (February 1996).

Table 1 : Observed delays, their stability and probability of appearance

Link routes	*Observed Delay* μs	*SD* μs	*Number of observations*	*Percentage*	*Delay Differences*
Longest	2499.670	0.029	13	0.3	0
	2499.217	0.049	2521	58.5	0.45
	2227.357	0.036	194	4.5	271.86
Shortest	2153.622	0.050	1582	36.7	73.74

The second problem is varying program sources/studios. E.g. in Finland there are more than 10 local studios. Fortunately all programs are first sent to main studio (Helsinki), saved in digital memory and sent forward synchronized to master oscillator .

Fig. 2 Observed constancy times of delays



Short term stability depends on signal to noise ratio. In good conditions 1 ns (0.3 m) resolution in one second is possible. As an example the swing on the transmitter antenna mast in windy conditions can be observed.
Empirical phase stability of a distant TV-signal (Lahti, 110 km) was 94 ns when a standard TV-receiver and time interval counter was used. During one second 15625 samples can be obtained lowering the noise below 1ns after smoothing.

Receiver design principles

Due to low signal level the bandwidth of the receiver has to be low compared to standard TV-set (around 5 MHz). In fact only 50 Hz frame-sync pulse and 15625 Hz line frequency has to be captured. Because the phase of line synchronization pulse is critical, several harmonics of it has to take into account. Therefore the bandwidth before detector has to be around 300 kHz, which means 10 dB improvement to S/N ratio. Assuming that one fix per second is good enough, an additional digital filtering

of line phase compared to local crystal oscillator phase decreases the phase noise by a factor of 40 dB.
Phase jumps due to different link routes are more troublesome. Big jumps (> 10 μs) are easily detected and corrected by data from a link route memory. Small jumps (< 1 μs) may be a real problem if the user is moving.

Conclusions

The measurement results and analyses show, that navigation utilizing TV-stations is possible at least in south Finland. Stable link network and precision timing using atomic time base facilitate the design of a corresponding navigation receiver. Small phase jumps creating problems may be avoided during crisis conditions or a special signal (text TV) can be added to TV-signal showing the link routes in use.
The accuracy seems to be about the same as with GPS. The receiver itself will be quite clumsy due to large antenna. Therefore the use is limited to military purposes only.
References:

Results and Prospects of Radio Meteor Method of Time Scale Comparison Application in Ukraine

B.L.Kashcheyev, Yu.A.Koval
Kharkov State Technical University of Radioelectronics
B.I. Makarenko
Scientific-Research Institute of Radio Engineering Measurements

Since 1986 a high precision radio meteor comparison system for the primary standards of Ukraine and Russia is operating. The system is equipped with the instruments specimens of "Metca-6" type [1] developed in Kharkov. The system items certification in 1988 and 1995 has confirmed high metroligical characteristics of the RMC.

At present, in accordance with Agreement on Collaboration bet- ween states of the UIS in the common time and frequency support (Bishkek, 9.10.1992), comparisons are being carried out regularly along the route Kharkov-Moskow. Experimental comparisons were car- ried out along the route Kiev-Moskow-Kharkov (1993) and regular sessions were started along the route Kiev-Kharkov. For the points of Ukraine (Kiev, Kharkov, Uzhgorod and the Ground-based automated spacecraft control complex) a modernized version of the apparatus ("Metca-6M") was developed.

The "convergence" of two points comparison results is in the limits of 5 ns and of the results in the triangle (for example, Kiev-Kharkov-Moskow-Kiev) amounts to 3 ns, this testifies to a high short-term stabihty of the equipment delays and the method potential.

A new generation of high-precise comparison radio meteor equipment "Metca-11" with an error less than 10 ns is developed. Introduction of this equipment can not only increase the effective- ness of the operating system but also contribute to its structure extension. As the analysis shows the distances from the sys- tem points to the leading metrological centres of the countries- members of UIS and Europe do not exceed the limiting ones for the RMC. That is why the radio meteor comparison system extension does not require the retransmission point organization.

References

1. Kashcheyev B.L., Dudnik B.S., Koval Yu.A., Semenov S.F. Radio meteor complexes for synchronization of time and frequencez standards with errors less than 10 ns // Izmeritelnaya Tekhnika, 1992.- N 12. - P. 31-32.

PROSPECTS IN USING GLONASS SYSTEM FOR SYNCHRONIZATION OF TIME SCALES WITH HIGH ACCURACY

V. Koutcherov, S. Pouchkin, E. Bykhanov, M. Lebedev

Gosstandart of Russia,
IMVP GP VNIIFTRI,
KNITS GLONASS

Information on Russian satellite navigation GLONASS system is presented in the paper. The system forms an orbital constellation consisting of 24 spacecrafts spaced in three orbital planes inolined in longitude at 120 degrees relative to each other. The spacecrafts orbit the Earth at 64,8 degrees. The orbit hight is about 19600 km. Radionavigation field produced by the system provides visibility of not less than 4 space vehicles on the earth surface. This allows a user to determine his location, velocity and time with a high accuracy. Cesium frequency standards with a diurnal instability of several parts in 10^{13} are on borad of the spacecrafts.

Since 1989 the Main Metrological Center (MMC) (located at the IMVP GP VNIIFTRI) of the State Time and Frequency Service of Russia transmits the National time scale, UTC (SU), to the GLONASS control Center. The Main Metrological Center controls time signal transmissions with GLONASS system and uses them for comparisons of the State primary standard with secondary time and frequency standards of Russia.

Basing on investigation results for the last years a long-term stability of the GLONASS system time scale is close to 1×10^{-14} for averaging time of 10 - 60 days.

The time difference between the system scale and UTC (SU) does not exceed 100-300 ns (taking into account a transmitted correction, τ_c).

At an early date some of the actions will be taken to close the system GLONASS scale to the UTC scale.

Considering the peculiarities of formation of the GLONASS system:
- the absence of an intentional degradation in the transmitted data;
- more hign inclination of the spacecraft orbits (65 degrees in comparison with 55 degrees in the GPS systems);

one can conclude that in some cases the accuracy of time scale synchronization with the GLONASS signals will be 1,5 - 2 times higher than that using GPS signals.

TESTING MOTOROLA ONCORE GPS RECEIVER FOR TIME METROLOGY

W. Lewandowski*, P. Moussay*, P. Guerin**, F. Meyer**, M. Vincent**

*Bureau International des Poids et Mesures, Sèvres, France
**Observatoire de Besançon, Besançon, France

Abstract

In the time-metrology community the GPS time receivers most commonly used are C/A code, one-channel, one-frequency devices. They were developed in the early 1980s and their high price, of about twenty thousand dollars, has not changed. But with GPS now fully operational, the market for GPS navigation receivers is booming and one can purchase a multi-channel, GPS pocket-sized receiver for only a few hundred dollars. One of them, the Motorola Oncore 8-channel, one-frequency receiver, is of special interest for timing because it provides a 1 pps output. Preliminary tests of zero baseline common views and 320 km baseline common views between Observatoire de Besançon and Bureau International des Poids et Mesures with the BIPM international GPS common-view schedule are described in this paper.

Keywords: GPS time receiver, time metrology.

INTRODUCTION

At the Observatoire de Besançon (OB) and the Bureau International des Poids et Mesures (BIPM), Motorola receivers are connected to local HP5071A caesium clocks, each fitted with an external time intervallometer and a microcomputer (Ref. 1). The set-ups at the individual laboratories are shown in Figures 1 and 2. In addition, at the BIPM, two one-channel 'classical' time receivers, Sercel and AOA TTR5, connected to the same clock as the Motorola were used in this exercise. At the OB a Sercel receiver was used.

At the OB the first series of tests were carried out using two co-located XT Motorola receivers. The objective of the tests was to verify that a low-cost device of this kind could be used for the synchronization of the Auger Observatory, a cosmic ray project designed to observe ultra high energy particles. The time offset between the 1 pps signal from the receiver and the corresponding signal from a HP5071A caesium clock was measured. Data were acquired every second (the Auger application requires this) for each receiver. No schedule was used for this series of tests. Sessions were performed using the highest satellite in view. Scanning of the constellation was repeated every 10 minutes. Differences between the two receivers at 1 second intervals show a standard deviation of about 7 ns for sessions of up to four days (Ref. 1).

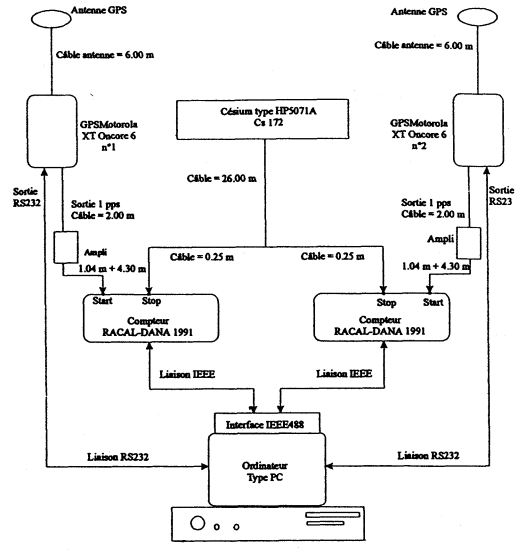


Figure 1. Experimental set-up at the Observatoire de Besançon.

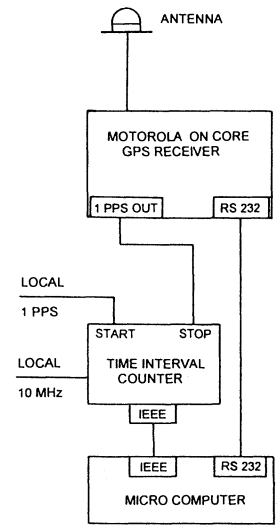


Figure 2. Experimental set-up at the BIPM.

Next, at the BIPM and the OB, tests of Motorola receivers were carried out under conditions as close as possible to those which obtain during GPS common-view clock comparisons for the generation of International Atomic Time (TAI). The one-second observations were statistically treated following a standard procedure (Ref. 2) using tracks which have a duration of thirteen minutes. However, all corrections added to the pseudo-range measurements were provided by the Motorola receiver software. It is not yet known if this software uses standard formulae and constants. This will be verified in coming tests and, most probably, software which includes standards for time metrology will be developed. For this test one-channel 'classical' Sercel time receiver and one of the eight channels of the VP Motorola receiver were programmed with BIPM international GPS common-view schedules Nos 27 and 28, each of 48 daily tracks. On 17 February 1997 the Sercel and Motorola receivers, on both sites, were programmed with an additional 41 tracks to allow permanent 24 hours tracking. The TTR5 receiver continued to observe the 48 tracks, as it is limited to this number by its software. Differential antenna coordinates for the receivers are known with an uncertainty of a few centimetres on each site, and an uncertainty of about 30 cm between the two sites.

ZERO BASELINE COMMON-VIEW RESULTS

Having thirteen-minute tracks in standard format for the Motorola receivers allowed a rigorous on-site, zero baseline, comparison in common-view with 'classical' GPS time-transfer receivers. Differences between the Motorolas and the Sercels at the BIPM and the OB, and standard deviations of individual 13-min common-view difference are shown on Figure 3. For reference, a comparison of two 'classical' GPS time receivers at the BIPM, a TTR5 and a Sercel, is reported on the same figure. It should be noted that Sercel antenna was protected by a temperature-stabilized oven, what significantly reduced its noise (Ref. 1). The noise exhibited by the time series of Figure 3 were analysed using a modified Allan variance. The series exhibit white phase noise up to an averaging interval of about twelve days.

It should be noted that the OB laboratory heating is turned off during weekends. This may explain why the OB Motorola receiver depart from the Sercel for these periods, Figure 3. It was already known that the response of some Motorola receivers depends strongly on temperature (Ref. 3). For this reason the BIPM receivers are located in an air-conditioned room.

In contrast comparison of Motorola and Sercel receivers at the BIPM, although showing better stability than that at the OB, exhibited more noise, Figure 3. This can be explained by the fact that Motorola antenna at the BIPM is covered by an oven, which diminishes the GPS signal level. This calls for further investigation. Unfortunately, the conditions under which the Motorola antenna is placed in the oven are such that the antenna cannot be considered to be temperature protected.

At the BIPM a procedure to remove constant biases between observations in different directions of the sky is used operationally for the treatment of GPS data. This makes it possible to determine if the measurements are affected by systematic diurnal variations. On Figure 4 we report comparisons of Figure 3 after removal of the biases; lower noise for the comparisons of Motorolas receivers with Sercels receivers being due, perhaps, to the use of non-standard software.

320 km BASELINE RESULTS

Figures 6 to 13 show precision determinations of GPS common-view time comparisons between clocks at the BIPM and the OB using Motorola and Sercel receivers. The interval covered is about one month. We have chosen to express the precision of a single 13-minute GPS common-view measurement in terms of the root mean square (rms) of the differences between raw and smoothed values. Vondrak smoothing (Ref. 4), which acts as a low-pass filter with cut-off periods ranging of about 1 day, was performed on the raw GPS common-view values. This cut-off period was chosen as representing, approximately, the limit between short time intervals, for which measurement noise is dominant, and longer intervals, for which clock noise prevails. This was determined by computation of modified Allan variance for the time series of Figures 6, 8, 10 and 12.

Atomic clocks at the BIPM and the OB were compared in terms of four time links: Motorola receivers at both sites, Sercel receivers at both sites, a Motorola receiver at the BIPM and a Sercel receiver at the OB, a Sercel receiver at the BIPM and a Motorola receiver at the OB. For each of the four time links we provide smoothing of raw data, and smoothing of data after the removal of the biases. After removal of the biases time links between Sercel receivers have a rms difference of 2,1 ns, and those between Motorola receivers a rms difference of 3,1 ns. This is an excellent result for the Motorola receivers as 'classical' receivers used for GPS common-view links show differences ranging from 1,7 ns rms to 3,5 ns rms, for distances up to 1000 km. For time links with 'crossed' receivers, Figures 10 to 13, we note that the one involving the BIPM Motorola receiver is noisier than the one involving the OB Motorola receiver. This confirms the results of the zero-baseline comparisons which indicate that the BIPM Motorola receiver is slightly noisier.

In Table 1 we compare time links using the Sercel and Motorola receivers. There is a constant shift of 445 ns between the two, certainly due to the use of uncalibrated Sercel and Motorola equipment. When the constant shift is removed, the values obtained agree to within a few

nanoseconds. Some of the remaining discrepancy can be explained by the instability of the OB Motorola receiver when subject to temperature variations. The result, however, remains within the performance of 'classical' time receivers.

CONCLUSIONS

- Tests of the GPS Motorola receiver reported in this paper demonstrate the metrological quality of this device and confirm the results of earlier work (Ref. 5).
- For metrological work it is a major advantage to operate the receiver with an external time intervallometer and a PC computer. This permits better control of the GPS time measurement ensemble.
- Development of software with standard formulae and constants is essential.
- Further investigation of short term and long term behaviour of the Motorola receiver is required. This should focus on the temperature dependence of the antenna and the receiver itself.

REFERENCES

[1] Lewandowski W., Moussay P., Guerin P., Meyer F., Vincent M., 1996, Testing Motorola Oncore GPS Receiver and Temperature Protected Antennas for Time Metrology, Proc. PTTI'96, in press.

[2] Allan D.W. and Thomas C., 1994, Technical Directives for Standardization of GPS Time Receiver Software, Metrologia, **31**, pp. 67-79.

[3] Gifford R.P., Personal communication, Dec. 1996.

[4] Vondrak J., 1969, A Contribution to the Problem of Smoothing Observational Data, Bull. Astron. Inst. Czechoslovakia.

[5] Gifford R.P., Cutler L.S., Kusters J.A., Miranian M., Allan D.W., 1996, Continuous, Multi-Channel Common-View, L1-GPS Time-Comparison Over A 4,000 km Baseline, Proc. Frequency Control Symposium, (in press).

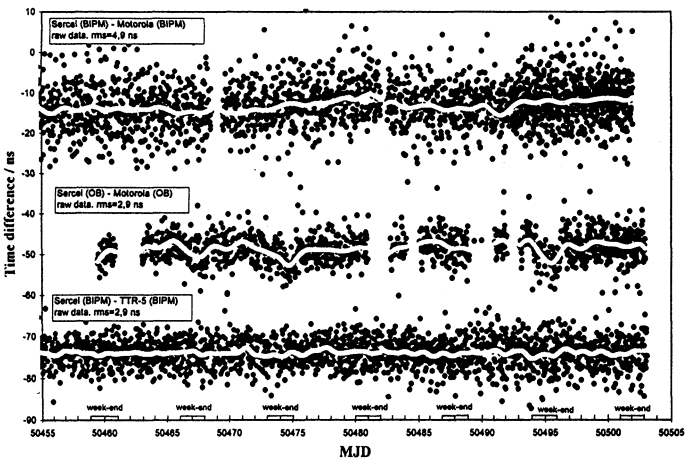


Figure 3. $[BIPM\ HP5071A - GPS\ time]_{SERCEL}$ - $[BIPM\ HP5071A - GPS\ time]_{MOTOROLA}$, $[OB\ HP5071A - GPS\ time]_{SERCEL}$ - $[OB\ HP5071A - GPS\ time]_{MOTOROLA}$, $[BIPM\ HP5071A - GPS\ time]_{SERCEL}$ - $[BIPM\ HP5071A - GPS\ time]_{TTR5}$, for individual 13-minute tracks and corresponding standard deviation.

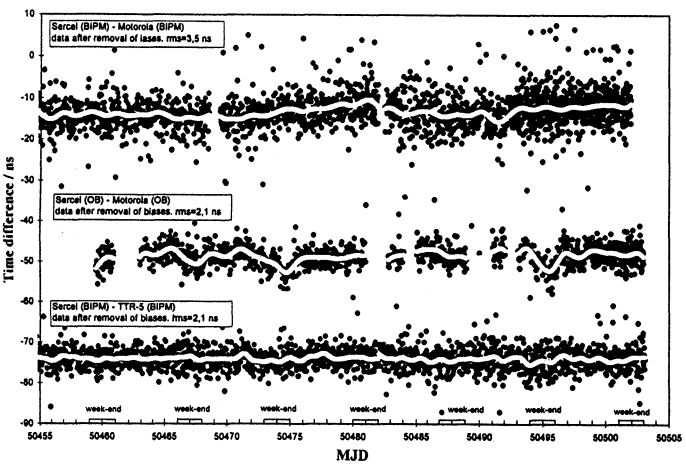


Figure 4. $[BIPM\ HP5071A - GPS\ time]_{SERCEL}$ - $[BIPM\ HP5071A - GPS\ time]_{MOTOROLA}$, $[OB\ HP5071A - GPS\ time]_{SERCEL}$ - $[OB\ HP5071A - GPS\ time]_{MOTOROLA}$, $[BIPM\ HP5071A - GPS\ time]_{SERCEL}$ - $[BIPM\ HP5071A - GPS\ time]_{TTR5}$, for individual 13-minute tracks and corresponding standard deviation, after removal of the biases.

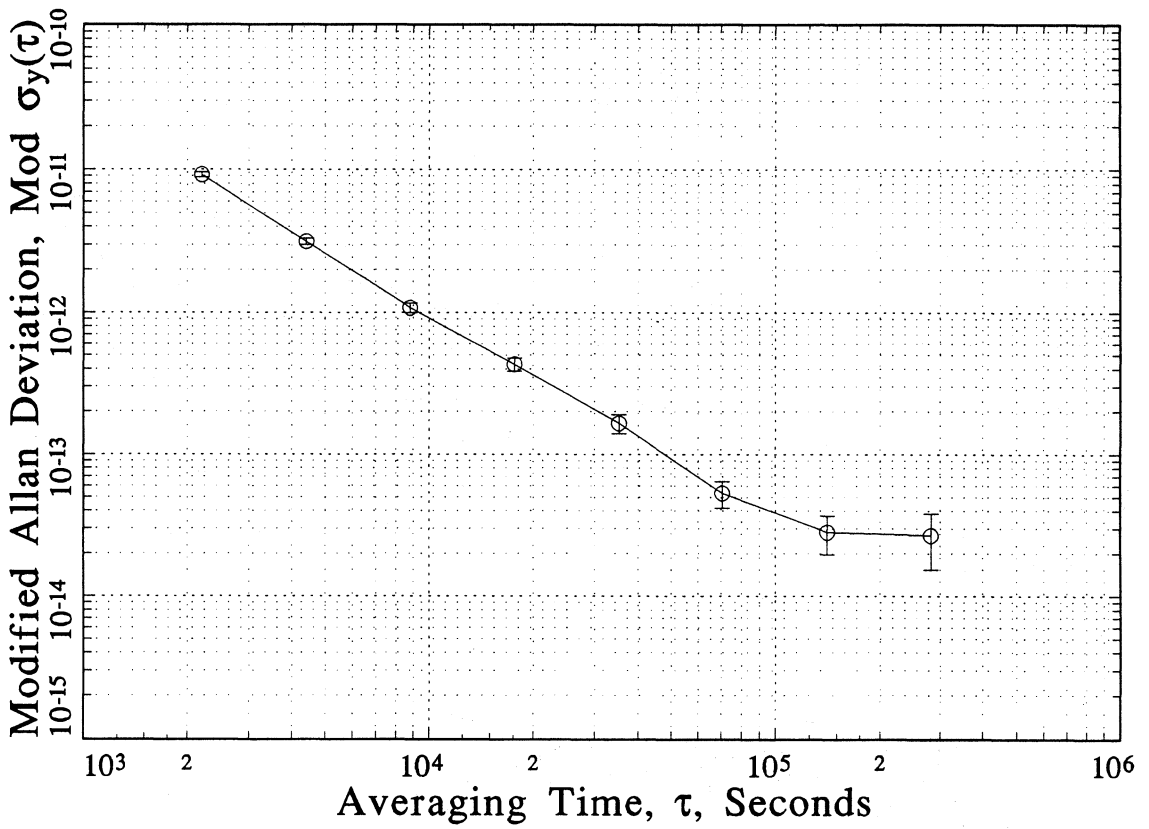


Figure 5. Square root of the modified Allan variance of the time series [*BIPM HP5071A - GPS time*]$_{\text{MOTOROLA}}$ [*OB HP5071A*]$_{\text{MOTOROLA}}$ reported on Figure 5.

Table 1. Comparison of [*BIPM HP5071A - OB HP5071A*]$_{\text{SERCEL}}$ - [*BIPM HP5071A - OB HP5071A*]$_{\text{MOTOROLA}}$.

MJD	(clock BIPM - clock OB)$_{\text{Sercel}}$ /ns	(clock BIPM - clock OB)$_{\text{Motorola}}$ - 445 /ns	(clock BIPM - clock OB)$_{\text{Sercel}}$ - ((clock BIPM - clock OB)$_{\text{Motorola}}$ - 445) /ns
50460	2615.5	2614.2	1.3
50463	2621.4	2621.6	-0.2
50464	2625.2	2625.8	-0.6
50465	2633.1	2634.7	-1.6
50466	2635.1	2637.1	-2.0
50467	2637.1	2637.1	0.0
50468	2640.5	2640.5	0.0
50469	2637.7	2638.9	-1.2
50470	2634.8	2635.8	-1.0
50471	2631.1	2634.3	-3.2
50472	2633.3	2634.1	-0.8
50473	2632.3	2632.7	-0.4
50474	2638.9	2635.6	3.3

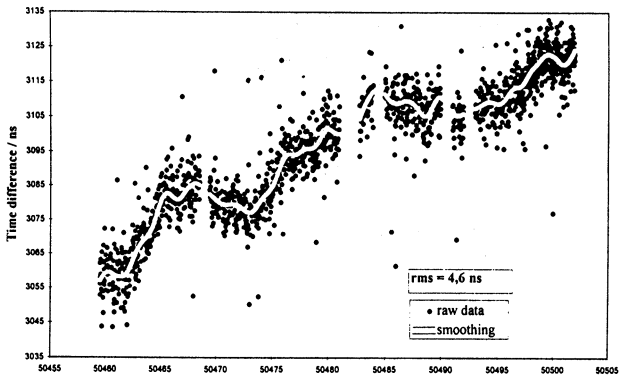


Figure 6. [*BIPM HP5071A-GPS time*]$_{Mot}$-[OB *HP5071A-GPS time*]$_{MOT}$, for individual 13-minute tracks and corresponding standard deviation.

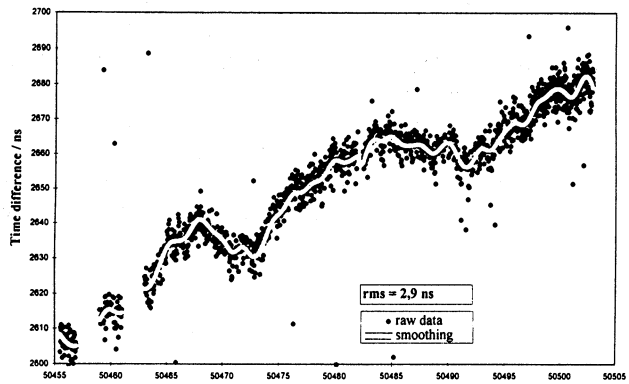


Figure 8. [*BIPM HP5071A-GPS time*]$_{SER}$-[OB *HP5071A-GPS time*]$_{SER}$, for individual 13-minute tracks and corresponding standard deviation.

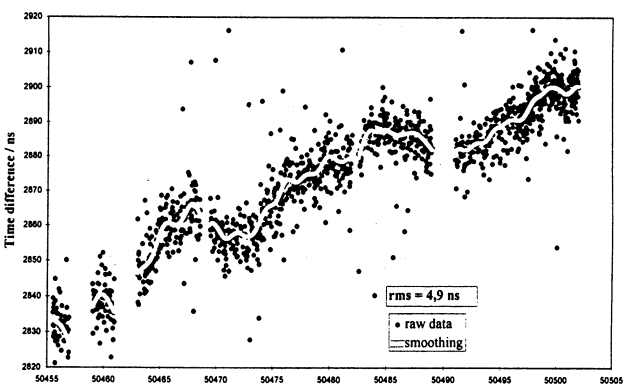


Figure 10. [*BIPM HP5071A-GPS time*]$_{MOT}$-[OB *HP5071A-GPS time*]$_{SER}$, for individual 13-minute tracks and corresponding standard deviation.

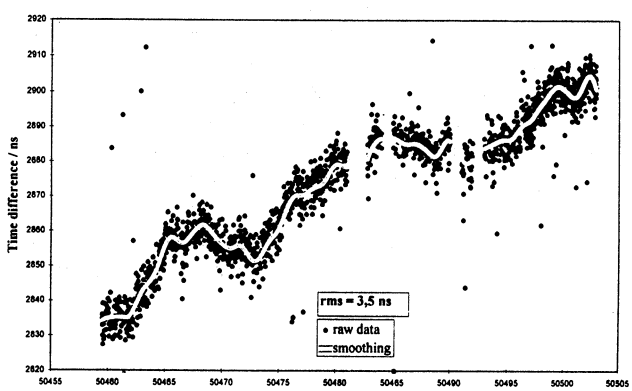


Figure 12. [*BIPM HP5071A-GPS time*]$_{SER}$-[OB *HP5071A-GPS time*]$_{MOT}$, for individual 13-minute tracks and corresponding standard deviation.

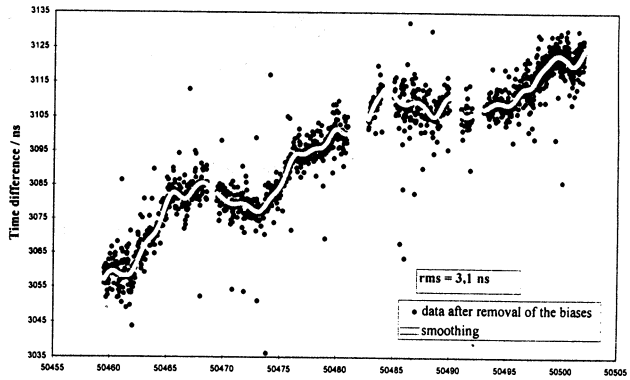


Figure 7. [*BIPM HP5071A-GPS time*]$_{MOT}$-[OB *HP5071A-GPS time*]$_{MOT}$, for individual 13-minute tracks and corresponding standard deviation, after removal of the biases.

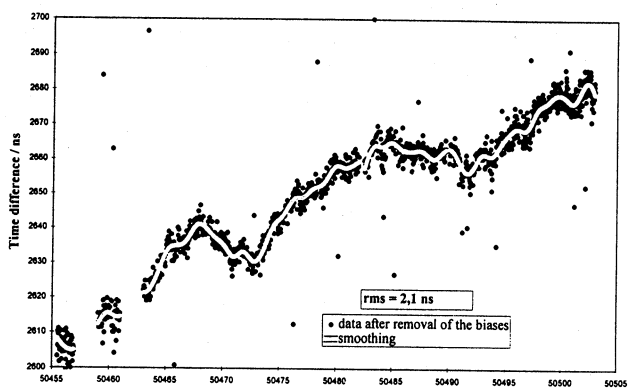


Figure 9. [*BIPM HP5071A-GPS time*]$_{SER}$-[OB *HP5071A-GPS time*]$_{SER}$, for individual 13-minute tracks and corresponding standard deviation, after removal of the biases.

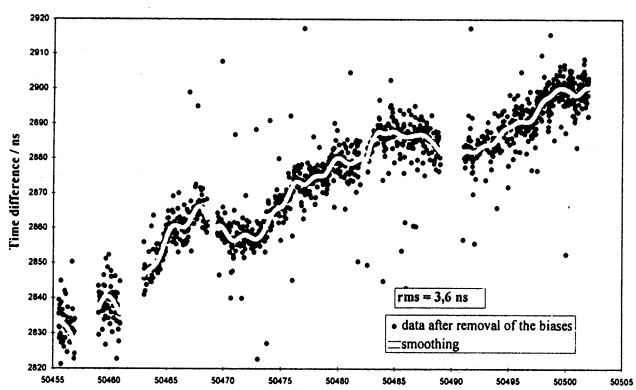


Figure 11. [*BIPM HP5071A-GPS time*]$_{MOT}$-[OB *HP5071A-GPS time*]$_{SER}$, for individual 13-minute tracks and corresponding standard deviation, after removal of the biases.

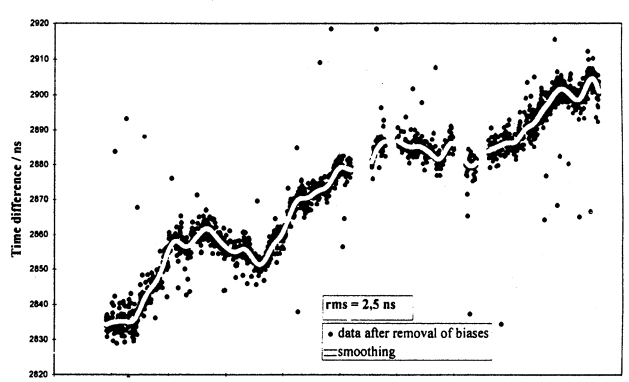


Figure 13. [*BIPM HP5071A-GPS time*]$_{SER}$-[OB *HP5071A-GPS time*]$_{MOT}$, for individual 13-minute tracks and corresponding standard deviation, after removal of the biases.

TEMPERATURE - PROTECTED ANTENNAS FOR SATELLITE TIME TRANSFER RECEIVERS

W. Lewandowski*, P. Moussay*, J. Danaher**, R. Gerlach**, E. LeVasseur**

*Bureau International des Poids et Mesures, Sèvres, France
**3S Navigation, Laguna Hills, California, USA

Abstract

The dependence of GPS time equipment on external temperature can be removed by enclosing the antennas in stabilized-temperature ovens. A detailed description of such ovens is provided. The phenomenon is illustrated by examples of temperature-correlated results. The improvement brought about by use of temperature-controlled antennas is described.

INTRODUCTION

During the last decade the performance of GPS common-view time transfer has improved by one order of magnitude through the use of high-accuracy ground-antenna coordinates, post-processed precise ephemerides, and double-frequency ionospheric measurements. In good cases the uncertainty of this time transfer can approach 2 ns, but further progress is limited by the performance of receiver hardware. It is now well documented, and generally admitted, that GPS time equipment is sensitive to external temperature (Ref. 1). The variation is typically of about 0,2 ns/°C and can approach 2 ns/°C for some types of receiver. Even for the lower value, the delay change can be the dominant contribution to the noise of time transfer by common-view for periods of several days over short baselines of several hundred kilometres. The higher value results in a diurnal effect of about 20 ns and a seasonal effect of several tens of nanoseconds. This makes it the dominant contributor to the noise of common-view time transfer even for intercontinental baselines. We illustrate this phenomenon for GPS equipment with results recorded during a recent calibration of a GPS reference receiver at the USNO covering a period of 11 months, Figure 1. For GLONASS, we report a comparison of two GLONASS time receivers recorded at the BIPM over a period of 10 months, Figure 3. It can be seen that the GPS and GLONASS receivers have similar behaviour, both over periods of several days and over the full period of the comparison. Short periods are characterized by a correlation with external temperature, Figures 2 and 4. A rough estimate of the correlation coefficient is 0,2 ns/°C for both GPS and GLONASS. No seasonal effect is evident.

The sensitivity to external temperature, suggests an effect linked to those parts of the time equipment located in the open-air, that is to the antenna and its cable. The receiver itself is usually located in an air-conditioned room. For several years different hypotheses were considered to explain the temperature dependence of timing equipment. All linked the problem to the electronics of the antenna, but none were verified. Experiments showed that the changes were not due to the changes in the antenna cable (Ref. 2), however the length and material of the cables are important and must be considered.

TEMPERATURE-PROTECTED ANTENNAS

As no practical way was found to resolve the problem electronically, another approach was suggested (Ref. 3): the antenna should be protected by an oven with a stabilized temperature. The primary objective of the antenna temperature stabilization process is to maintain the critical components at some constant temperature. The exact temperature is not critical, but the lower the better since higher temperatures have a negative impact on both the preamplifier Noise Figure and the overall reliability of the entire assembly. This desire to choose a low set point temperature for the amplifier is complicated by the need to use a set point that is higher than the highest anticipated temperature from the sun plus self heating. The sun's radiation can easily raise the temperature by 17°C, so that operation at 50°C would require a set point well in excess of 67°C.

Self heating is a more complex variable being dependent on both the internal heat dissipation in the preamplifier electronics and the amount of thermal insulation between the preamplifier assembly and the outside world. Extensive use of high quality insulating materials would provide excellent insulation from outside thermal fluctuations, but would result in a high temperature rise due to self-heating from the internal preamplifier. Minimal insulation would minimize self-heating effects but would make the unit susceptible to rapid ambient temperature changes due to such elements as shifting

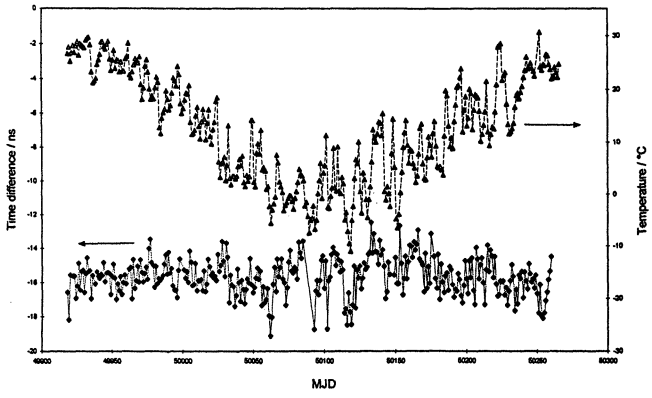


Figure 1. Daily averages of [*UTC(USNO) - GPS time*]$_{TTR6}$ - [*UTC(USNO) - GPS time*]$_{Stel502}$, and daily average temperature at USNO for a period of 11 months.

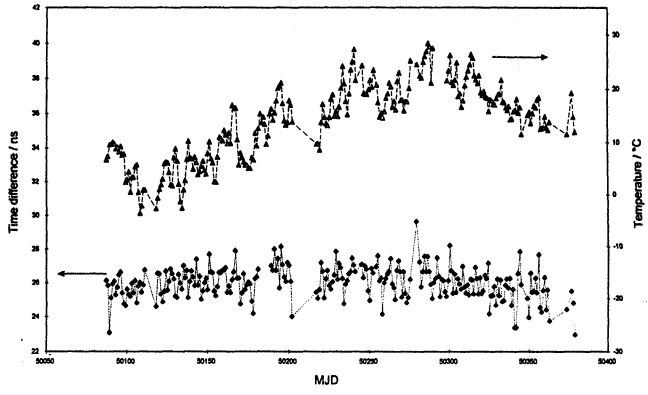


Figure 3. Daily averages of [*BIPM clock - GLONASS time*]$_{R\text{-}100/10}$ - [*BIPM clock - GLONASS time*]$_{R\text{-}100/30}$, and daily average temperature at BIPM for a period of 10 months.

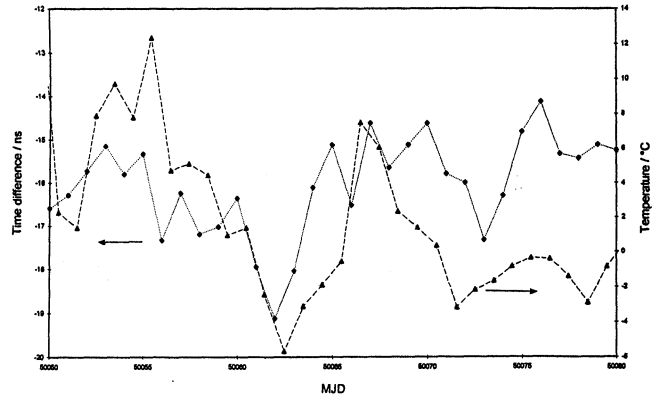


Figure 2. Daily averages of [*UTC(USNO) - GPS time*]$_{TTR6}$ - [*UTC(USNO) - GPS time*]$_{Stel502}$, and daily average temperature at USNO for a period of about 1 month.

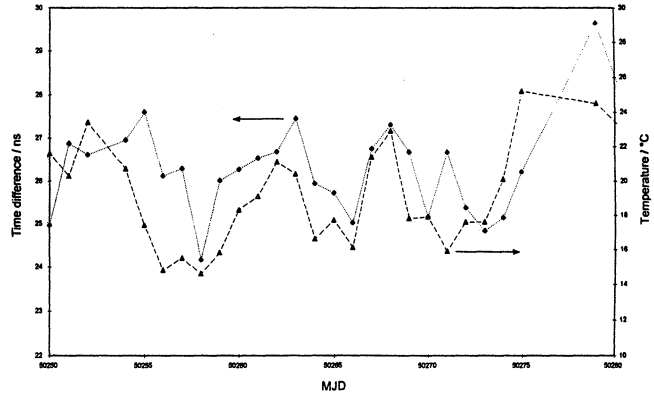


Figure 4. Daily averages of [*BIPM clock - GLONASS time*]$_{R\text{-}100/10}$ - [*BIPM clock - GLONASS time*]$_{R\text{-}100/30}$, and daily average temperature at BIPM for a period of about 1 month.

shadows, wind gusts, thunder showers ... Minimal insulation would call for a high-power electric heating system to handle low ambient temperatures during the winter months.

At the BIPM two ovens were built, both being simple chambers with stabilized temperature. Detailed descriptions of these ovens are in Figures 5 and 6. The temperature of the oven used at the BIPM was set at 38°C. This is the highest temperature recorded at Sèvres, which implies that only heating is required: cooling systems are much more complicated to build in a laboratory.

The 3S Temperature Stabilized Antenna model TSA-100, shown in Figure 7, solves the temperature problem with a dual chamber design.

The preamplifier assembly is maintained at the relatively mild temperature of 40,5°C. This assembly is surrounded by minimal amounts of insulation and enclosed by a outer chamber which is maintained at a maximum temperature of 27°C. This outer chamber is heavily insulated so that it is thermally isolated from the external ambient air. When ambient temperatures rise above approximately 21°C, this chamber is cooled by use of a Thermo-Electric Cooler (TEC) which pumps the heat to an exterior heat sink without the use of moving parts. As ambient temperatures drop below 21°C the cooling is not required, the TEC is turned off and the chamber temperature is allowed to drop along with the ambient. As ambient temperatures drop further, various thermostats engage heating systems to prevent the chamber temperature from dropping below approximately 7°C.

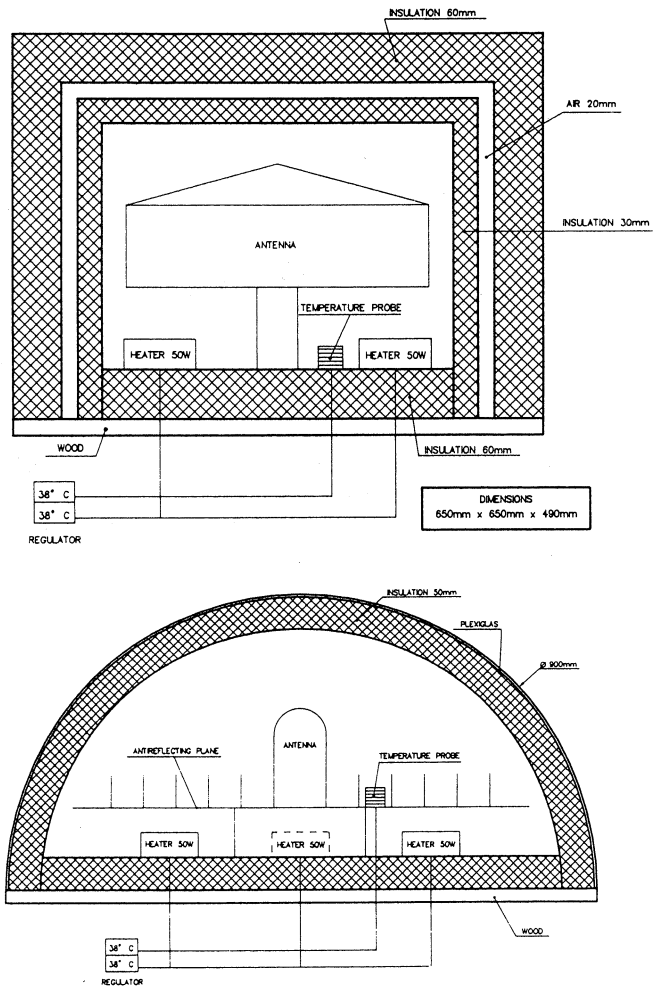


Figure 5. Two ovens built at the BIPM.

The low end temperature control of the outer chamber does not need to be precise. The first level of temperature control occurs as the ambient air temperature approaches 10°C at which point a thermostatic switch applies a reverse current to the TEC which reverses the direction of heat flow and pumps heat into the outer chamber from the outside heat sink. As the outside ambient temperature drops to the point at which reverse heat point action can no longer maintain the outer chamber above 10°C, a second thermostatic switch engages heaters attached to the walls of the outer chamber. These heaters, with the aid of the heavy insulating layer, can easily maintain the outer chamber temperature at this level for external ambient temperatures down to -23°C.

The net result is an inner chamber with embedded heaters that operates with minimal insulation at a temperature approximately 14°C above that of the outer chamber. The outer chamber is heavily insulated: it is cooled at higher ambient temperatures and heated at lower ones. Control for both chambers is remotely maintained by Proportional-Integral-Derivative (PID) controllers located in the indoor/laboratory Temperature Control Unit (TCU).

The temperatures of the inner chamber, containing the preamplifier, and of the outer chamber which surrounds it, are both transmitted to the PID controllers in the TCU down a remote control cable via proportional currents derived from RTD temperature sensors. The operating or process temperatures appear on the PID controllers with red LED displays. The set points for these controllers appear in green LEDs. The set points are set at the factory but can be varied within certain bounds to perform various temperature related experiments from the convenience of the laboratory.

RESULTS

Initial observations show that temperature stabilization of the antenna assembly reduces or even eliminates the diurnal delay variation. It is thought that the observed stabilization results from control of the temperature of the filters and amplifiers rather than of the antenna element itself. To demonstrate this two GPS time receivers, AOA TTR5 and AOA TTR6, connected to the same clock were compared in common-view using the standard BIPM international schedule. The receiver antennas are separated by a few metres and their differential coordinates are known with an uncertainty of several centimetres. Smoothing with a cut-off period of about 1 day was applied to raw data in order to identify receiver delay variations caused by daily temperature changes. In Figure 8 we show a comparison of the two receivers covering a period of about 1 month when the antennas were not protected. There is a clear correlation with external temperature. Figure 9 shows a similar comparison for 1 month when the AOA TTR6 antenna was protected by a BIPM oven. The temperature dependence is greatly reduced. The AOA TTR5 antenna was not protected by an oven, but this type of antenna is known to have very small temperature dependence.

The initial test results for the 3S Navigation TSA-100 antenna are shown in Figures 10 and 11. The test setup consists of two separate R-100/40 GPS/GLONASS receivers, using TSA-100 antennas separated by 2 metres. The observation site was at Laguna Hills, California. The two receivers are attached to a common frequency reference and clock, but are otherwise operated independently, in common view mode. The raw differences in 13-minute L1 wide-band GLONASS common tracks over two days are shown in Figure 10. This is the difference in time reference measurements between the two receivers, with a constant time bias subtracted from the difference. Almost all observations fall within +/-5 ns of the average and the diurnal variation, seen with non-temperature controlled antennas is not present.

The transmission frequency of the L1 GLONASS signals is 1602 MHz + $9k/16$ MHz where k is 1 to 24. It is expected that there will be a constant calibration delay which depends upon the GLONASS frequency. When these frequency dependent constants are removed, the

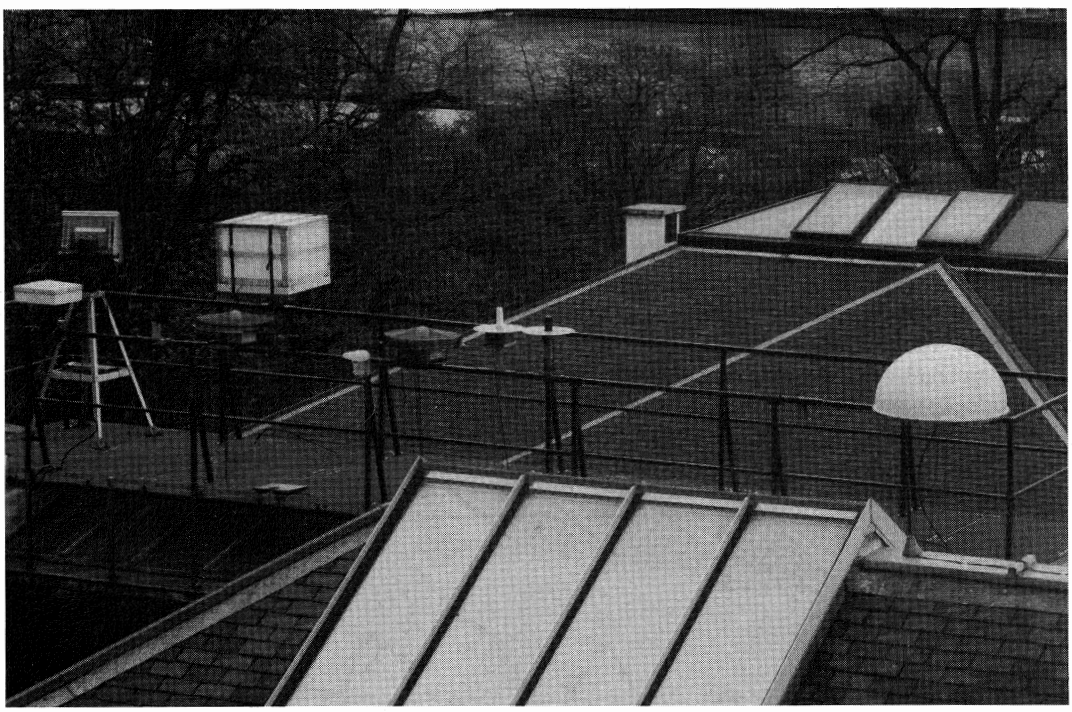


Figure 6. GPS and GLONASS antennas at the BIPM. Two GPS antennas are covered by ovens.

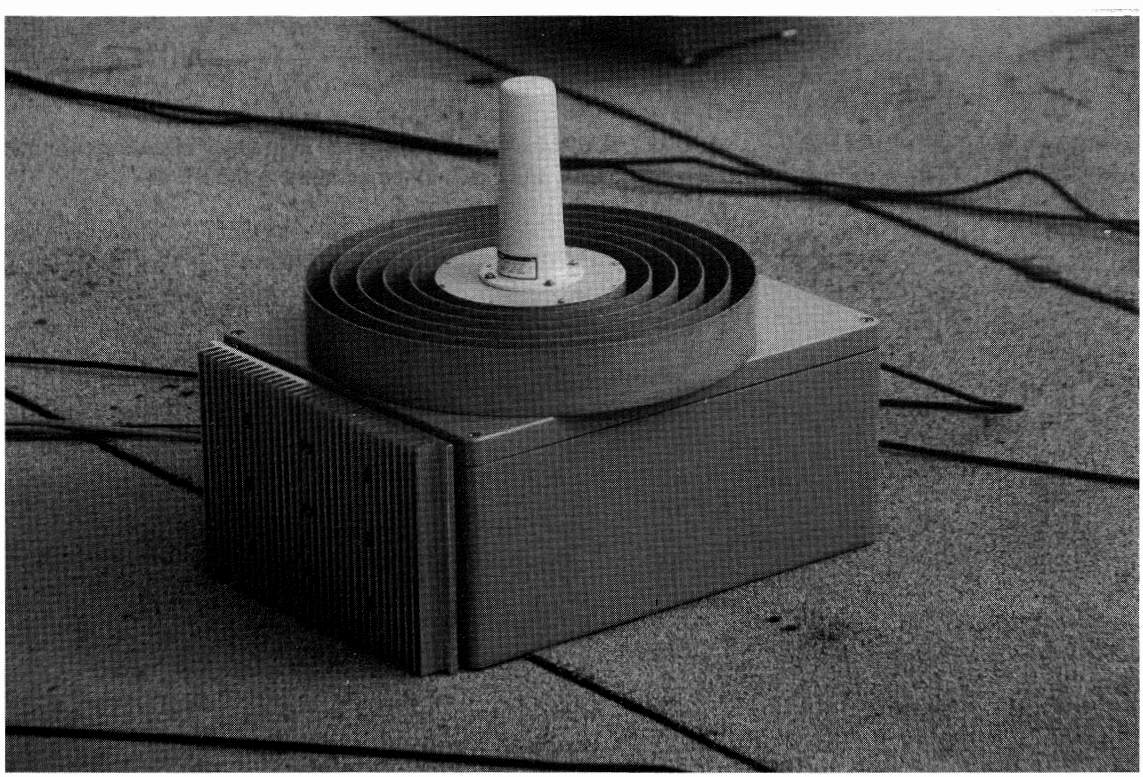


Figure 7. The TSA-100, the 3S Temperature Stabilized GPS/GLONASS/WAAS L1+ L2 Antenna System.

two day common view differences are as shown in Figure 11. This shows that the range of 13-minute common view results is now about +/- 2 ns and that any remaining diurnal delay variation must be in the sub-nanosecond level.

CONCLUSION

- This study shows that GPS, GLONASS and GPS/GLONASS antenna electronics and outdoor in-line amplifiers should be temperature stabilized when they are used for precision time applications.
- The use of temperature-stabilized enclosures should improve not only common-view time transfer and time dissemination, but also frequency comparison by phase measurements.
- The development of a built-in time calibration system for time receivers is a challenge for the timing community. Such a solution is that best adapted to resolve current difficulties with the delay stability of GPS and GLONASS time equipment.

REFERENCES

[1] Lewandowski W., and Tourde R., 1990, Sensitivity to the external temperature of some GPS time receivers , Proc. 22nd PTTI Meeting, pp.307-316.

[2] Freon G., Personal communication, 1990.

[3] Lewandowski W., Moussay P., Guerin P., Meyer F., Vincent M., 1996, Testing Motorola Oncore GPS Receiver and Temperature Protected Antennas for Time Metrology, Proc. PTTI'96, in press.

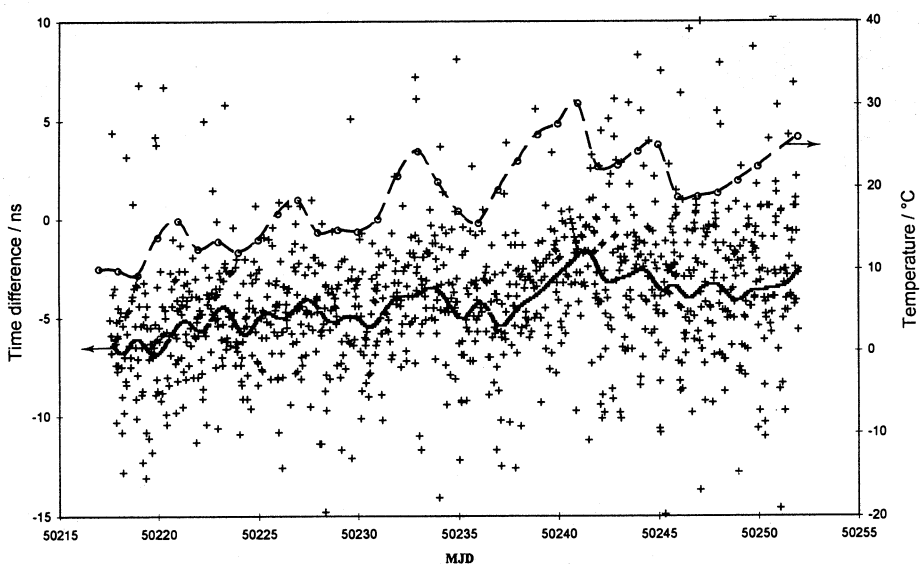


Figure 8. [*BIPM HP5071A - GPS time*]$_{TTR5}$ - [*BIPM HP5071A - GPS time*]$_{TTR6}$ for individual 13-minute common-view tracks, with TTR5 and TTR6 antennae no-protected by ovens, and external temperature at the BIPM.

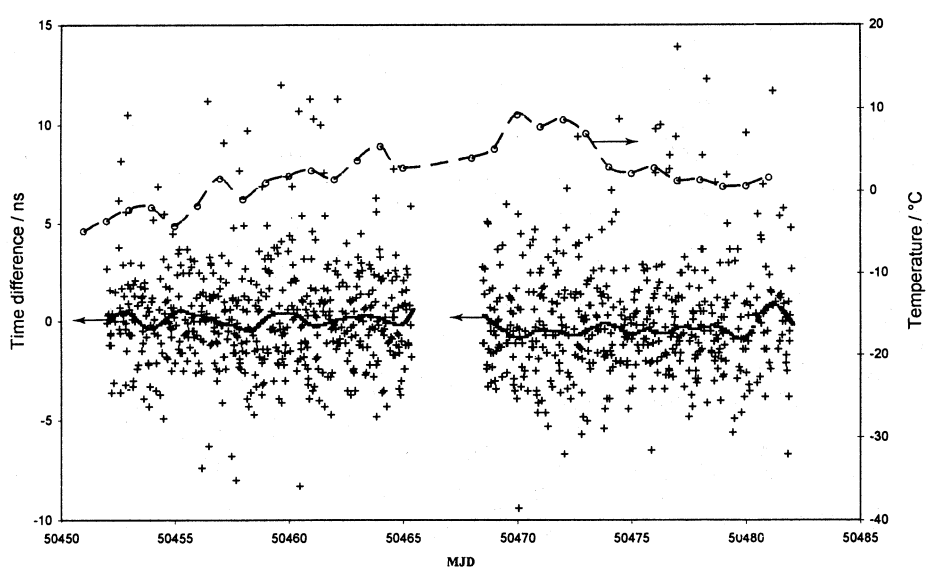


Figure 9. [*BIPM HP5071A - GPS time*]$_{TTR5}$ - [*BIPM HP5071A - GPS time*]$_{TTR6}$ for individual 13-minute common-view tracks, with TTR6 antenna protected by an oven, and external temperature at the BIPM.

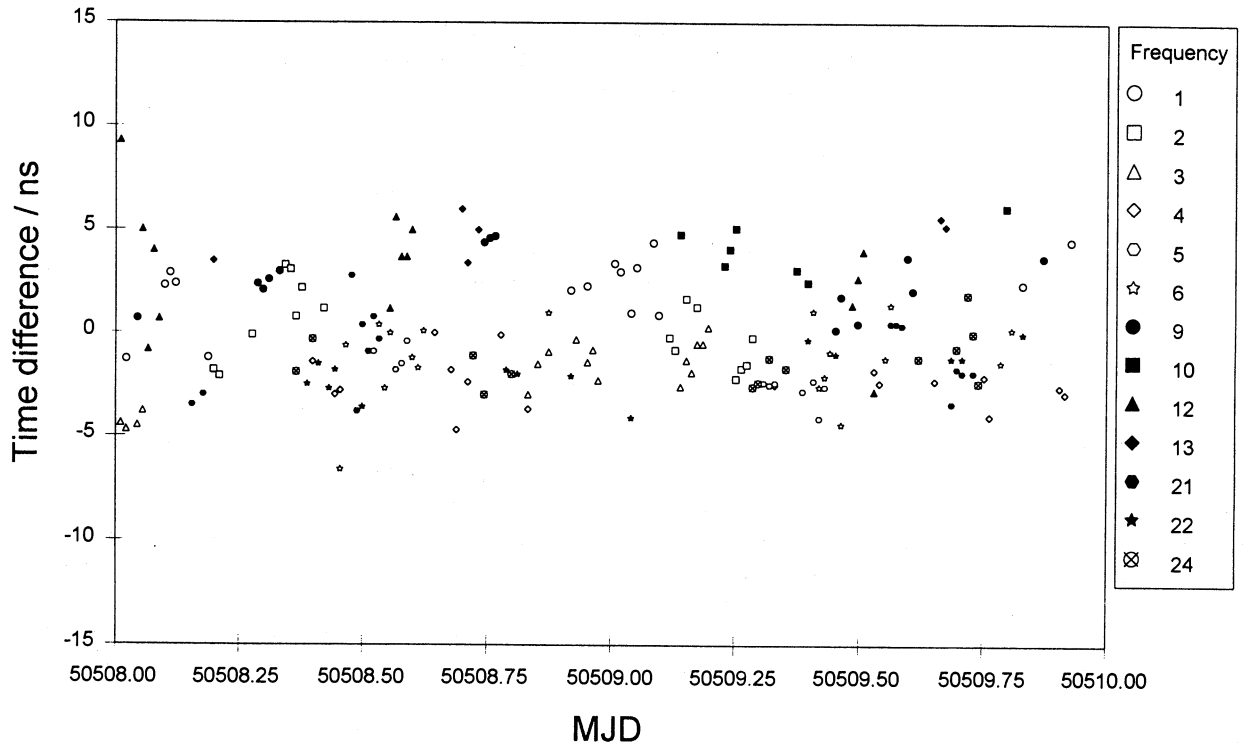


Figure 10. [3S *Clock - GLONASS time*]$_{\text{R-100/40 No 1}}$ - [3S *Clock - GLONASS time*]$_{\text{R-100/40 No 2}}$ for individual 13-minute common-view tracks, with TSA-100 antennas.

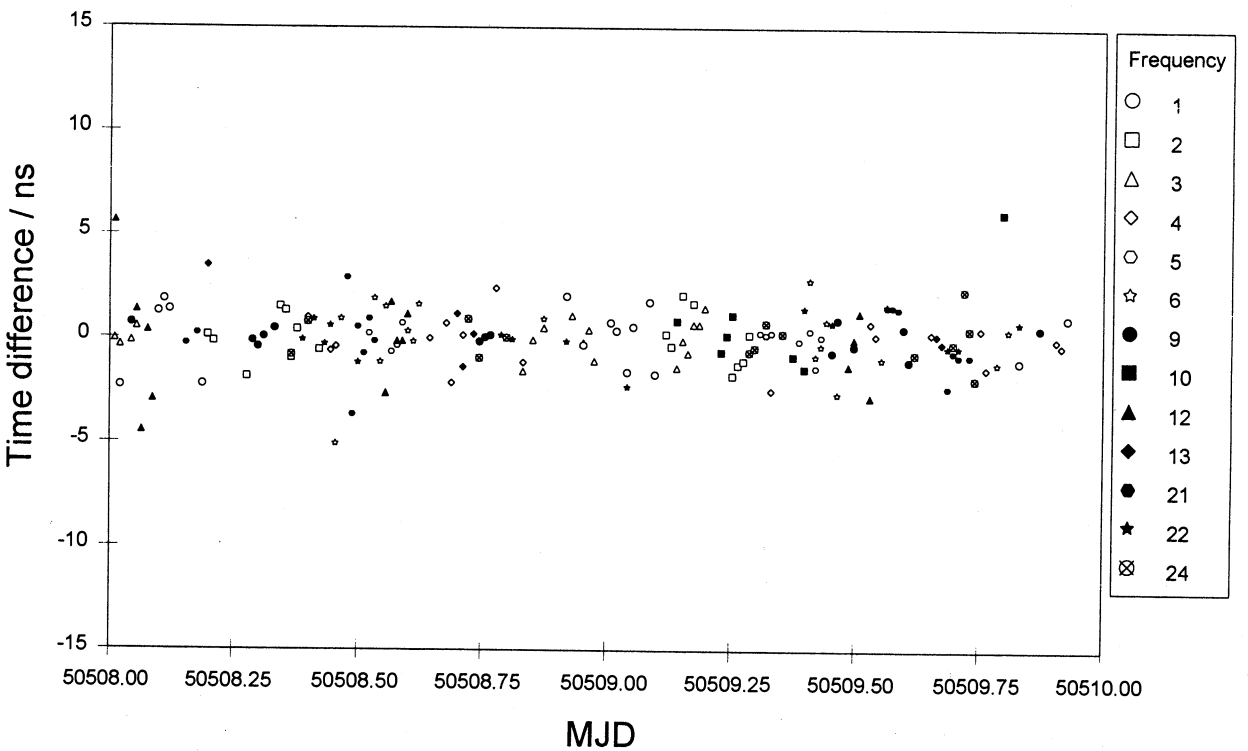


Figure 11. [3S *Clock - GLONASS time*]$_{\text{R-100/40 No 1}}$ - [3S *Clock - GLONASS time*]$_{\text{R-100/40 No 2}}$ for individual 13-minute common-view tracks, with TSA-100 antennas after bias removal due to GLONASS frequencies.

GPS TIME TRANSFER USING GEODETIC RECEIVERS: MIDDLE-TERM STABILITY AND TEMPERATURE DEPENDENCE OF THE SIGNAL DELAYS

F. Overney, Th. Schildknecht, G. Beutler
Astronomical Institue, University of Berne

L. Prost, U. Feller
Swiss Federal Office of Metrology

1. ABSTRACT

It has already been shown that time transfer with a theoretical accuracy of some 10 ps can be achieved by using the tools developed in the field of geodesy (geodetic GPS receivers and data processing in interferometric mode) [5].
In order to reach the required sub-nanosecond accuracy, it is necessary to know the delay in the GPS signal path from antenna to the receiver and the associated temperature dependence.
After a short description of the time transfer system, we will give a synthesis of the results from earlier experiments and a new experiment yielding the temperature and humidity dependence of the signal delay. In this study, the dependence of the signal path from ambient conditions is determined for each part of the signal path: antenna, cable, electronics and GPS receiver.
In addition, the results of stability measurements for a short base line (6 m) over a period of 12 days will be presented.

Keywords: time transfer, GPS receiver, temperature dependence, GeTT.

2. INTRODUCTION

In parallel with the increasing quality of atomic frequency standards it is necessary to improve the methods to compare them.
One possibility is the use of the Global Positioning System (GPS) for transferring time between two remote clocks.
The classical approach is the so-called "Common-View" method (CV) [1]. However, this method does not exploit the full capacity of GPS signal processing and thus the accuracy in the time comparison is limited to few nanoseconds.
In order to improve the accuracy of GPS time transfer to below the sub-nanosecond level it was suggested to use a geodetic approach [2,3].
We developed a new Geodetic Time Transfer (GeTT) method that uses geodetic GPS receivers (which can use all GPS observables) and a interferometric data processing method adapted from geodetic technique.
In order to reach the required sub-nanosecond accuracy, it is not only necessary to know the delay in the GPS signal path from the antenna to the receiver but also the effect of variation of local parameters like the temperature.
The purpose of this paper is to make a synthesis of the results from earlier experiments and new experiments yielding the temperature dependence of the signal delay through the antenna, antenna cable and the GPS receiver.
We also give some results about the stability and the accuracy of GeTT.

3. DESCRIPTION OF STATIONS

Two prototype GeTT terminals were built in 19" mobile racks [4]. Figure 1 shows the schematic of one terminal. We can discern two types of components: Firstly, the components involved in the GPS signal acquisition like the antenna, the antenna cable and the receiver box that contains a modified geodetic GPS receiver and two electronic boards. The latter contain the PPS signal distribution (PPS fan out) and a 5 to 20 MHz multiplier. Secondly we find auxiliary components used for delay measurement and data transfer. These components are a Time Interval Counter (SR620) and a PC.
The signal/data flow in Figure 1 begins on the top and goes down to the bottom. The receiver box has three input connectors and three output connectors. One input for the antenna cable that receives GPS signals from the antenna and two clock synchronisation inputs (5 MHz and 1 PPS) used to synchronise the GPS receiver with the station reference clock. The outputs are firstly used to download data from the GPS receiver to the PC (through an RS232 link) and secondly to measure the time interval between the PPS from the fan out board and from the GPS receiver. This measurement gives some in-

formation about the PPS delay through the GPS receiver.

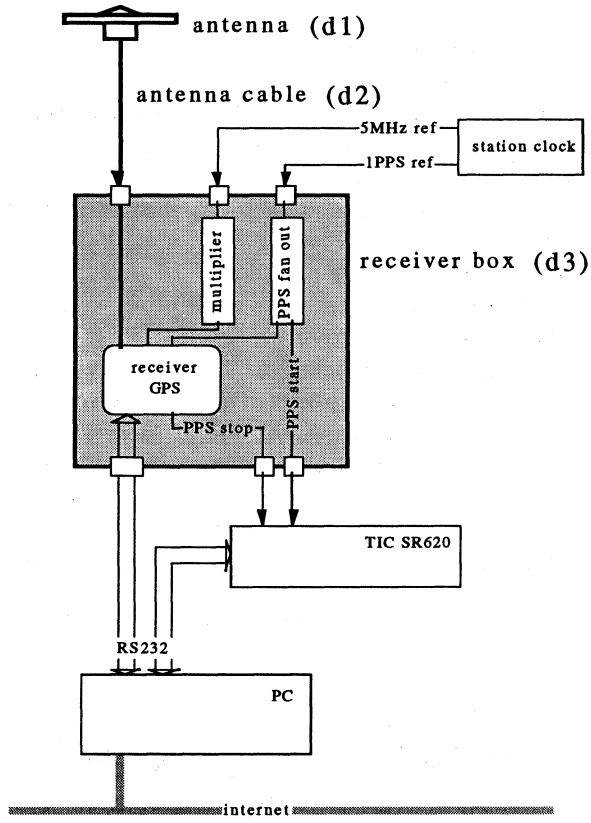


Figure 1: Schematic of a time transfer terminal

The data downloaded from the GPS receiver are sent to a central computer (once per day) where they are processed (together with data from the second terminal) to obtain the time (and frequency) differences between the station clocks driving the terminals.
Each part of the local signal path introduces a delay (d1 to d3) which must be considered in the final reduction of the time transfer. In fact, it is the difference between these delays from each terminal which is critical.
In the next section we study the variation of these delays with the temperature.

4. TEMPERATURE DEPENDENCE

The dominant parameter that influences local delays is the temperature. We tried to evaluate this dependence and its contribution to the final error budget.

4.1 Antenna

We have not yet conducted a new experiment to measure the temperature dependence of the delay through the antenna (and its amplifier). However, some measurements made by the manufacturer (Ashtech, Sunnyvale, CA) show a decrease of the delay with increasing temperature between -20°C and 40°C. The magnitude of this dependence is less than -0.04 ns/°C.
Another estimate of this dependence was done for an Allen Osborne antenna [6] and gives a magnitude of 0.1 ns/°C. Despite the difference between these two estimates, we can expect that our two Ashtech antennas show the same temperature dependence. As we will indeed see in section 5.1, no outdoor temperature correlation has been observed when two antennas were exposed to the same temperature variation.
For a longer base line the two antennas will not be exposed to the same temperature variations and we will have to know exactly the thermal coefficients. So, more investigations about the antenna temperature dependence will be necessary.

4.2 Antenna Cable

To determine the antenna cable (type RG213/U) temperature dependence we made a GeTT experiment between two terminals (TT1 and TT2) placed side by side and driven by the same station clock. We also used the same antenna for both terminals but in one case a 20 m long part of the antenna cable was placed in a thermal box that can be stabilised in temperature T and relative humidity rH.
We had also measured the temperature dependence of RG58 and RG316 cables but for those we measured the transmission delay of 5 MHz square pulses.

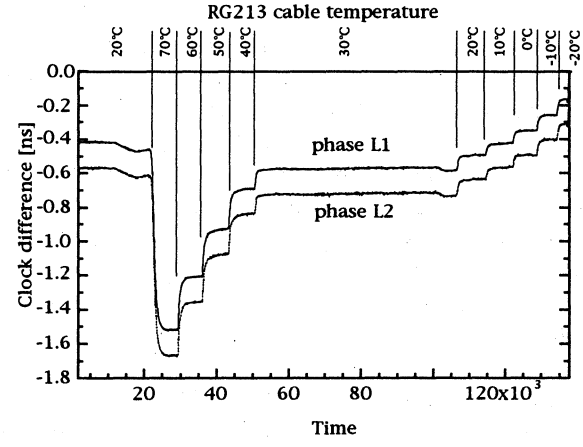


Figure 2: Clock difference computed from L1 and L2 carrier phase measurements by GeTT (with 1 antenna and 1 station clock for both terminals)

Figure 2 shows the variation of the clock difference computed from the L1 and L2 carrier phases measurements by the GeTT method.

There is, of course, an arbitrary offset due to the initial phase ambiguity.
After each temperature step, the clock differences (computed from L1 and L2 measurements) are settled to a new value. Both frequencies show the same temperature dependence. The stabilisation is an exponential process and we can calculate a time constant τ that is about 10 minutes. The delay through the cable decreases if the temperature increase. To explain this decrease of the delay we must not only take into account the lengthening of the cable (that increases the delay), but also the variation of the dielectric constant ε_r of the cable (PE for RG213).

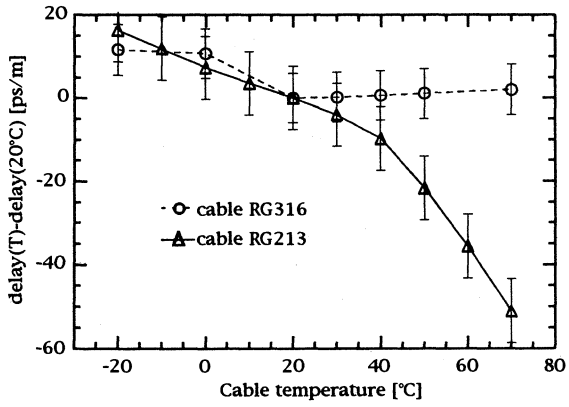


Figure 3: Variation of the signal transmission delay through a cable (RG213 and RG316) with respect to the temperature

Figure 3 shows the variation of the delay with respect to the temperature for a RG213 cable (measured by GeTT) and for a RG316 cable (transmission delay of 5 MHz pulse).
Again, we must consider the effect of the dielectric constant to explain the difference between these two dependencies. The dielectric used in the RG213 cable is polyethylene (PE) whereas for the RG316 cable is polytetrafluorethylene (PTFE). For a RG58 cable that uses PE dielectric material, we obtain the same temperature dependence as for RG213.
For the antenna cable (RG213) that is exposed to the largest temperature variation, we obtain a linear temperature dependence between -20°C and 40°C with a slope of -0.42 ps/°C/m. Above 40°C, the delay decreases more strongly (-1.38 ps/°C/m).
For the RG316 cable, we have no temperature dependence for temperatures above 20°C and below 0°C but there is a step of about 11 ps/m between 0°C and 20°C.

For RG213 cable we have also measured the dependence of the delay with respect to the relative humidity rH. We have put a 20 m cable in water. No significant variation of the delay was measured for 30 days.

4.3 Receiver box

The receiver box is composed of two electronic boards (frequency multiplier and 1 PPS fan out] and a GPS receiver.
The temperature dependence of the electronic boards is linear between 20°C and 50°C. The temperature coefficients are 26 ps/°C and 14 ps/°C for the frequency multiplier and the 1 PPS fan out, respectively .
To measure the temperature dependence of the GPS receiver, we made a GeTT between two terminals with the same set up as described in section 4.2 but with one GPS receiver placed in a thermal box.
Figure 4 shows the variation of the clock difference, computed for the code P1 and the phase L1, when the temperature of the GPS receiver (receiver #2) is increased from 20°C to 50°C.

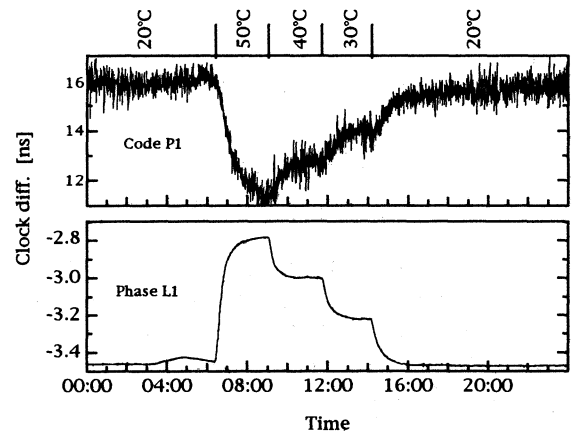


Figure 4: Clock difference computed by GeTT for code P1 and phase L1

After each temperature step, the clock difference stabilises at a new value by an exponential process with a time constant of about 15 min. The clock difference computed with the code decreases if the temperature increases while it is the opposite for the phase. The code P2 and the phase L2 show the same dependence.
It was interesting to know if the second GPS receiver of same type (Ashtech Z-XII modified) would behave similarly. Therefore we repeated this measurement with the GPS receiver #1 in the thermal box. Results for both GPS receivers computed for both codes (P1 and P2) and both phases (L1 and L2) are given in Figure 5.

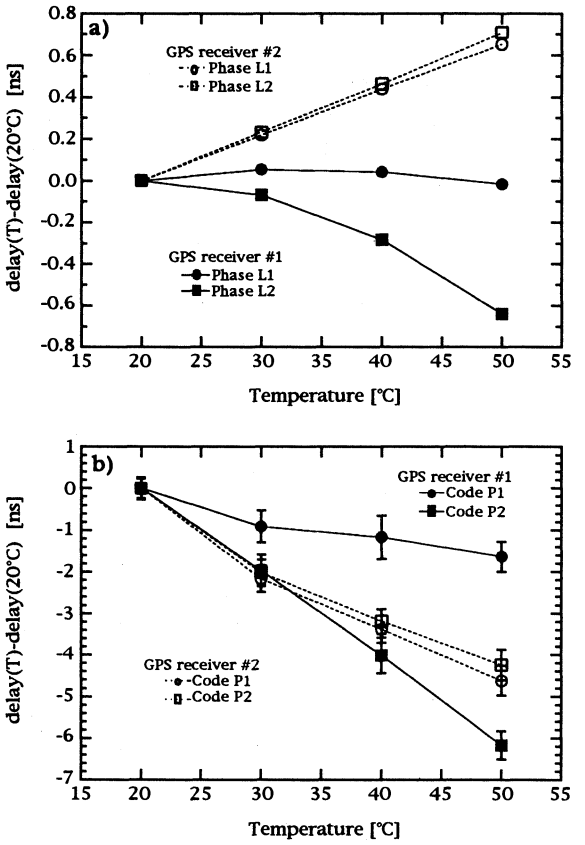


Figure 5: Temperature dependence of two GPS receivers (Ashtech Z-XII modified): a) for phase L1 and L2, b) for code P1 and P2

For receiver #2 the dependence is nearly linear for either code or phase. We obtain a thermal coefficient of about 0.022 ns/°C for the phase and -0.14 ns/°C for the code. Furthermore these coefficients are nearly frequency independent (almost same coefficient for L1 and L2 or P1 and P2). However, for receiver #1 we find a non-linear dependence for all observables, except for P2 (where the thermal coefficient is about -0.21 ns/°C) and a strong frequency dependence. Thus two GPS receivers of the same type do not necessarily have the same temperature dependence and the thermal coefficient can reach -0.21 ns/°C for the code.
It is therefore necessary to keep the temperature of the receiver box stable.

5. MIDDLE-TERM STABILITY OF GeTT

To verify the stability of the GeTT method we made a time transfer between our two terminals placed side by side, driven by the same station clock (Cs clock) on a very short base line of 6 m.
The clock difference computed for the code P1 and the phase L1 through a period of 12 days are shown in Figure 6.
Although we use the same clock signals for both terminals, an offset of 15.8 ns appears for the code P1. The corresponding offset for P2 is only 7.2 ns. These offsets are the result of different internal delays in both GPS receivers. They are not a problem as long as they remain constant. However, they must be taken into account in the final clock difference computation. The standard deviation is about 0.53 ns.
For the phase L1 (that is affected by an arbitrary offset, the phase ambiguity) the standard deviation is only 0.017 ns. We can also identify a periodic variation that appears only for the working days (from Monday to Friday, day of year 344 was a Monday). These variations can be explain by the fact that one receiver box was not temperature stabilised and thus sensible to variations of the room temperature. We do not find these variations for Saturday and Sunday when nobody was in the room. If we compute the standard deviation for days of the weekend we obtain a value below 10 ps that we expected [5].

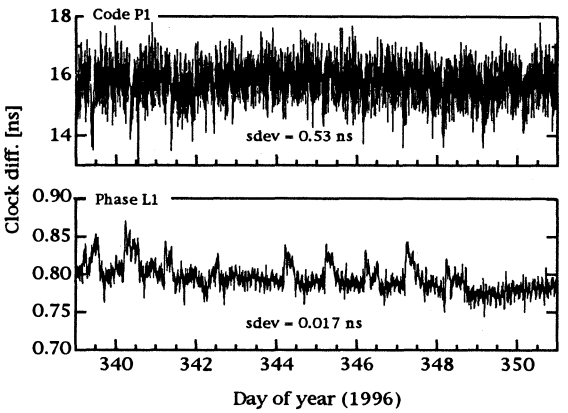


Figure 6: Clock difference between two terminals driven by the same station clock and for a base line of 6 m.

Since we found no correlation with the outdoor temperature we can conclude that either the temperature coefficients of the antennas are below that expected value of - 0.04 ns/°C or both antennas have an identical temperature dependence.
Figure 7 shows the Allan deviation [8] σ_y calculated from the code P1 and the phase L1 of the figure 6.

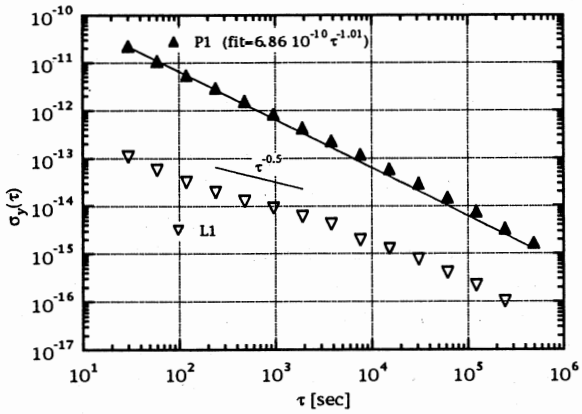


Figure 7: Allan variance calculated from the code P1 and the phase L1

The value of the slope for the code P1 is about -1. That indicates the intrinsic noise of the system is White PM and/or Flicker PM type for τ between 30 seconds and 5 days.

For the phase L1 the short-term noise (t<1000 seconds) is two magnitude less than for the code. The slope is about -0.8 for $\tau < 100$ and $\tau > 10^4$ seconds and about -0.5 for $100 < \tau < 10^4$ seconds.

6. CONCLUSION

In order to reach a sub-nanosecond level in the accuracy of the time transfer using GPS receivers, it is mandatory to take the temperature dependence of the delays through the local signal path into account. For our GeTT system we obtain:

— A maximum of -0.04 ns/°C for the antenna, or our two antennas have the same temperature dependence (to be examined in near future).

— About -0.42 ps/°C/m for the antenna cable (RG213) for temperatures between -20°C and 40°C. The temperature dependence is determined by the dielectric of the cable.

— The temperature coefficient of the GPS receiver can reach -0.21 ns/°C for the code and 0.022 ns/°C for the phase. It is different for the two carrier frequencies.

— Tow receivers of the same type (Ashtech Z-XII) showed surprisingly different temperature dependencies.

The stability of the two GeTT terminals over a period of 12 days is below 1 ns; a standard deviation of 0.54 ns for the codes and below 0.01 ns for the phases is observed. Offsets of about 15.8 ns for the code P1 and 7.2 ns for P2 are present when two terminals are driven by the same clock signals.

The next steps within this study will be a time transfer on a larger base line (5 km and more) and a comparison of the method with other techniques (common-view, TWSTFT, ...).

REFERENCES

[1] N.Ashby, M. Weiss, *"Accurate time transfer during common-view of a GPS satellite"*, Proc. 34th Ann. Symp. on Frequency Control 1980, pp. 334-346.

[2] Th. Schildknecht, G. Beutler, W. Gurtner, M. Rothacher,*"Towards sub-nanosecond GPS Time Transfer using Geodetic Processing Technique"*, Proc. 4th EFTF 1990, pp. 335-346.

[3] P. Baeriswyl, Th. Schildknecht, W. Gurtner, G. Beutler, *"Frequency and Time Transfer with Geodetic GPS Receivers"*, Proc. 7th EFTF 1993, pp. 119-124.

[4] P. Baeriswyl, Th. Schildknecht, J. Utzinger, G. Beutler, *"Frequency and Time Transfer with Geodetic GPS Receivers: First Results"*, Proc. 9th EFTF 1995, pp. 46-51.

[5] P. Baeriswyl, Th. Schildknecht, T. Springer, G. Beutler, *"Time Transfer with Geodetic GPS Receivers: using Code and Phase Observation"*, Proc. 10th EFTF 1996, pp. 430-435.

[6] Jan Cermák, *"A GPS Parallel Common-View Experiment"*, Proc. 10th EFTF 1996, pp. 296-301.

[7] W. Lewandowski, *"Sensitivity to the External Temperature of some GPS Time Receivers"*, Proc. of the 22nd PTTI Appl. and Planing Meeting 1990, pp. 307-316

[8] M. A. Weiss et al, *"Confidence on the Modified Allan Variance And the Time Variance"*, 9th EFTF 1995, pp. 153-165.

SATELLITE EARTH STATIONS FOR TWO-WAY TIME TRANSFER AT THE TECHNICAL UNIVERSITY GRAZ

H Ressler[1], D Kirchner[2], R Robnik[1]

[1]Space Research Institute, Graz, Austria
[2]Technical University Graz, Austria

ABSTRACT

At the Technical University Graz (TUG) there are two earth stations for two-way time and frequency transfer (TWSTFT) using geostationary communication satellites available, both equipped with a delay monitoring system. The hardware implementation is described and the facilities offered by these two co-located earth stations are demonstrated. Maximum delay variations in normal operation and some signal dependencies on operational parameter variations are reported.

Keywords: Two-way satellite time and frequency transfer (TWSTFT), satellite earth station.

1. INTRODUCTION

A two-way measuring system mainly consists of a satellite earth station (E/S), a modem and counter, and in most cases distribution amplifiers providing frequency and one pulses per second (1pps) of the local time scale. The knowledge of the signal delay, in particular, the differential E/S delay (transmit minus receive path) and of the relation of the 1pps transmitted by the modem to the local time scale is essential for the uncertainty of the TWSTFT method. The differential E/S delay can be determined comparing the earth station with a calibrating station at the same location and referred to the same time reference. In most cases the relation of the transmitted 1pps to the local time scale can be measured by means of a time interval counter using long cables. Common to both measurements is the fact that they are snapshots at a certain time, that cannot detect delay changes due to long term effects such as environmental effects and aging. Therefore a permanent delay monitoring should be performed, whose results - applied to the time transfer measurements - improve the accuracy of this method.

2. MEASUREMENT SET-UP

At the TUG there are two TWSTFT systems available, both systems can be operated in a fully automated way. In addition to the delay monitoring, various signals are measured to study the behavior of parts of the system. The setup of two systems improves the reliability using the same components and employing the same operational parameters, and makes it possible to evaluate the stability in a common clock mode.

2.1 Principle of operation

Fig. 1 shows a block diagram of the two earth stations with its interconnection possibilities.
The 1pps and frequency distribution amplifiers provide signals which are phase coherent to UTC(TUG). The modem generates an internal 1pps (1PPSTX), which is derived from the frequency but synchronized to an external 1pps and modulated onto the stations intermediate frequency (IF), which is usually at 70 MHz. The TX signal is adjusted to appropriate power level and frequency by means of the IF unit, up-converted to RF frequency, amplified, and finally transmitted to the satellite. The RF signal received from the satellite is amplified and down-converted to the IF, again adjusted in power and frequency to match the requirements of the RX input of the modem. In a two-way time transfer, the time difference between the 1PPSREF and the demodulated 1PPSRX representing the transmitted 1PPSTX of the second laboratory can be measured by means of a time interval counter.
In order to check the performance of the equipment switches are installed connecting the transmit part directly with the receive one at various stages beginning with the modem up to the satellite simulator. The test loop of the modem, a part of the modem itself, gives the sum of the transmit and receive delay of the modem. The indoor loop (ID loop) includes the IF unit. The outdoor loop (OD loop) adds the cables. The loop via the satellite simulator covers the VSAT too. If the modulated signal generated by the modem passes the transfer switch (X switch) directly to the satellite simulator and the 70 MHz signal provided by the PLO is transmitted on the VSAT, then the loop includes the receive part of the earth station and the cables to the satellite simulator. Thus, on some assumptions, the differential E/S delay can be calculated (Ref. 1). A synthesizer provides the satellite simulator with a local oscillator (LO) signal, whose frequency corresponds to the translation frequency f_T of the satellite (receive minus transmit frequency) minus 70 MHz.
The last accessible 1pps related to the local time scale is the 1PPSTX of the modem, which has to be measured to get the transmit epoch of the signal. A multiplexer channels the signals to the stop input of the counter in order to perform time interval measurements and to check the frequencies of the LOs and the performance of the counter itself measuring the delay of a cable.

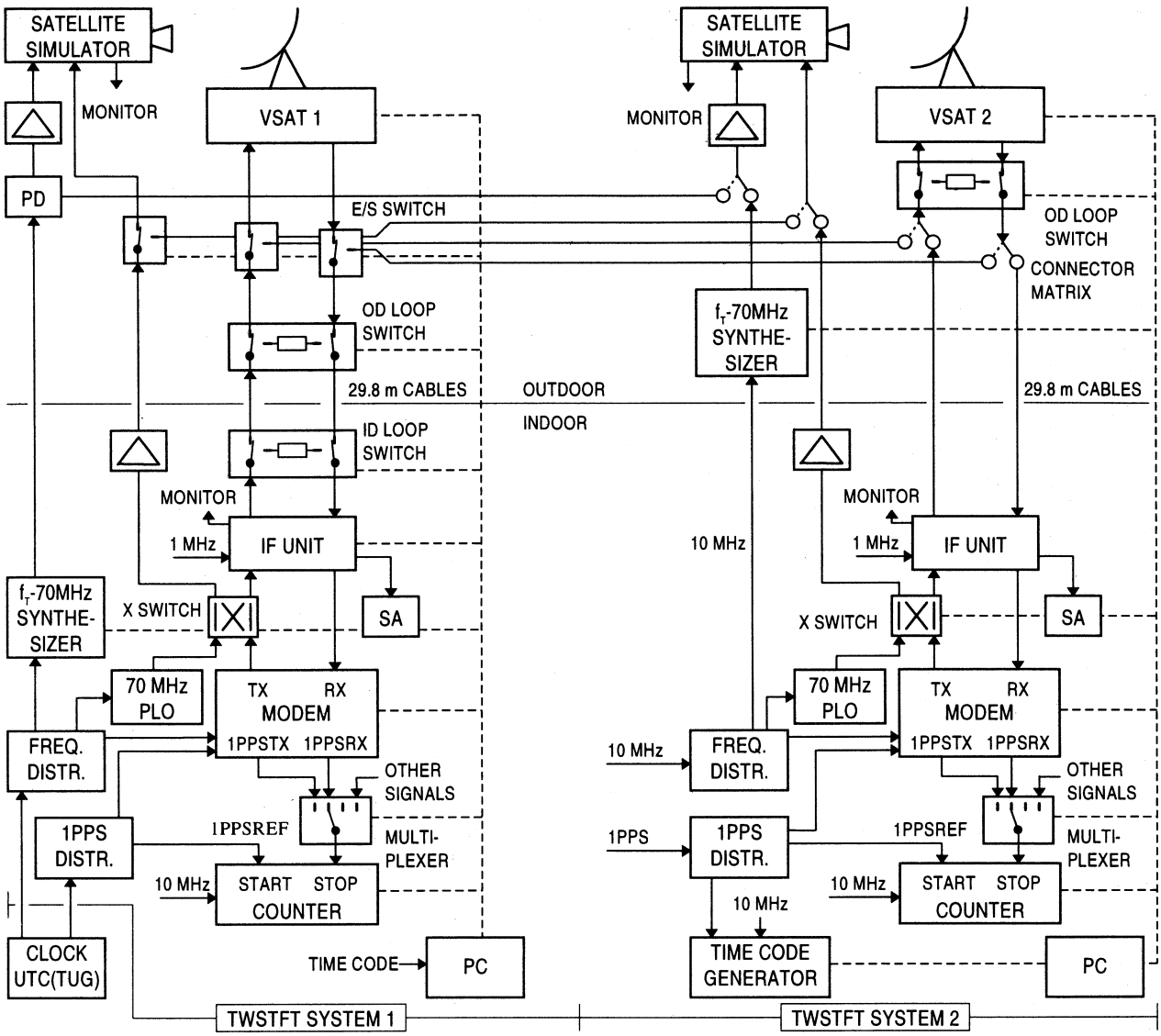


Figure 1: Measurement set-up

The spectrum analyzer (SA) is used to monitor the spectrum, in particular, the carrier to noise density ratio and the variation of the noise level.

2.2 Automation

The devices linked with dashed lines in Fig. 1 are connected to the PC employing RS-232, RS-485 and IEC-488 interfaces. For some devices it was necessary to develop special interfaces capable to communicate with the serial port of a PC. The in-house made control equipment was realized by means of a microcontroller and programmable logic devices (EPLDs). The software, running under DOS, allows unattended operation. It consists of two modules, one to configure the tasks of the measurement program and the measurement schedule, stored on the hard disk, and the other one to execute the measurements and to control the system in a quasi real time multitasking mode with a time slicing of 20 ms. The use of the European time code (Ref. 2) guarantees an exact time tagging of the data.

2.3 Earth station arrangement

Both VSATs are located side by side on the roof of the laboratory. All the cables are ending in a commonly used cabinet suitable to place the E/S switch and the connector matrix. This arrangement provides two main modes of operation:

1. Independent operation of both earth stations.
2. Mixed operation of the indoor equipment of system 1 and the outdoor equipment of system 2. IF the cables are connected according the dashed lines in the connector matrix of Fig. 1 then the IF signals can be routed to the VSAT 2 and the 70 MHz signal to the satellite simulator by means of the remote con-

trolled E/S switch. The f_T-70 MHz signal can be supplied either by means of a power divider (PD) from VSAT 1 or by a separate synthesizer. To get access on the control lines, they can to be switched too.

For test purposes it is possible to connect both pieces of indoor equipment by means of the connector matrix as well as to operate the indoor equipment of system 2 with the outdoor equipment of system 1.

3. SYSTEM COMPONENTS

Both TWSTFT systems - set up according to the design considerations given in Ref. 3 - are equipped with a delay monitoring system using satellite simulators of the de Jong type (Refs. 4, 5). One system is designed for stationary use only, the second one for stationary use and/or calibration purposes consisting of few transportable pre-assembled pieces.

3.1 1pps and frequency distribution

For the TWSTFT system exclusively in-house made distribution amplifiers are used. Ground loops are avoided by means of RF transformers. As a result of permanent checks, some upgrades of the frequency distribution were carried out in the past. This work is still in progress.

3.2 Counter and modems

The time interval counter currently in use is the model SR620 from Stanford Research Systems, Inc. with an ovenized VCXO. Using an external frequency source the specified resolution is typically 25 ps but actually performing better, the accuracy ±500 ps for absolute and ±50 ps for relative measurements, respectively.
For the modem (MITREX 2500) a separate synchronization unit is used because dependencies of the synchronisation (1PPSTX of the modem referred to an external 1pps) on signal level and puls width were observed with the built in one. A remote controlled module, also imitating the front panel features of the modem, was developed to integrate the modem into the automated measuring system. This was easily accomplished routing the signals from the control board mounted on the rear site of the front panel to the TX and RX assembly through the external control unit, replacing the top cover of the modem.
The TWSTFT system 2 is designed to operate with a SATRE modem, at the present under test, which offers frequency agility and signal level adjustment and further features such as selectable digital filters in the transmit part, additional codes with higher chip rates, data transfer capabilities, a built-in time interval counter and of course remote control.

3.3 IF unit

Both VSATs provide frequency agility in steps of 1 MHz only and the IF of the MITREX modem is fixed to 70 MHz. The allocated frequencies on satellites may be assigned on a 22.5 kHz grid and additionally the receive frequency may change by some kHz up to about ±50 kHz due to the deviation of the transponder translation frequency from its nominal value (Ref. 3). To solve this problem at the TUG an IF unit was developed, interfacing the modem to the VSAT. The purpose of the IF unit is the adaptation of the signal power and frequency provided by the modem to the requirements of the

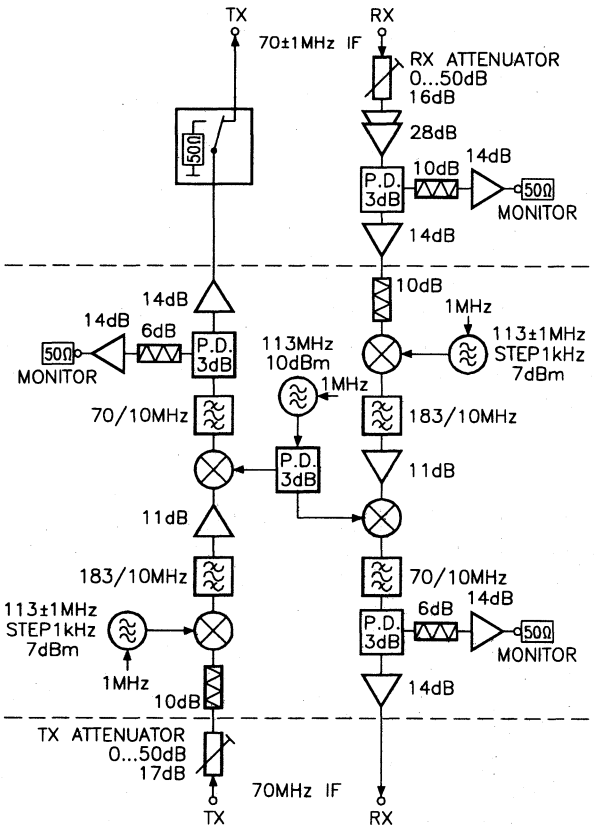


Figure 2a: IF unit of TWSTFT system 1

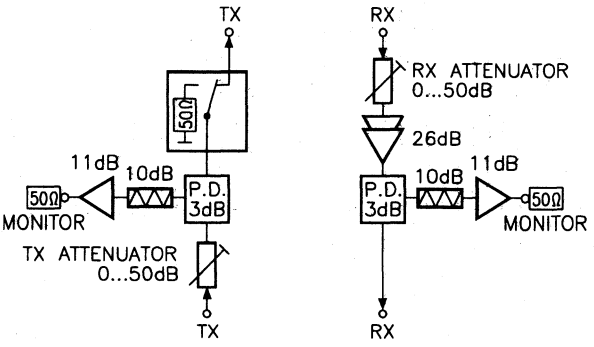


Figure 2b: IF unit of TWSTFT system 2

VSAT and subsequently the satellite and vice versa. Additionally it provides access to the IF signals during operation to monitor signal level, frequency and carrier to noise density ratio. The frequency conversion module - indicated by the dashed lines in Fig. 2a - is omitted in the IF unit of system 2 because it is designed to operate with modems providing this capability (Fig. 2b). There is one VHF switch installed that deliberately can only be activated by hand to turn on and off the transmitted signal independent from the current status of the program controlling the system.

3.4 VHF Switches

The VHF switches have to be suitable to operate with frequencies up to at least 70 MHz (IF signals) and are designed for 50 Ω termination. There are two types in use, commercial available coax relays for signal power up to 0.3 W with a specified frequency range from DC to 500 MHz, an isolation of 65 dB and an insertion loss of less than 0.5 dB. The second ones are in-house made switches (loop switch, transfer switch and multiplexer) with an isolation of at least 40 dB measured at 100 MHz and an insertion loss of less than 0.5 dB. The loop switches contain attenuators to match the power to that one usually received from the satellite at the appropriate location. To avoid mismatch on the connected coax cables all non routed inputs and outputs are terminated with 50 Ω.

3.5 VSATs

Both VSATs (ASAT 1214 from SSE Technologies with several modifications) are designed for two-way satellite communications system that operate in the Ku-band frequency range. The function of the terminal is the up-conversion of the 70 MHz IF signal to an RF signal in the 14.0 to 14.5 GHz band for transmission via an antenna and the down-conversion of the received RF signal (10.95 to 12.75 GHz) to an 70 MHz IF signal (Ref. 3). Both VSATs are provided with a 5 W solid state amplifier with a stated 1 dB compression point of 4 W. Fans mounted on the heat sinks prevent the devices from overheating. In case of VSAT 2 the fan is temperature controlled. VSAT 1 is equipped with a programmable attenuator (0 to 31 dB) that permits the output power to be adjusted in steps of 1 dB. Several different LNCs with a noise temperature of 110 K are available covering a frequency range from 10.95 to 12.75 GHz. The use of 1.8 m offset antennas (Andrew for VSAT 1, Prodelin for VSAT 2) gives a maximum EIRP of about 52 dBW and a G/T of about 22 dB/K. The monitor and control assembly, installed in both VSATs permits the remote controlled monitoring of various operational parameters and is used to set the transmit and receive frequencies as well as the transmit power (VSAT 1 only).
It should be emphasized that the protection of the outdoor unit against rain and snow is very important to avoid moisture in the outdoor units and corrosion of the connectors. This was done by means of self bonding sealing tapes and individually formed aluminium metal sheets.
Apart from the antenna mount and the antenna itself, the entire outdoor equipment of VSAT 2 is partitioned in two parts. The first one, consisting of the feed boom with pre-assembled feed, HPA, LNC, satellite simulator and all the cables belonging to it, can be stored in a wooden box for transport purposes. The power supply and driver unit of the ASAT 1214, the f_T-70 MHz LO and the outdoor loop switch are housed in metal cabinet with a feedthrough for the cables, thus all the connections can be done in a protected environment. The box is prepared for use with a temperature controlled fan.

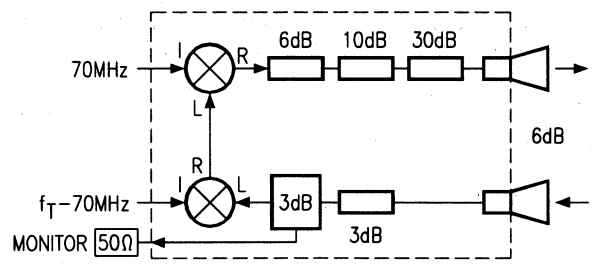


Figure 3a: Satellite simulator of the delay monitoring system

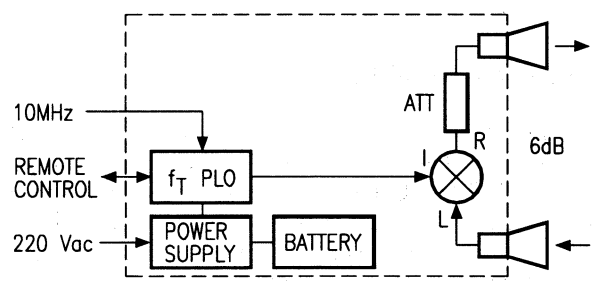


Figure 3b: Satellite simulator for test purposes

3.6 Delay monitoring equipment

By means of the delay monitoring equipment it is possible to measure the sum of the E/S transmit and receive delay (from the modem 1PPSTX to the satellite simulator and back to the modem 1PPSRX) and the receive part only, thus allowing to calculate the differential E/S delay (Ref. 1). The equipment consists of a the transfer switch, the satellite simulator (Fig. 3a), a synthesizer providing the f_T-70 MHz signal, a 70 MHz local oscillator and amplifiers.
Recently at the TUG a remote controlled PLO providing the satellite translation frequency ±15 MHz programmable in steps of 1 MHz with adjustable power level up to 10 dBm was developed. The center frequency is selectable by means of different ceramic resonators soldered on the printed circuit board. The small RF sealed aluminium case (8x5x3 cm) containing the PLO can be installed in a stand-alone device which offers operation

from keyboard as well as remote control through a serial interface. A third satellite simulator using this PLO was assembled for a test set-up at which both antennas are pointed to the satellite simulator (Fig. 3b).

4. OPERATIONAL EXPERIENCES

The TWSTFT system 1 was set up in 1990 and is in continuous operation since end of 1993, system 2 since end of 1995, using an uninterruptible power supply. During this period there only occurred three hardware faults (LO in VSAT 1, LO in the IF unit of system 1 and a relay in a modem) proving the high reliability of the system.
Delay dependencies of system components on various parameters were tested resulting in operating procedures one has to stick to. Some of the results are given below.

Changing loads (mismatch) on signal outputs:
Delay changes of the indoor loop up to 600 ps were measured depending on the termination - 50 Ω, not connected, connected spectrum anlyzer turned off - of the monitor output of the IF unit of system 1. Phase changes up to the ns level were observed on various frequency distribution amplifiers if an unused output was terminated with 50 Ω. This circumstance and the periodical effects measured from time to time caused step by step improvement of the 1pps and frequency distribution during the last years using in-house made distribution amplifiers (Ref. 6).

Changing LO power level:
A power variation of 4 dB at the LO inputs of the satellite simulator resulted in a delay change of 220 ps in the receive loop and 520 ps in the receive plus transmit loop, respectively.

Counter:
To check the performance of the counter, test measurements are done using a 1pps to start and the same 1pps but passing a distribution amplifier to stop the counter. The nominal time interval is determined performing relative measurements against a third 1pps. Measurements during the last 3 years on a regular basis (1 single shot measurement each hour) revealed that counter calibrations according to the procedure described in the user manual of the manufacturer (tuning of the oscillator included) as well as phase changes of the frequency at the reference input of the counter could cause permanent changes of the counter readings up to ±50 ps. After the initial calibration the offset referred to the nominal time interval was of the same order.

To summarize the behaviour of the two-way measuring system 1 the peak-to-peak variations of different measurements during a period of about 450 days (210 days for 1pps and frequency distribution) taken on three days per week are given in table 1. One point represents the mean of 100 samples and in case of the counter test the mean of 10 samples. During the reported period of the counter check there were no calibrations but phase steps in the frequency fed into the reference input of the counter. The indoor temperature was stabilized to about 1 °C, the change of the outdoor temperature amounted to 35 °C. The results of table 1 demonstrate the high precision of the TWSTFT as in use but it has to be pointed out, that the given measurements - apart from the modem loop - can be used to correct the two-way time transfer measurements allowing an significant improvement of the measurement uncertainty. According to an error budget the estimated uncertainty of the differential E/S delay is below 50 ps (Ref. 6) and the uncertainty of the other contributions is even smaller. Thus the modem seems to be the limiting factor as long as the signal delay variations of the modem are of the order given in table 1 and cannot be measured in course of the two-way measurements.

Measurement	Samples	Peak-to-peak variation ns
Counter check	387	0.07
1pps and freq. distr. including 1PPSTX generation of the modem	175	0.19
Modem loop	167	0.15
Diff. E/S delay of system 1	163	1.52

Table 1: Peak-to-peak variation of measurements during TWSTFT using the INTELSAT satellite at 307°E

REFERENCES

[1] D. Kirchner, H. Ressler and R. Robnik, "An automated signal delay monitoring system for a two-way satellite time transfer terminal", Proc. 9th European Frequency and Time Forum, pp.75-79, 1995.

[2] D. Kirchner, H. Ressler, "Zeituebertragung ueber Telefonmodems", Verlag Sprache und Technik, Funkuhren, Zeitsignale Normalfrequenzen, pp. 185-199, 1993

[3] D. Kirchner, "Design considerations for Ku-band two-way satellite time transfer terminals", Proc. 4th European Frequency and Time Forum, pp.631-637, 1990.

[4] G. de Jong, "Accurate delay calibration of satellite ground stations for two-way time transfer", Proc. 9rd European Frequency and Time Forum, pp.198-203, 1989.

[5] G. de Jong, "Automated delay measurement system for an earth station for two-way satellite time and frequency transfer", Proc. 26th PTTI, 1994.

[6] D. Kirchner, H. Ressler and R. Robnik, "Signal stability of a Ku-band two-way satellite time transfer terminal", Proc. 27th PTTI, pp. 303-311, 1995

The Time and Frequency National Service of Ukraine Pecularities of Organization and Development

O.N. Velitchko
State committee Ukraine for standartization, metrology and sertification - Kiev - Ukraine

There are three metrology centres Kiev, Kharkov and Uzhgorod in Ukraine. They are equiped by working time and frequency standards (their accuracy $4\star10^{-13}$-$5\star10^{-14}$) and several astrometrycal and laser observation posts. But we haven't State Primary Time and Frequency Standard, Atomic and Coordinated Time National Scales, high accuracy systems of transference Time and Freguency Standard Signals by National broadcast and other communication channels including satellite (GPS-receiver) and radio-meteor.

In 1994 the State Programme of Organization and Development the Time and Frequency National Service on 1995-1996 years was worked out in Ukraine. This Programme was confirmed by Goverment of Ukraine resolution in 1995. The main tasks of the Programme are organization of the United National Service, that will satisfy needs of all consumers, ensure maximum use of existing scientific-technical complexes of the country, use of the results of international collaboration scientific-technical complexes and systems of other countries fist of all European, Russia, USA and participation in the international organization COOMET and other international organizations work.

The object of research and elaboration in limits of Programme is creation sufficient scientific-technical base and optimum organization ot the structure of service, technical base for creation of the National Reference Group of Time and Frequency (to an accurasy of $1\star10^{-14}$) that is need.

A STUDY EXAMINING THE POSSIBILITY OF OBTAINING TRACEABILITY TO UK NATIONAL STANDARDS USING GPS DISCIPLINED OSCILLATORS.

J A Davis and J M Furlong

Centre for Electromagnetic and Time Metrology,
National Physical Laboratory,
Queens Road, Teddington, Middlesex,TW11 OLW.

ABSTRACT

NPL is undertaking an extensive study to examine the suitability of using GPS Disciplined Oscillators (GPSDOs) as standards traceable to UTC(NPL). The performance of 15 GPSDOs loaned by UK manufacturers and suppliers are being examined at NPL. Continuous measurements are made over several weeks of both the 1 Pulse Per Second (1PPS) and the phase of the 5 or 10MHz output signals from each GPSDO. NPL is continuously monitoring the GPS system using three multichannel GPS receivers, and also making measurements of the local environment. Any correlation between changes in the GPSDOs under test and changes in either the GPS system or local environment may then be examined. The GPSDOs are also being examined under non-ideal operating environments, for example operating with only a limited view of the sky. A description of the various aspects of the study is provided, along with some preliminary results and a discussion of future work.

Keywords:
Time transfer, time dissemination, GPS Disciplined Oscillators, traceable time standards.

1. INTRODUCTION

The use of GPS Disciplined Oscillators (GPSDOs) as time and frequency standards has been increasing in popularity within the United Kingdom. However the growth of the use of GPSDOs in calibration laboratories is being limited because GPSDOs are not recognised as standards traceable to the national time scale UTC(NPL). The use of GPSDOs offer advantages over more conventional methods of time and frequency dissemination, e.g., the reception of NPL's MSF 60 kHz signals, both in increased time and frequency dissemination accuracy and in global coverage. GPSDOs are widely used internationally, in particular within the telecommunications industry [Ref. 1].

In response to growing interest in using GPSDOs, NPL is undertaking a study of their performance. This study aims to determine the suitability of using GPSDOs as standards traceable to UTC(NPL), and also the time and frequency dissemination uncertainties associated with their use. NPL is working in close collaboration with the United Kingdom Accreditation Service (UKAS), who may ultimately use results from this study in their own evaluation of the suitability of using GPSDOs as traceable standards. Similar studies have been performed by other countries [Ref. 2].

In this paper, an account is given of NPL's study of GPSDOs. This includes a description of the measurements under-way and preliminary results. An outline of the future direction of the GPSDO study at NPL is also provided.

2. ELEMENTS OF THE GPSDO STUDY

Measurements on the GPSDO project started at the beginning of January 1997 and will continue until the end of March 1997. Eleven organisations have provided NPL with 15 GPSDOs for test during this period. These are: Absolute Time, Datum, Efratom, Hewlett Packard, Navstar, Radiocode Clocks, Rapco Electronics, Quartzlock, Tekelec, Trak and Truetime. In addition, the performance of three GPSDOs are being monitored simultaneously by Motorola in Swindon, Wiltshire (two) and by Hewlett Packard, Winnersh (one).

Extensive measurements are being made on each of the GPSDOs. A 16 channel phase comparator (TimeTech, model PComp 16-001/96) is used to compare the phase of the GPSDO's 5 or 10MHz standard frequency output, against a standard frequency derived from NPL's Active Hydrogen Maser (Sigma Tau, model MHM 2010) generating UTC(NPL). The phase difference is recorded every second.

The 1 Pulse Per Second (1PPS) signal generated by each GPSDO is compared against a 1PPS signal derived from UTC(NPL). The output from the 15 GPSDOs are fed simultaneously into two switch boxes. Each switch box has two sets of four-way switches. The output from each four-way switch is fed as a 'stop pulse' into a Universal Counter Time (Racal Dana 1991). The 'start' pulse is derived from a 1PPS signal generated from UTC(NPL). This switching arrangement results in a duty cycle for logging the GPSDOs 1PPS of one minute on, three minutes off.

The GPSDO laboratory is located in an adjacent building to the one which houses the UTC(NPL) signal. Changes in cable delays between the two buildings are monitored by recording the phase difference between a direct 5MHz input into a phase comparator, and a 5MHz signal which is looped between the two buildings. Delay changes will be recorded as phase changes in the returned signal. The GPSDO logging system is shown in figure 1. The time of the 1PPS pulses are calibrated against UTC(NPL) by using a portable HP5071A caesium clock.

Each GPSDO antenna is mounted on a 5m high aluminium pole, which typically extends 1-2 m above the height of the roof. There is a minimum spacing of 1 or 2m between any two poles depending on whether or not the antenna is *passive* or *active* respectively.

Both the GPSDO laboratory and the external roof environmental conditions are monitored continuously e.g., air temperature, humidity and surface temperature using temperarture/humidity loggers (Grant Squirrel Loggers model no. 1200).

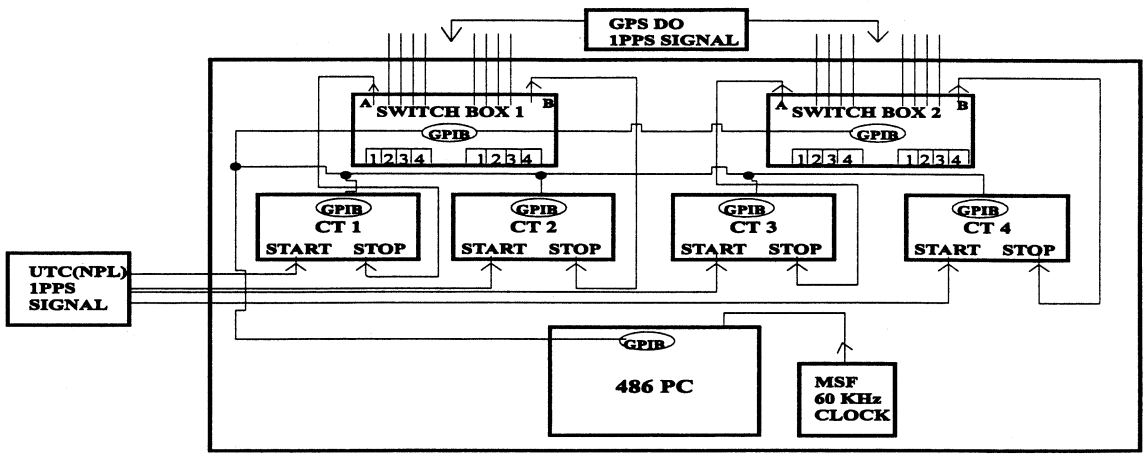


Figure 1: Experimental arrangement at NPL for the 1PPS logging system

The status of each GPSDO is being monitored by collecting data available via the GPSDO's RS232 port. The information varies considerably between the different receivers. A typical GPSDO may provide information on the number of satellites tracked, their azimuth, elevation and signal strength, health of the GPS constellation, dilution of precision (DOP) values, DAC corrections, GPS lock and instrument status information. The antennae coordinates e.g., latitude, longitude and vertical height are also monitored continuously. An example of the information is shown in figure 2, which shows the satellites tracked by a typical GPSDO unit.

NPL is monitoring the status of the GPS constellation throughout the course of the study. An eight channel receiver (Allen Osborne Associates model TTR-4P) is used to track the GPS constellation and receive the Course Acquisition (C/A) coded signals. The TTR-4P is set to record several data sets every 30 seconds. These include measurements of the time transfer (Local clock - UTC(USNO)) calculated for each GPS satellite, satellite azimuth and elevation angles, modelled and measured codeless ionospheric delay corrections, satellite signal strengths, and receiver position. An example of the measurement of the local clock - UTC(USNO) for three satellites are shown in figure 3.

The short term periodic variations are due principally to the effects of S/A on the GPS satellites. This may be clearly seen in figure 3, where the variations of satellite PRN 15 (free from S/A) are significantly smaller. This effect is likely to also limit the performance of the GPSDOs. The local clock has been offset by approximately 2 microseconds from UTC(NPL), ensuring that the 1PPS(UTC) signal is always ahead of the 1PPS(GPS) signal.

Two, eight channel C/A code, Motorola VP Oncore GPS receivers [Ref. 3] are also used to monitor the GPS constellation. These receivers act both as a back-up to the TTR-4P receiver, and also provide a separate measure of the status of the GPSDO signals received at NPL. The Oncore receivers have been set to record several sets of information: pseudo-range correction data, positional information and RAIM data. Measurements of the Local Clock - UTC(USNO) differences and signal strength for each satellite, along with the user position (entered fixed position) have been made. Plots of pseudo-range correction data for three satellites are shown in figure 4.

NPL is complementing its own GPS measurements with data from the USNO WWW site. Information is obtained on the health and outages of each GPS satellite, and the mean values of GPS Time - UTC(USNO) for each satellite. These latter GPS Time - UTC(USNO) values are of interest for comparison against GPS measurements made at NPL. The value of this data is limited because the observable satellites at any one time will be different at NPL and USNO. Also, the data from USNO's WWW site does not provide a continuous record for each satellite.

NPL is performing measurements of (UTC(NPL) - UTC(USNO)) using both the common-view of GPS satellites [Ref. 4] and Two-Way Satellite Time and Frequency Transfer [Ref. 5]. The accuracies obtained from these techniques are substantially higher than those obtainable using the direct reception of GPS satellites. These measurements are being used to accurately relate UTC(NPL) to UTC(USNO), and also to determine any changes in the value of UTC(NPL)-UTC(USNO). The two time scales UTC(NPL) and UTC(USNO) have been directly compared using a portable HP5071A caesium clock [Ref. 6]. NPL intends to borrow the BIPM calibrating GPS receivers in the near future, to obtain a further measure of UTC(NPL) -UTC(USNO).

The local environment of the GPSDOs is being continuously monitored throughout the study. Figure 5 shows the typical temperature variation of an external antenna pole.

3. PROGRAMME OF GPSDO MEASUREMENTS

A wide variety of measurements are being performed on the GPSDOs under study at NPL. Initially the units have been operated continuously for several weeks. The GPSDOs are normally operated in a position hold mode, although some use an average or even instantaneous position. These comparisons against UTC(NPL) are being used to statistically assess the performance of GPSDOs.

The use of inaccurate antenna coordinates may degrade the performance of the GPSDOs. The performance of the GPSDOs operating using instantaneously determined, averaged andonstant antenna positions are under comparison, as is the effect of operating GPSDO with slightly offset (inaccurate) fixed antennae coordinates. The determination of the antenna

Figure 2: Satellites Tracked by a Typical GPSDO unit

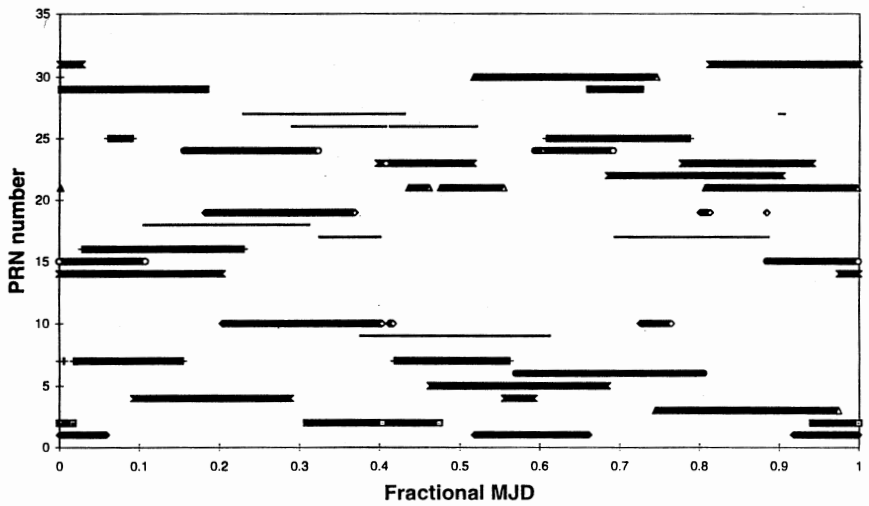


Figure 4: Pseudo-range correction data for an Oncore receiver

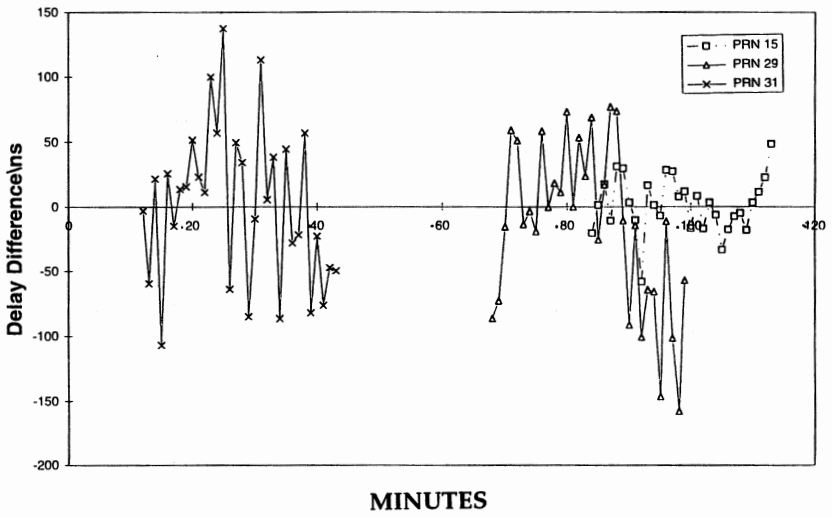


Figure 3: Local Clock - UTC(USNO) measured for three different GPS satellites using the TTR-4P receiver

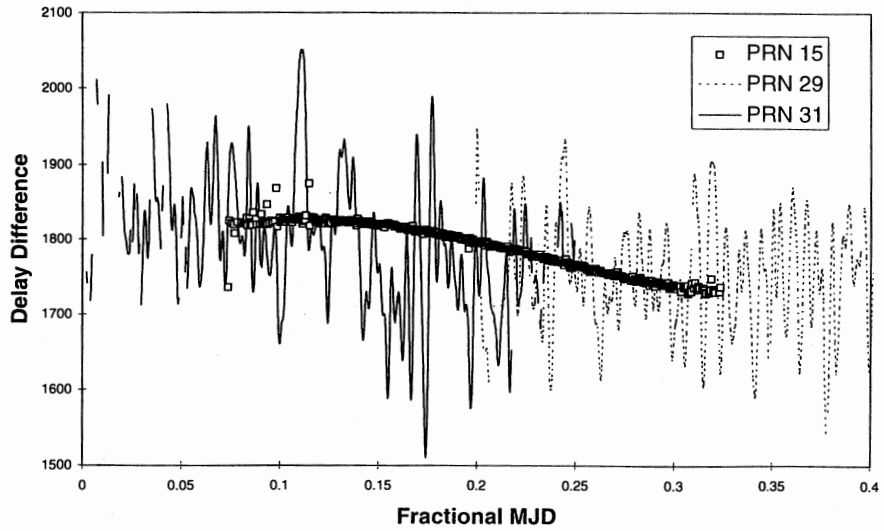
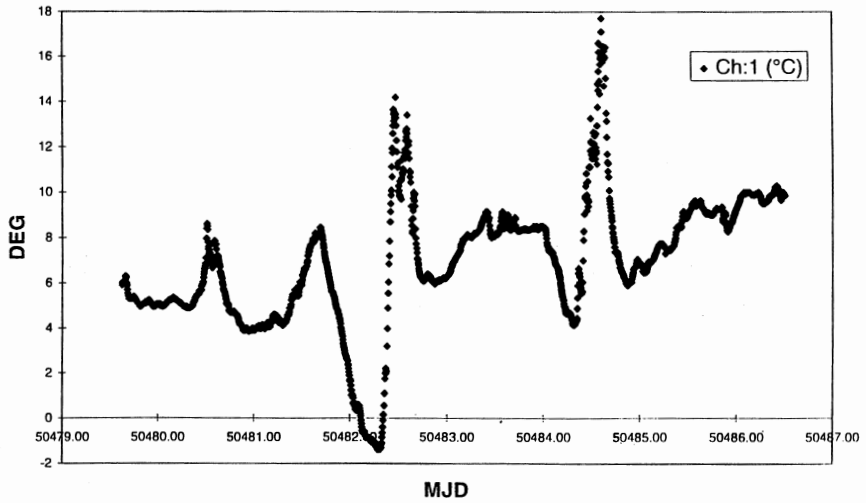


Figure 5: Antenna Pole Surface Temperature



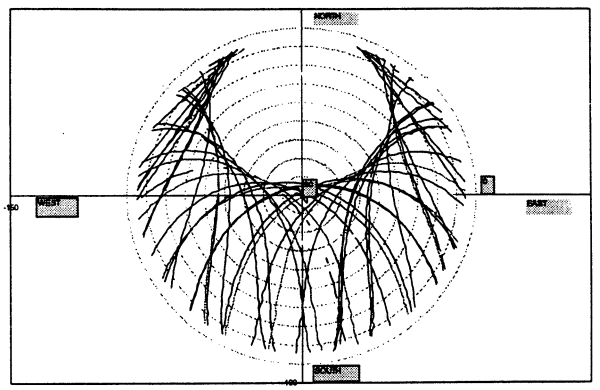


Figure 6: Polar plot of healthy GPS satellite tracks across the sky as viewed from NPL. Dashed lines correspond to 10 ^{0} elevation contours.

coordinates is likely to be an important part of the installation procedure for GPSDOs claiming traceability to UTC(NPL).

The GPSDOs are being studied when operating with a restricted view of the sky. This part of the study will be important for determining the effects of setting up a traceable GPSDO installation at a site where the whole sky is not visible. For many GPSDOs, the simplest way to perform this test is to change the satellite elevation mask angle. Alternatively, it is possible to remove satellites from the tracking schedules of the GPSDOs.

These tests also examine the effect of operating the GPSDOs when only a partial constellation of GPS satellites are available, which may be useful if in the future the GPS system were to operate with less than a full constellation of satellites. Antennae locations have been prepared at NPL where local obstacles are masking part of the sky. Operating the GPSDOs from these locations will not only examine the effect of a limited field of view, but also examine the effect of local multipath. Figure 6 shows the trajectories of all healthy GPS satellites observable from the NPL site over one day. The GPS satellite orbits reach a maximum latitude of 55 degrees. This results in an absence of satellites over a large portion of sky when looking north from NPL.

The stability of the free running oscillator is a limiting factor in the performance of GPSDOs. Almost all of the GPSDOs will operate in the absence of GPS signals this will allow the performance of the free running oscillators to be characterised.

The Time-To-Alarm of each GPSDO will also be determined, that is the time between the loss of GPSDO signal and the reporting of the loss by the GPSDO. This is important for high accuracy applications of GPSDO where it is not acceptable to operate for any length of time on a free running oscillator. Tests are also being made on the effects of reacquiring a GPS signal after the loss of the signal for a short period, and also on the length of time a GPSDO may operate in the absence of a GPS signal, before its performance is outside is normal operating range.

The start-up performance of GPSDOs are being examined under a variety of conditions, both under a "cold start" where both previous position and satellite ephemeris data is lost and under a "warm start" where both position and ephemeris data is retained. This test may determine the length of time that a GPSDO must be left after switching on, so as to ensure that it is performing within its normal operating range.

4. PRELIMINARY RESULTS

Examples of the variation of the 1PPS signals and phase of the 5 or 10MHz reference are shown in Figures 7 and 8 for a high quality Rubidium GPSDO and a less expensive Quartz GPSDO respectively. The phase angle has been converted to nanoseconds delay.

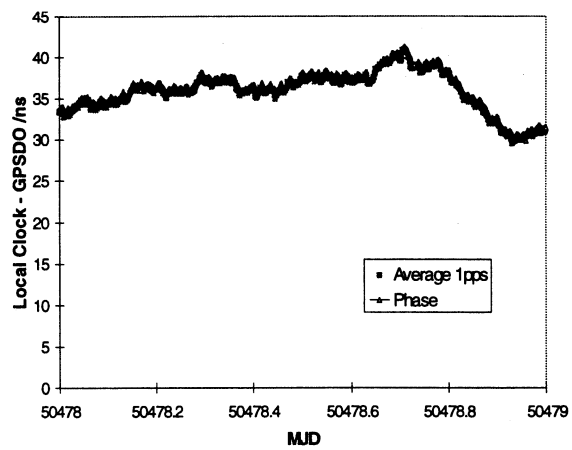


Figure 7: Variation of 1PPS signal and phase of standard frequency output from a Rubidium GPSDO, over a 24 hour period.

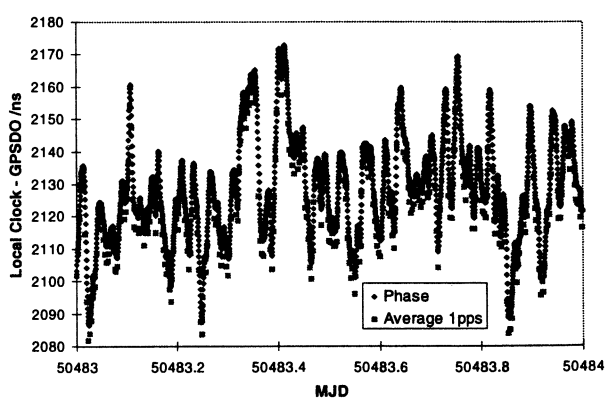


Figure 8: Variation of 1PPS signal and phase of standard frequency output from a Quartz GPSDO , over a 24 hour period.

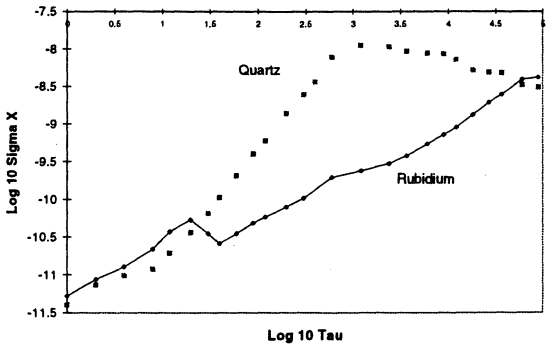


Figure 9: Plots of $Log_{10}(\sigma_x)$ against $Log_{10}(\tau)$, from the phase data of a quartz and a rubidium GPSDO unit.

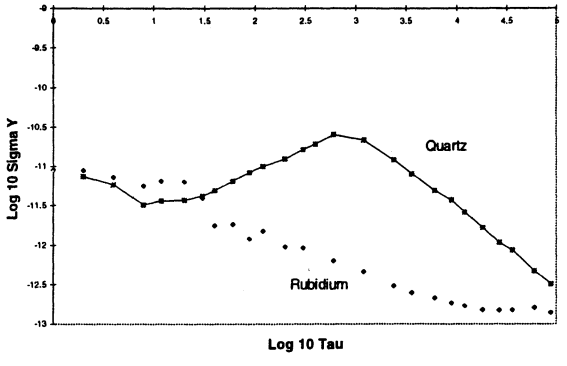


Figure 10: Plots of $Log_{10}(\sigma_y)$ against $Log_{10}(\tau)$, from the phase data of a quartz and a rubidium GPSDO unit.

The 1PPS and phase data correlate well in these two examples, however this is not the case for all the GPSDOs under test. The 1PPS signal originating from a GPSDO with a rubidium oscillator is inherently more stable.

Curves of $Log_{10}(\sigma_x)$ against $Log_{10}(\tau)$ and $Log_{10}(\sigma_y)$ against $Log_{10}(\tau)$, are shown in figures 9 and 10 respectively for the same GPSDOs in figure 7 and 8. These plots are produced using the phase output from each GPSDO. The large differences between the two curves which exist at τ=1000s, correspond to the behaviour of a typical quartz and rubidium oscillator inside the GPSDO. For averaging times (τ) of 1000s, the advantages of using a high quality rubidium oscillator is demonstrated as the disciplining of the GPS system is limited by the effects of SA.

6. FUTURE WORK

NPL's study of GPSDOs is still in its early stages, and only a limited analysis has been performed on a small subsection of the data. NPL intends to report on the completed study later. The immediate objective of the study is to determine if GPSDOs may be used as standards traceable to UTC(NPL). A report and correlation will be forwarded to UKAS on the completion of this study. It is expected that it will be necessary to develop procedure documents; in particular to describe the procedures necessary to set up a GPSDO at sites where they are intended to be used as traceable standards. Additional analysis will be required to determine the level of traceability that may be given to individual models of GPSDOs.

5. CONCLUSIONS

NPL has obtained a substantial quantity of information on the performance of GPSDOs. Initial results show significant variables in both the accuracy of the 1PPS outputs and frequency stability between different models of GPSDOs. The results of the study will be needed to assess the suitability of using GPSDOs as standards traceable to UTC(NPL).

ACKNOWLEDGEMENTS

The authors thank the manufactures and suppliers of GPSDOs for their support and loan of equipment.

This work has been supported under the UK National Measurement System Programme for Time and Frequency.

REFERENCES

[1] Wong S., 'How To Synchronise Telecommunications Networks', Proc. 11th EFTF, Neuchatel, this issue.

[2] Fisk P., Quigg S., 'Preliminary Report: Evaluation and Comparison of GPS-Disciplined Oscillators for use in Traceable Frequency Dissemination', Proc. 27th PTTI meetings, 1995.

[3] Lewandowski W., Moussay P., Guerin P., Meyer F., Vincent M., 'Testing Motorola Oncore GPS Receiver and Temperature-Stabilized Antennas for Time Metrology', Proc 28th PTTI meetings, 1996.

[4] Lewandowski W., Thomas C., 'GPS Time Transfer', Proc. IEEE, vol 79, No. 7, 1991.

[5] Kirchner D., 'Two-Way Time Transfer via Communication Satellites', Proc. Of the IEEE, vol 79, No. 7, 1991.

[6] Davis J. A., McA Steele J., 'A Caesium Flying Clock Experiment Between NPL and USNO', Proc. 11th EFTF, Neuchatel, this issue.

RUBIDIUM FREQUENCY STANDARD USING THE RADIO NAVIGATION SYSTEMS "GLONASS" AND "NAVSTAR"

A.S.Kleyman

Doctor of Science (Techn.) Manager of The Laboratory
of Time and Frequency Measurement
Kleyman Alexandr Samuilovich SSIA "Metrologi"
fl.133, Kosmycheskaya st.27, Kharkov 310145, Ukraine
tel: 057-2-303-129, 057-2-403-020.

The rubidium frequency and time standard (RFTS) was realized by tne group of the scientists and engineers on tne national element base and it passed the long tests (~ 620 days) in the standard of time and frequency of SSIA "Metrology". The created standard (RFTS-1) has the high precision characteristics and is not inferior to foreing specimens. The root-mean-square value of two-choice dispersion doesn't exceed 2.0ò10 systematic frequency variation doesn't exceed 1.5ò10
However, the insufficient precision of frequency reproducibility (1.0-2.0ò10 of further perfecting rubidium frequency standards. Such a task can be solved using the rubidium frequency standard along with a receiver of radio navigation systems "GLONASS" and "NAVSTAR".
Then such a standard working in the differential mode with the time and frequency standard of SSIA "Metrology" can have precision not less than 5ò10 standard and its correction is made according to the standard of SSIA "Metrology". Preliminary research of such a standard with the receiver A-724 (radio navigation system "GLONASS") showed that the correction apparatus error was 5ò10 per day. Such equipment realization promising is undoubtable, and the preliminary tests show the possibility of realization of a standard with precision which is more than 5ò10

GPS CONTROLLED FREQUENCY AND TIME STANDARD WITH HIGH IMMUNITY FROM TEMPORARY ABSENCE OF 1 PPS SIGNAL.

A. Lisowiec, A. Czarnecki
Tele and Radio Research Institute, Ratuszowa 11, 03-450 Warsaw, Poland.

1. ABSTRACT

GPS controlled frequency and time source with local oscillator specifically designed for best medium and short term stability has been presented. The source enables monitoring of accumulated phase time error with 5 ns resolution in the absence of GPS signal for practically unlimited period of time. The control algorithm is capable of maintaining absolute phase accuracy (i.e. without accumulation of phase time error) in the case of frequent GPS pulse slips as long as the accumulated phase time error doesn't exceed preset value. After very long absence of GPS signal when the leveling of phase time error may be impractical, the accumulated phase time error is zeroed and the control algorithm goes into mode optimized for fast resynchronisation of the local oscillator.

The source can be equipped with three independently synchronized local oscillators that may be switched to the output without phase shifts, which greatly improves reliability.

2. INTRODUCTION

Although GPS controlled sources offer frequency inaccuracy of the order 10^{-12} when averaged over one day period virtually independent of the quality of the local oscillator, the short and medium term frequency instability is determined primarily by the local VCXO. While the low end of the market used TVCXO and medium stability OCVCXO, the instruments that offer highest performance employ rubidium oscillators or some sophisticated and expensive quartz technologies like BVA.

Rubidium oscillators offer better long time behavior than quartz oscillators. However in applications when the long time behavior is controlled by GPS, the advantage of rubidium over very high stability quartz oscillator shows itself only when the GPS controlled source enters hold over mode. Techniques have been developed to bring the medium term stability of quartz oscillators to the level comparable with rubidium.

As the main influence on the quartz oscillator frequency stability over time intervals ranging from 1s to a few hours is temperature, two solutions have been developed. One is to compensate for temperature changes e.g. Smart Clock algorithm of Hewlett-Packard. The other is to develop better oven.

At Tele and Radio Research Institute there has been developed over the last three years a new quartz oscillator with greatly improved oven. The temperature coefficient of the oscillator is of the order of $3x10^{-10}$ over whole temperature range. On the basis of this oscillator there has been developed a new generation of GPS controlled sources.

3. FUNCTIONAL DESCRIPTION

3. 1. Principle of operation

The block diagram of GPS controlled clock set is presented in fig. 1. The set consists of several functional groups. There are from 1 to 3 synchronization units (called channels) each consisting of oscillator module BGn (n=1,2,3) and frequency comparison module KPn (n=1,2,3). The pulses of VCXO frequency multiplied to 200 MHz are counted by two counters in time intervals between successive GPS 1pps pulses. The counters count the pulses alternately with zero dead time [1], which means that none of the pulses are omitted or counted twice. Thus the number of counted pulses after subtracting ideal reference is a measure of the accumulated phase time drift. The deviation of the oscillator's frequency from nominal is determined by averaging the number of pulses counted over loop time interval. The VCXO control voltage consists of two parts, one proportional to frequency deviation averaged over loop time and the other proportional to the slope of the regression line fitting the phase time drift curve over several most recent measurement results. The control algorithm is essentially the same as the one implemented in the older generation of GPS controlled source which has been described in [4].

When locked, the instrument learns the aging rate of VCXO in each channel. As the aging rate of the VCXOs is of the order of $5 \cdot 10^{-11}$ per day a

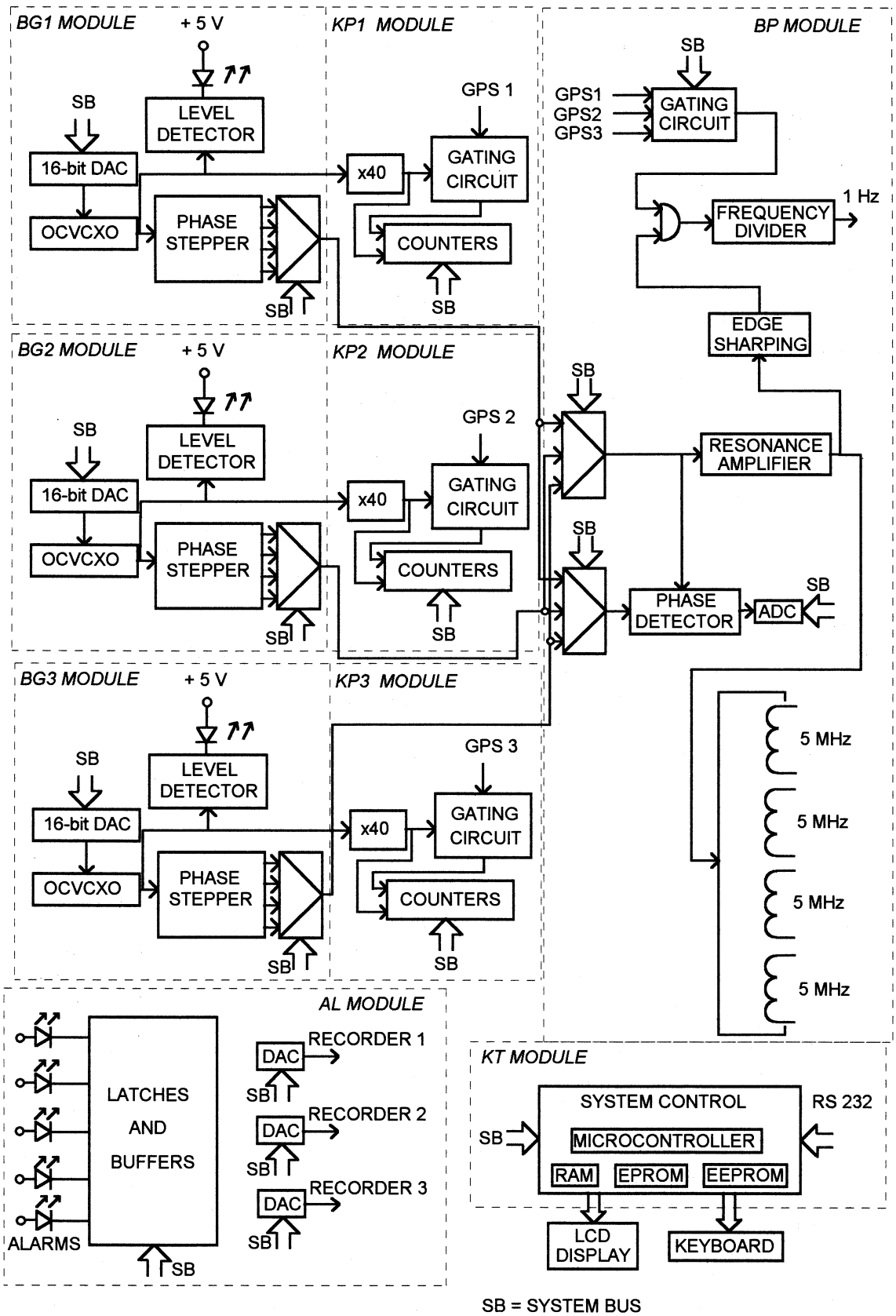


Fig. 1 Block diagram of the clock set

minimum of 10 days of uninterrupted locked in mode operation is required. After the initial 10 days period the aging rate coefficient is updated once per day.

The next functional group is the output signal switching and 1pps output signal generation module BP. To minimize phase transients during switchover from one channel to the other, the output signals of the VCXOs are cyclically phase stepped and their phase compared with the phase of the output signal. The microcontroller stores the state of the phase stepper in each channel that minimizes phase difference with output signal. The TTL 1pps output signal of duration 100 ns is generated by synchronically dividing the frequency of sinusoidal output signal.

The control of the clock set is centralized in the microcontroller unit KT. The interface with the user is via 80 character LCD display and a numeric keyboard. The set can work as an autonomous instrument or can be fully remotely controlled via RS-232 serial interface by an PC IBM computer. All parameters of the control loop can be changed and the user can implement its own control algorithm.

3. 2. Hold-over mode

In the absence of GPS signal the instrument enters hold-over mode in which the control voltage is incremented every hour according to the aging rate coefficient and slope of the voltage/frequency characteristic of the VCXO. Linear drift in time is assumed. It has been confirmed experimentally that frequency accuracy after ten days of operation in hold over mode was better then $2 \cdot 10^{-10}$. With LFE resonator option it is hoped to come down even below 10^{-10} [2], [3], [5].

Additionally the GPS controlled clock set enables monitoring of phase time error (with accuracy of 5 ns) even in the absence of GPS pulses. It works as follows.

As the GPS 1 pps pulses switch the counting of multiplied VCXO frequency pulses between two counters, in their absence the counter that has last been active eventually overflows. When this overflow is detected by monitoring the msb bit of the counter the control algorithm enters the loop shown in fig. 2 with initial value of R set to zero. This loop has the special form that is dictated first by limited computational speed of the 8 bit microcontroller and second by the length of the registers used in the computation process as all calculations have to be done in integer arithmetic so as not to lose accuracy.

First GPS pulse that arrives stops the active counter and starts the other one. The phase time error is computed as $(N + R - 2 \cdot 10^8) \cdot 5$ [ns] where N is the number of pulses counted by the counter that has been stopped.

The monitoring of phase time in the absence of GPS pulses can be done indefinitely. It is up to the software what to do when the GPS signal appears again and the accumulated phase time error is large. A drift of $3 \cdot 10^{-11}$ leads to phase time error equal to 2.6 μs after one day period. It is impractical although possible for control algorithm to level out such phase time error as this would lead to considerable temporary frequency deviation.

The control algorithm has the fast tuning option in which the accuracy of τ/T where τ is RMS GPS 1 pps phase jitter and T is loop time constant is attained in T seconds independent of initial VCXO inaccuracy. With τ = 100 ns the accuracy of $5 \cdot 10^{-11}$ is attained in about half an hour.

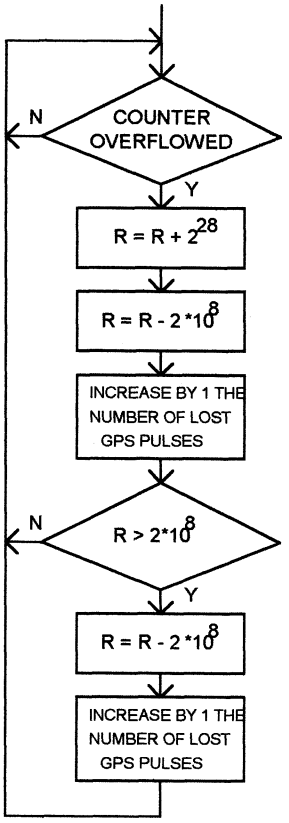


Fig. 2. Phase time monitoring algorithm in the absence of GPS signal.

4. CONCLUSION

Very high stability local VCXO of 2001 type enables to obtain exceptionally good clock set parameters when locked in. With loop time constant greater than 1 hour medium term frequency fluctuations (over time period from 1 s to 1 hour) don't exceed $2 \cdot 10^{-11}$. Temperature effects are virtually unobservable. Using the clock set as frequency reference it was possible to

observe the cesium source inaccuracy of about $2 \cdot 10^{-12}$ in test setup shown in fig. 3 within 10 hour period.

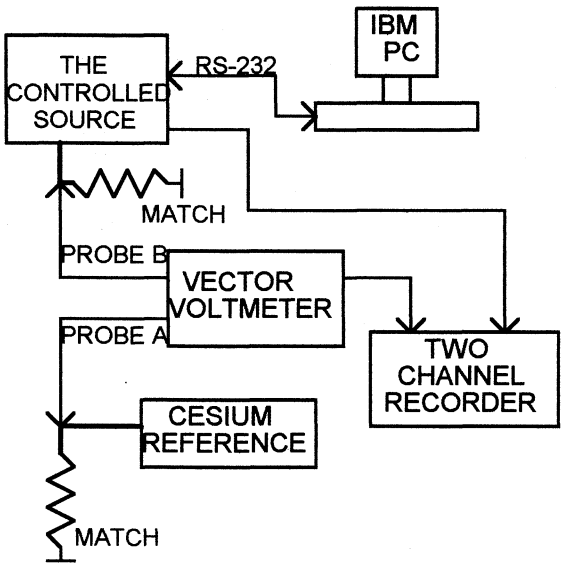


Fig. 3. Measurement setup.

As the described above model of GPS controlled clock set needs an external GPS receiver of no specific type, special measures have also been taken to make the control algorithm as immune as possible to various faults in 1 pps signal.

There is a need in telecommunication services for field type frequency reference source. One of the main requirements on such source is as short as possible synchronization time to within a few times 10^{-11} accuracy. Time to first fix for typical GPS receiver with no data known (a situation when the receiver is carried to a distant location) is about 15 minutes. On the other hand time needed for initially cold 2001 type oscillator to attain 10^{-8} accuracy (minimum accuracy to start automatic fine frequency tuning) is of the order of 1 hour. Adding to it a minimum of half an hour of synchronization time makes an unacceptable long time period for the source to become operational. To overcome this problem a battery which keeps the oscillators in warm up state is offered as an option.

5. REFERENCES

1. Lisowiec A., Czarnecki A., 1994, „Digital time comparator of 1ns resolution for 10 s observation time", 8th Piezoelectric Conference PIEZO'94 Proc.
2. Masiukiewicz A., Gniewinska B., Szulc W., 1994, „Application of LFE-SC cut resonators at high stability oscillators, Proc.of IEEE FCS, Boston, USA, 1994
3. Masiukiewicz A., Gniewinska B., Szulc W., Weiss K., „Study of SC - cut lateral field excited resonators features", Proc. of 9th EFTF, Besancon, 1995
4. Lisowiec A., Czarnecki A., Rau Z., „A method for quartz oscillator synchronization by GPS signal", Proc. of 10th EFTF, Brighton, 1996
5. Masiukiewicz A., Gniewinska B., Kalinowska B., „Analysis of low frequency spectrum of OCXO oscillator employing LFE resonator", Proc. of PIEZO'96, Waplewo, Poland, 1996.

DESIGN OF A HYDROGEN MASER FOR SPACE

Edward M. Mattison and Robert F. C. Vessot

Smithsonian Astrophysical Observatory
Cambridge Massachusetts 02138, USA

1. ABSTRACT

An active atomic hydrogen maser for long-term use in space has been designed and built as part of the Smithsonian Astrophysical Observatory's Hydrogen Maser Clock (HMC) project. We describe aspects of the maser's mechanical, magnetic and thermal design that are important to its performance in space. The flight hardware has been tested in a laboratory thermal/vacuum chamber; we report performance measurements of the magnetic and thermal control systems.

Keywords: hydrogen maser, atomic clock, thermal control

2. CLOCKS IN SPACE; THE HYDROGEN MASER CLOCK PROGRAM

Frequency references – high stability clocks – increasingly find applications in space missions. Atomic clocks of ever increasing stability have present and potential uses as frequency references for the GLONASS and Global Positioning System navigation systems, local oscillators for space-based Very Long Baseline Interferometry, "proper" clocks for tests of general relativity, frequency references for detection of gravitational radiation, and "traveling clocks" for worldwide time transfer.

Clocks for use in space must satisfy several restrictions and requirements, many of which are also requirements or desirable features of earth-based clocks. These requirements include:

- Limitations on mass, size and power
- Requirements for reliable long-term unattended operation
- Ability to withstand vibrational loads during launch
- Ability to tolerate varying magnetic fields
- Ability to cope with a varying thermal environment

An active atomic hydrogen maser for long-term use in space has been designed and built as part of the Smithsonian Astrophysical Observatory's Hydrogen Maser Clock (HMC) project. HMC is a NASA-sponsored program with the goal of producing and demonstrating a space-qualified hydrogen maser with drift-removed fractional frequency stability of 10^{-15} or better in one day. The HMC maser is an evolutionary outgrowth of a two-decade long SAO program of research and development of hydrogen masers for earth and space use.[1,2,3] The maser and its control electronics have been designed as an integrated system to cope with the requirements of space flight. We discuss below characteristics of its mechanical, magnetic and thermal design that are particularly relevant to use in space.

The HMC maser is designed for use with a variety of spacecraft, requiring only an appropriate mechanical connection and electrical interface. It was originally to be tested aboard the European Space Agency's Eureca spacecraft, and then, following cancellation of the planned Eureca reflight, on the Russian Mir space station[4]. At present the flight portion of the HMC program has been terminated, and the flight model maser and its electronics are undergoing laboratory testing at SAO.

3. MECHANICAL AND STRUCTURAL CHARACTERISTICS

The HMC maser's physics unit, shown in in cross-section in Figures 1 and 2, takes the general form of a cylinder with 84 cm long and 43 cm in diameter. The maser's main components are its quartz storage bulb and low-expansion resonant cavity; the titanium vacuum tank that contains the cavity; a stainless steel vacuum manifold that includes two sorption pumps for scavenging hydrogen and two small ion pumps for removing other gases; a $LiAlH_4$ hydrogen source and a glass dissociator chamber for producing a beam of hydrogen atoms; electrical heaters, insulation and thermistors for temperature control; and magnetic shields and solenoids for magnetic field control. In addition, the physics unit contains electronic components that amplify the 1420 MHz maser signal from the cavity and electrically isolate the cavity from external perturbations. Separate units contain analog and digital control and monitoring electronics, the R.F. receiver that phase-locks a 100 MHz crystal oscillator to the maser signal, and a microprocessor that controls the maser's electronics and acts as an interface with the spacecraft's data and telecommand system. The masses of the major instrument elements are given in Table 1. Additional elements, whose masses depend upon the specific spacecraft used, are the bracket that mounts the maser to the spacecraft, and any additional spacecraft-specific electronics.

Table 1. HMC Instrument Mass Summary

Element	Mass (kg)
Maser physics unit	70.7
Control and RF electronics	27.9

Structurally, the maser is supported from a circular aluminum midplane plate, which supports the maser's resonant cavity and vacuum tank on one side, and its vacuum manifold and hydrogen source on the other. The midplane plate is the main structure for mounting the maser to the spacecraft. A titanium "aft neck" tube connects one end of the vacuum tank to the midplane plate and the vacuum manifold, while a similar "forward neck" connects the other end of the vacuum tank to the maser's cylindrical outer aluminum housing. The housing, in turn, transfers the forward neck's load to the midplane plate. By means of an ANSYS finite element model with approximately 2800 nodes, the HMC maser has been designed to cope with the vibrational and accelerational loads of a Space Shuttle launch. It can withstand at least 15 g's r.m.s., in all axes acting simultaneously, in a spectrum from 20 Hz to 2 kHz. The maser's lowest mechanical resonant frequency is 46 Hz. The flight cavity and vacuum tank, which are the most critical components, have been tested to flight input vibrational levels.

4. MAGNETIC FIELD CONTROL

A spacecraft in low earth orbit experiences the earth's magnetic field, with a magnitude of about 0.5 gauss and a variation over an orbit of up to ±0.5 gauss, depending upon the spacecraft's attitude in orbit. In addition, some spacecraft create variable magnetic fields themselves, for example by magnetic torquers used for attitude control. The magnetic field within the maser's storage bulb must be maintained at a level on the order of 0.3 milligauss. To achieve frequency stability of better than $\Delta f/f<1\times10^{-15}$, the temporal variation of the internal magnetic field must be less than $\Delta H<0.8\times10^{-6}$ gauss. To achieve these conditions the HMC maser utilizes passive magnetic shields, internal solenoids and an active magnetic compensation system.

As shown in Figure 1, the maser's resonant cavity and titanium vacuum tank are surrounded by a three-section, two-layer cylindrical printed circuit solenoid that creates the internal magnetic field of approximately 0.3 milligauss, and by four layers of concentric magnetic shields that attenuate external fields. The outermost shield extends to enclose the vacuum pump manifold and atomic hydrogen dissociator, reducing external fields that could perturb the state-selected atomic hydrogen beam. The measured shielding factor of these Hypernom shields is

$$S_{passive} = \frac{\Delta H_{ext}}{\Delta H_{int}} \approx 3.4 \times 10^5$$

The passive shields are augmented by an active magnetic compensation system. A single-axis fluxgate magnetometer sensor is mounted inside the outer shield to sense the axial field near the end of the maser. A compensation coil is wound on the outside cylindrical surface of the next shield, and a feedback circuit drives the coil to keep the field sensed by the magnetometer constant. The shielding factor for the total magnetic control system, determined by measuring the transverse ("Zeeman") resonance frequency in the oscillating maser' storage bulb, is

$$S_{total} \approx 2.8 \times 10^6$$

With this shielding factor, the expected maximum fractional frequency variation due to movement through the earth's field is on the order of $\Delta f/f \sim 2\times10^{-16}$.

5. THERMAL CONTROL SYSTEM DESIGN FEATURES

Temperature changes of the maser's resonant cavity and storage bulb affect the maser's output frequency. To keep frequency variations below the level of 1 part in 10^{15}, the cavity temperature must be maintained constant to approximately 10^{-4} °C. The HMC maser employs several strategies to achieve this level of temperature control. To control heat flow from the vacuum tank, the maser's structure is divided into three concentric isothermal control regions. Thermal gradients are controlled by subdividing each isothermal region into multiple independently controlled zones, by mounting controlled guard heaters on heat leakage paths, by separating heaters from the primary controlled structure (the vacuum tank) and by carefully calibrating and matching thermistors and setpoint resistors to ensure that all zones of an isothermal region control at the same temperature. Radiative heat flow is reduced by means of multilayer insulation in the spaces between the regions, which are evacuated by being open to the space environment, while conductive heat flow is controlled by design of the segmented nylon rings that support the magnetic shields.

As shown in Figure 2, the innermost isothermal region, which is the titanium vacuum tank that surrounds the resonant cavity, is maintained at 50°C. The resolution of the tank control system is 1×10^{-4} degrees. To reduce thermal gradients in the tank, the three tank heaters are separate from the tank itself, one being located on the outside surface of the inner magnetic shield that is directly outside the tank and the others on the titanium neck tubes where they connect to either end of the tank.

The tank, in turn, is surrounded by an aluminum oven that is located directly over the third magnetic shield and whose temperature is maintained at 41°C. The oven region acts as a guard to control heat that flows from the tank region both radiatively from the tank surface and conductively along the magnetic shield supports and the titanium support necks. The oven region consists of three control zones located on the cylinder and end surfaces of the oven, and two zones mounted on the outer ends of the support necks.

The third isothermal region consists of the midplane plate and an outer aluminum support shell that directly surrounds the fourth magnetic shield. This zone is maintained at approximately 27°C by a control thermistor and set of heaters mounted on the midplane plate.

In addition to the thermal control zones that are integral with the maser, the system includes a controlled temperature guard station on the structure that mounts the maser to the spacecraft, to act as a first stage of isolation from the conductive environment. The entire instrument is surrounded with multilayer insulation to isolate it from the radiative environment.

The thermal control system incorporates several electronic and hardware features to achieve the high degree of thermal stability required. The digital electronic control system is based upon four 68HC11 microcontrollers, each of which can control up to five thermal zones. Each 68HC11 includes a microprocessor, an 8-bit analog-to-digital converter with eight-channel multiplexer, and timer registers that are used as pulse-width modulators (PWM) for high-efficiency switched heater power control. The vacuum tank heaters, which are closest to the maser's resonant cavity, are powered by high-frequency (~8 kHz) PWMs to avoid perturbation of the maser oscillation; the other heaters are switched at a 6 Hz rate. The thermal control program incorporates a three-mode PID (proportional, integral and differential) algorithm to eliminate proportional offset. (Differential control is included in the algorithm, but has not been found to be useful in this application.)

Components of the thermal control system have been chosen for thermal stability and low magnetic field production. Thermistors are glass-encapsulated, high stability units that have been burned in. Monitor and control thermistors for each zone are chosen to be matched. Temperature setpoint resistors are chosen to have low temperature coefficients, and are physically mounted on a temperature controlled zone within the maser for minimum temperature perturbation. Heaters are flexible printed circuits with Kapton film insulation. For each heater identical

etched foil elements are overlayed with opposite current flow, to minimize magnetic field production.

The ability of the thermal control system to stabilize the tank zone temperatures in the face of external temperature changes is shown by the data of Table 2. For these measurements, which were made on the engineering model of the maser, the temperatures of the maser support structure and the forward neck guard zone were separately lowered by 2°C.

Table 2. Response of tank control zones to external temperature change

	ΔT(Support) = −2 °C	ΔT(Fwd neck) = −2 °C
ΔT(Tank forward):	-0.1×10^{-4} °C	$+0.2 \times 10^{-4}$ °C
ΔT(Tank cylinder):	$+1 \times 10^{-4}$ °C	-1×10^{-4} °C
ΔT(Tank aft):	$+3 \times 10^{-4}$ °C	-2×10^{-4} °C

6. PRELIMINARY TEST RESULTS

The maser physics package has been operated in a thermal-vacuum tank continuously since August 1996. Analysis of the data under varying degrees of thermal perturbation is still in process and will be reported in a future journal. From a rough evaluation we can say that the immunity of the maser to thermal and magnetic environmental changes is well within specifications.

A plot of the measured frequency stability data is shown in Figure3.

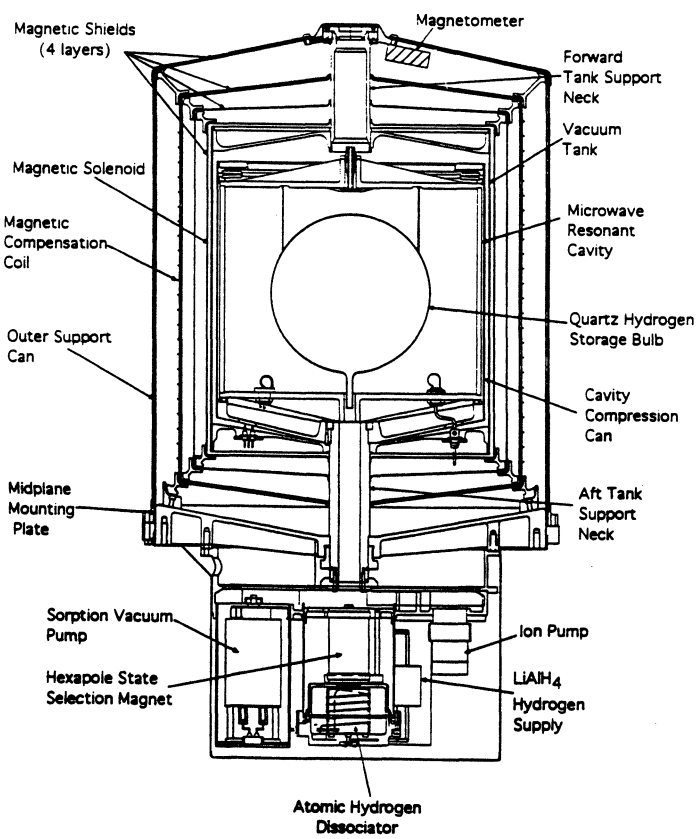


Figure 1. HMC maser - major components

6. REFERENCES

1. "A hydrogen maser design for ground applications." M.W. Levine, R.F.C. Vessot, E.M. Mattison, E. Blomberg, T.E. Hoffman, G. Nystrom, D.F. Graveline, R.L. Nicoll, C. Dovidio and W. Brymer, *Proc. Eighth Annual Precise Time and Time Interval Applications and Planning Meeting*, Naval Research Laboratory, Washington, D.C., pp. 249–276 (1976).

"A new generation of SAO hydrogen masers." M.W. Levine, R.F.C. Vessot, E.M. Mattison, G. Nystrom, T.E. Hoffman, and E. Blomberg, *Proc. 31st Annual Frequency Control Symposium*, USAERADCOM, Ft. Monmouth, N.J. 07703, pp. 525–534 (1977).

2. "Test of relativistic gravitation with a space-borne hydrogen maser." R.F.C. Vessot, M.W. Levine, E.M. Mattison, E.L. Blomberg, T.E. Hoffman, G.U. Nystrom, B.F. Farrell, R. Decher, P.B. Eby, C.R. Baugher, J.W. Watts, D.L. Teubr, and F.D. Wills, *Phys. Rev. Lett.* 45, 2081 (1980).

3. "Ultra-stable clocks for use in space." E.M. Mattison, *Advances in Space Research* 9, pp. (9)13–(9)19 (1989)

4. "High precision time transfer to test a hydrogen maser on Mir." E.M. Mattison and R.F.C. Vessot, *Proc. 27th Annual Precise Time and Time Interval Applications and Planning Meeting* (San Diego, CA, 25 Nov-1 Dec 1995) (NASA Conference Publication 3334), pp. 181-192 (1995).

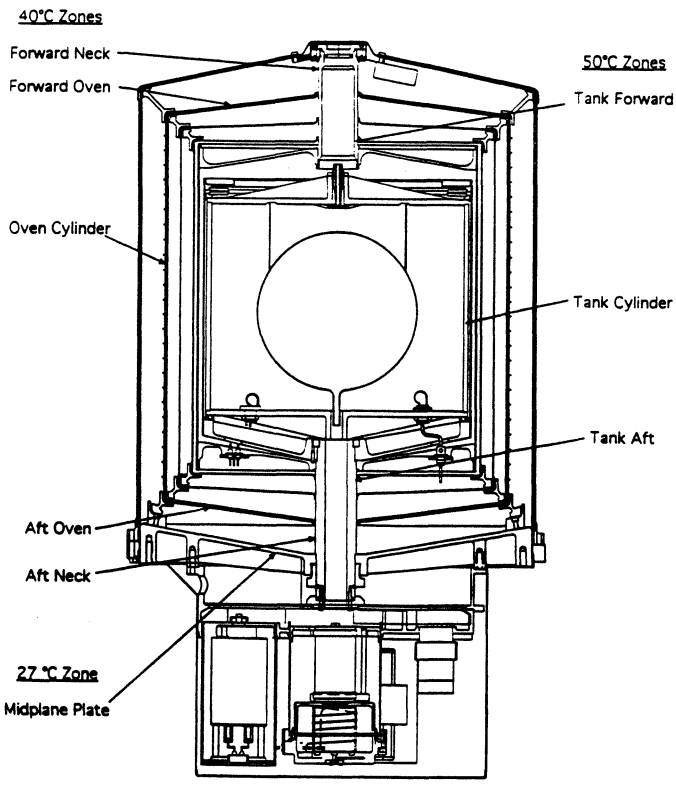


Figure 2. Thermal Control Zone Locations

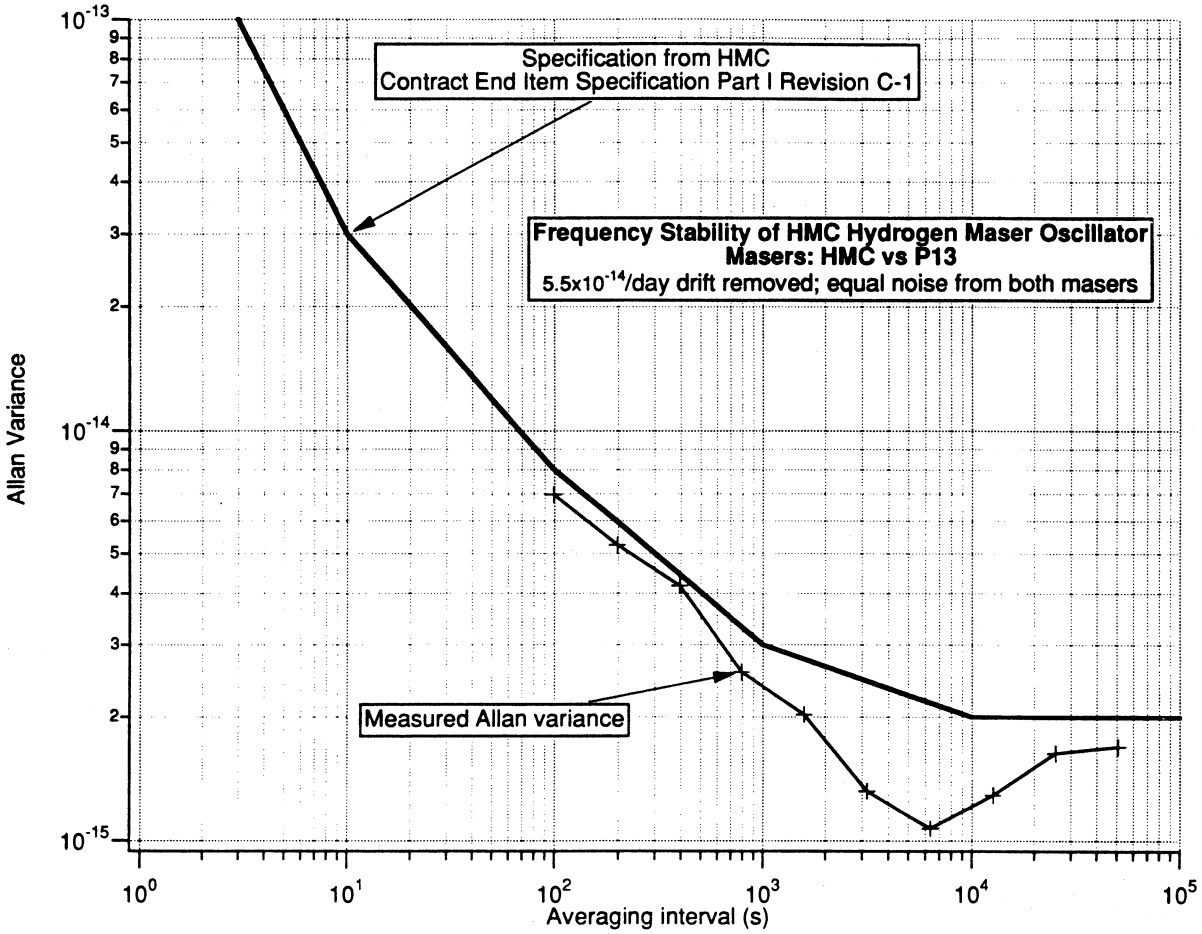


Figure 3 Frequency Stability Data of HMC (1996)

Influence of the ionospheric refraction on time transfer using GPS.

René WARNANT
Royal Observatory of Belgium
Avenue Circulaire, 3
B-1180 Brussels
Belgium

The time transfer using GPS is affected by the ionopheric refraction. This error is usually taken into account by correcting GPS observations with the Klobuchar model. In addition, differences between observations collected in neighbouring stations are made; one assumes that the residual ionospheric error remaining in the differences is negligible.

The paper presents some results of a PH. D. thesis dedicated to the study of the ionospheric refraction effect on GPS measurements. This effect mainly depends on the ionospheric Total Electron Content (TEC). At the Royal Observatory of Belgium, a method has been developed in order to compute the Total Electron Content using a combination of GPS dual frequency measurements. This method has been applied to the large GPS data bank accumulated at Brussels: the Royal Observatory has participated in many international GPS campaigns from May 89 to August 92; a permanent GPS station is in operation at Brussels since April 93.

This study clearly shows that even in a mid-latitude "quiet" station as Brussels, the TEC has a very irregular behaviour. For this reason, the use of the Klobuchar model often give rise to an important ionospheric residual error. In addition, the occurrence of Travelling Ionospheric Disturbances and scintillations effects have also been detected using GPS measurements. The paper shows that these ionospheric phenomena often cause large gradients in the Total Electron Content even between neighbouring stations. On the other hand, Travelling Ionospheric Disturbances usually appear at specific moments in the day and in the year. The statistics of occurrence of these disturbances could be taken into account to plan the sessions used in GPS time transfer.

The solar minimum will probably be reached before the end of 1996. During the next years, the solar activity will become more important giving rise to higher ionospheric perturbations. For this reason, the procedures used in GPS time transfer should take into account the ionospheric disturbances presented in this paper.

INFLUENCE OF THE LOCAL OSCILLATOR FREQUENCY FLUCTUATIONS ON THE MEASUREMENTS WITH DMTD SYSTEMS

Mihai P. Dinca

University of Bucharest, Department of Physics, PO BOX MG-11, Bucharest-Magurele, 76900, ROMANIA,
e-mail mpdi@fizica.fizica.unibuc.ro

ABSTRACT

Dual Mixer Time Difference system is frequently used for comparing frequency standards. It is generally known that, for long averaging time, the frequency fluctuations of the local oscillator can be neglected, provided that this oscillator is of the same quality as the compared ones. In this paper, the contribution of each system component to the result of the measurement is analyzed. Weight function relating the local oscillator frequency deviations and its corresponding term in the measurement result is calculated and a factor characterizing the local oscillator contribution, in terms of Allan variance, is introduced. Then, this factor is numerically computed for several kinds of noise, taking in consideration the pass-band of the measuring system, and the conditions for neglecting the local oscillator contribution are obtained.

Keywords: frequency standards, frequency stability, DMTD systems, local oscillator

1. INTRODUCTION

In comparing frequency standards the Dual Mixer Time Difference (DMTD) system (Ref. 1) offers the advantage to measure the time difference between oscillators, from which time fluctuations, frequency and frequency fluctuations can be deduced. The system structure is sketched in Fig. 1. Signals from the compared oscillators, assumed to have the same frequency ν_0, are mixed with the local oscillator signal which has the frequency slightly shifted, by ν_b. The beat signals are amplified and shaped and their zero crossing moments are detected. Own to variable phase shifter, these signals are slightly time shifted and the counter measures their time difference. The results are random spread around an average value Δt_{av} and from their analysis the frequency stability can be deduced. The contribution of the system itself (system noise) can be simply measured by driving the both system inputs with the same signal.

2. PROCESSING OF THE SYSTEM OUTPUT DATA

We shall denote by $\varphi_{o1}(t)$, $\varphi_{o2}(t)$ and $\varphi_{LO}(t)$ the phase fluctuations of the compared oscillators and local oscillator, respectively. On channel 1 the isolation amplifier produces the phase fluctuations $\varphi_{a1}(t)$ while the mixer and following processing chain introduce the phase fluctuations $\varphi_{f1}(t)$. If the local oscillator frequency is less than ν_0, the phase fluctuations for the processed signal, whose zero crossing moments are detected, can be expressed as

$$\begin{aligned}\varphi_{b1}(t) &= \varphi_{o1}(t)-\varphi_{a1}(t)+\varphi_{f1}(t)-\varphi_{LO}(t)\\ &= \varphi_1(t)-\varphi_{LO}(t)\end{aligned} \quad (1)$$

A similar equation holds for the other channel

$$\begin{aligned}\varphi_{b2}(t) &= \varphi_{o2}(t)-\varphi_{a2}(t)+\varphi_{f2}(t)-\varphi_{LO}(t)\\ &= \varphi_2(t)-\varphi_{LO}(t)\end{aligned} \quad (2)$$

All terms in $\varphi_1(t)$ and $\varphi_2(t)$ are mutually uncorrelated and uncorrelated with $\varphi_{LO}(t)$. The fluctuations Δt_k of the time difference between the beat signals have the form

$$\Delta t_k = \frac{1}{2\pi\nu_b}\left[\varphi_{b1}({}^1t_k)-\varphi_{b2}({}^2t_k)\right] \quad (3)$$

where ν_b is the beat frequency. If Δt_{av} is negligible compared against the beat period, the last equation becomes

$$\Delta t_k \cong \frac{1}{2\pi\nu_b}\left[\varphi_{b1}(t_k)-\varphi_{b2}(t_k)\right] \quad (4)$$

and, using (1-2), the canceling of the local oscillator contribution results.

In order to measure the Allan variance (Ref. 2) of the oscillator under test, defined by

$$\sigma^2_{y1}(\tau) = \frac{1}{2(2\pi\nu_0\tau)^2}\left\langle\left[\varphi_{o1}(t+2\tau)-2\varphi_{o1}(t+\tau)+\varphi_{o1}(t)\right]^2\right\rangle \quad (5)$$

an adequate statistical processing is implemented, namely

$$\frac{1}{2\tau^2}\left(\frac{\nu_b}{\nu_0}\right)^2\left\langle\left[\Delta t_{k+2m}-2\Delta t_{k+m}+\Delta t_k\right]^2\right\rangle = \sigma^2_{y1}(\tau)+\sigma^2_{y2}(\tau) \quad (6)$$

where $\tau = m/(2\pi\nu_b)$ with m integer. The previous equation can be interpreted as the average power of the random process

$$z(t) = \frac{\nu_b}{\nu_0}\left(y_{o1}(t) - y_{o1}(t)\right) \otimes h_A(t) \qquad (7)$$

where $y_{o1}(t)$ and $y_{o2}(t)$ are the fractional frequency fluctuations, $h_A(t)$ is the Allan variance weight function and symbol $\otimes$ denotes the convolution product.

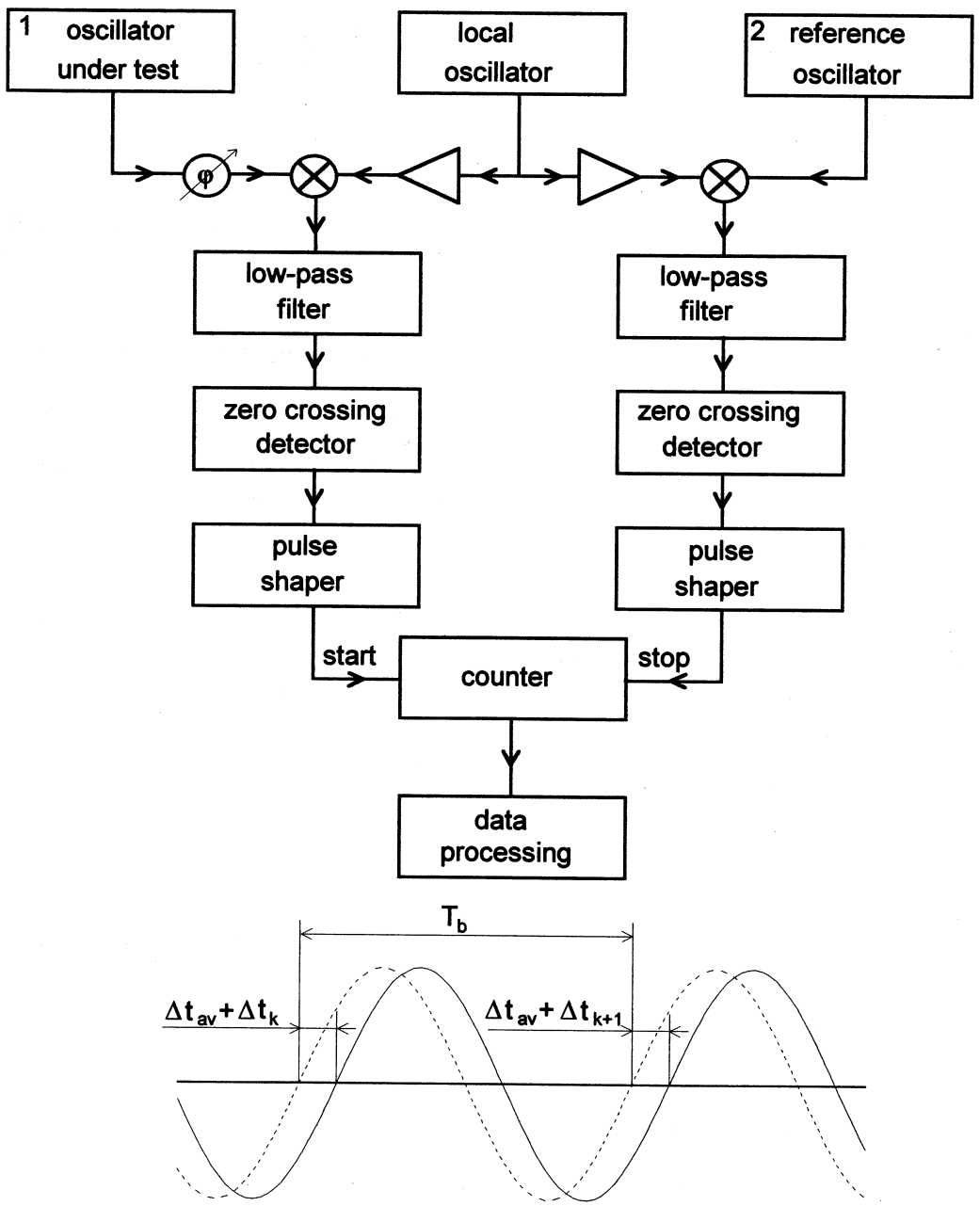


Fig. 1. Block diagram of DMTD system.

3. EXPRESSION OF THE MEASUREMENT RESULT

If the same data processing algorithm is used, but none of the above approximations are made, the measurement result is the average power of the process

$$z(t) = \left[\frac{\nu_b}{\nu_0}\left(y_{o1}(t) - y_{a1}(t)\right) + y_{f1}(t)\right] \otimes h_1(t)$$
$$- \left[\frac{\nu_b}{\nu_0}\left(y_{o2}(t) - y_{a2}(t)\right) + y_{f2}(t)\right] \otimes h_2(t)$$
$$+ \frac{\nu_b}{\nu_0} y_{LO}(t) \otimes h_3(t) \qquad (8)$$

where the y quantities are fractional frequency fluctuations and the subscripts have the meaning previously given. Here, all fluctuation terms are mutually uncorrelated. The weight functions $h_1 - h_3$ are depicted in Fig. 2.

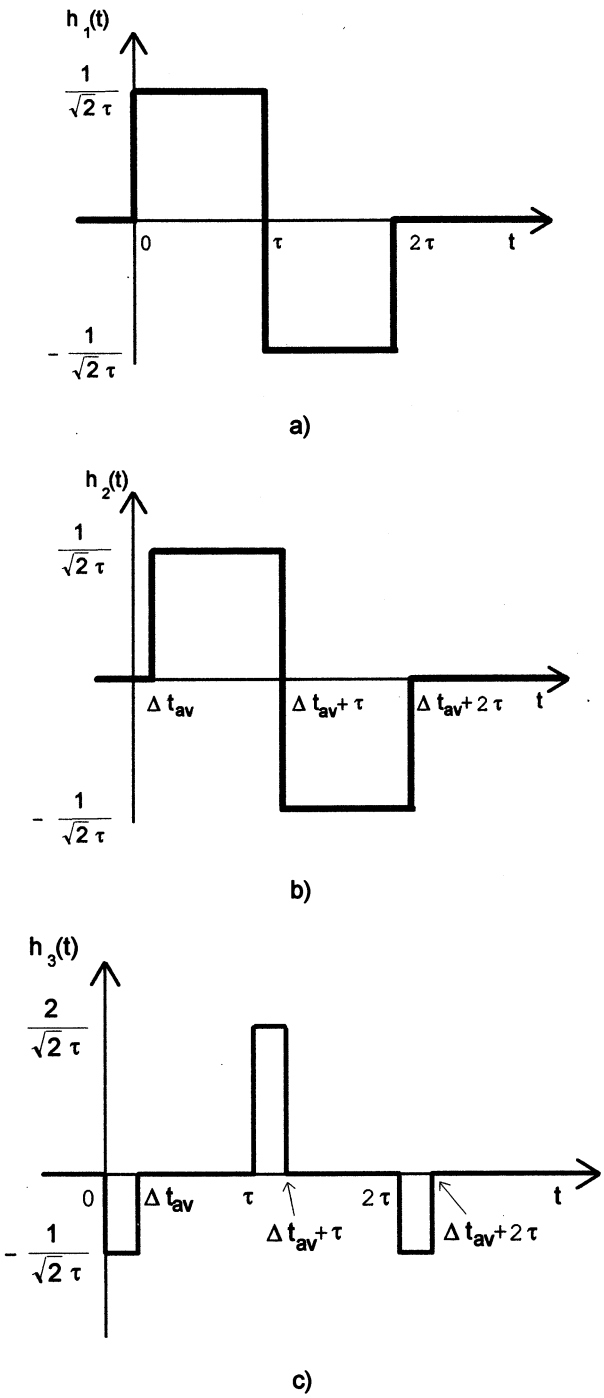


Fig. 2. Weight functions h_1-h_3.

The measurement result can be expressed in terms of Allan variances for all perturbing processes

$$\left\langle z^2(t)\right\rangle = \left(\frac{\nu_b}{\nu_0}\right)^2 \left[\sigma_{o1}^2 + \sigma_{o2}^2 + B\sigma_{LO}^2 + \sigma_{ya1}^2 + \sigma_{ya2}^2\right] + \sigma_{yf1}^2 + \sigma_{yf2}^2 \tag{9}$$

by introducing a factor B which characterizes the influence of the local oscillator

$$\left\langle z_{LO}^2(t)\right\rangle = \left(\frac{\nu_b}{\nu_0}\right)^2 B\,\sigma_{yLO}^2(\tau) \quad . \tag{10}$$

The contribution of the terms corresponding to isolation amplifiers and processing chains was studied elsewhere (Ref. 3).

4. INFLUENCE OF THE LOCAL OSCILLATOR

As the local oscillator has performances similar with those of the compared oscillators, its contribution on the measurement result can be neglected if the factor B is much less than unity.

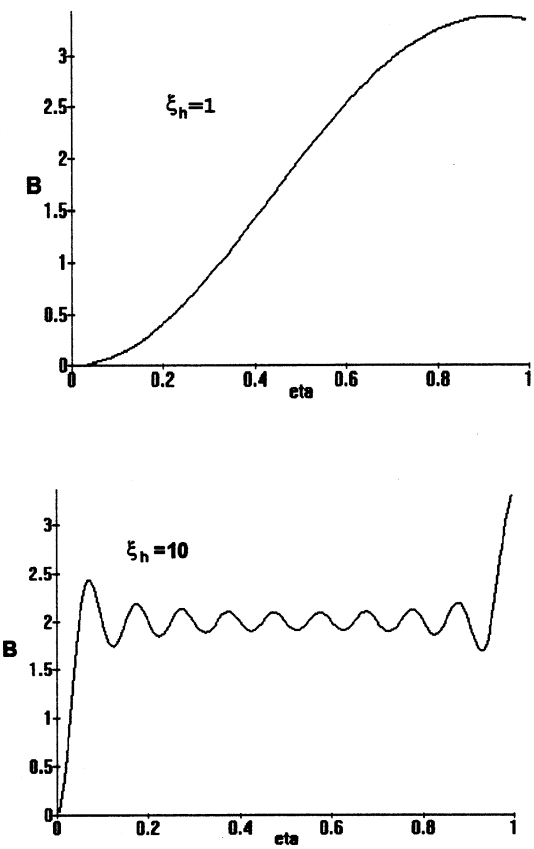


Fig. 3 a) and b). $B(\eta)$ dependence for several ξ_h values

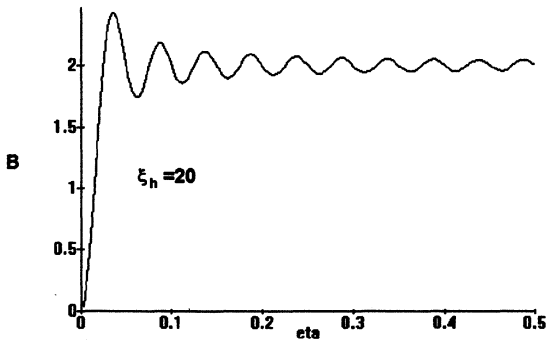


Fig. 3. c) $B(\eta)$ dependence for several ξ_h values.

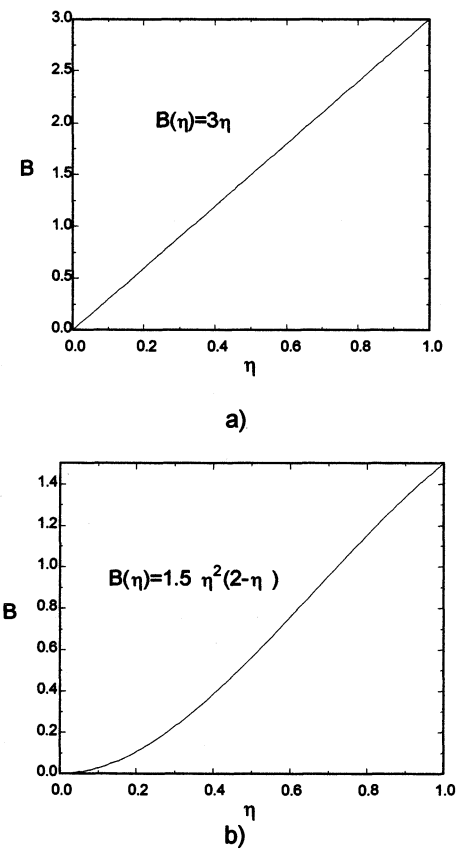


Fig. 4. $B(\eta)$ dependence for white frequency (a) and random walk frequency (b) noises.

The factor value can be computed from the power spectral density of $y_{LO}(t)$ by using the transfer functions corresponding to Allan variance and weight function $h_3(t)$. With a spectral density of the form f^{α} one obtains

$$B = \frac{4 \cdot \int_0^{f_h\tau} \xi^{\alpha-2} \sin^4(\pi\xi)\sin^2(\pi\eta\xi)\,d\xi}{\int_0^{f_h\tau} \xi^{\alpha-2} \sin^4(\pi\xi)\,d\xi} \qquad (11)$$

where f_h is the upper cut frequency of the measuring system (considered as ideal low-pass filter) and $\eta = \Delta t_{av}/\tau$.

For white phase noise, the dependence $B(\eta)$ is significantly affected by the value of $\xi_h = f_h\tau$. In Fig. 3 is presented this dependence for ξ_h=1, 10 and 20. One can conclude that, for this kind of noise, B is negligible when $\Delta t_{av} << 1/f_h$.

In the case when the Allan variance does not depend on the system pass-band, the reducing of B requires more relaxed conditions. Thus, for white frequency and random walk frequency noises, the dependence $B(\eta)$ depicted in Fig. 4, shows that if $\eta << 1$ the factor B can be neglected. This condition is equivalent with $\Delta t_{av} << m/(2\pi\nu_b)$, which is automatically fulfilled for large m values.

5. CONCLUSIONS

For DMTD system, the contribution of each component to the overall phase fluctuations measured at the system output was calculated in terms of Allan variance. A factor which characterizes the influence of the local oscillator was introduced and its value was computed for different kind of noises. The results emphasized that the requirements for neglecting this influence depends on noise kind. While for white phase noise the average time difference between the beat signals must be much less than the inverse of the system pass-band, for the noise whose Allan variance does not depend upon the pass-band the average time difference must be much lower than τ.

REFERENCES

1. D. W. Allan and H. Daams, "Picosecond time difference measurement system", Proc. 29th Annual Frequency Control Symposium, May 1975, pp. 404-411.
2. D. W. Allan, "Statistics of atomic frequency standards", Proc. IEEE, vol. 54, no. 2, pp. 221-230, Feb. 1966.
3. M. P. Dinca, "Low frequency noise reduction in DMTD systems", Proc. 9th European Frequency and Time Forum, Besancon, 1995, pp. 369-373.

A NONLINEAR APPROACH OF THE SPECTRUM AND STABILITY OF FREQUENCY LOCKINGS

S. DOS SANTOS, J. F. MOUGIN, F. LARDET-VIEUDRIN, J. GROSLAMBERT and M. PLANAT.
Laboratoire de Physique et Métrologie des Oscillateurs
associé à l'Université de Franche Comté
32, Av. de l'Observatoire, 25044 BESANÇON CEDEX, FRANCE

ABSTRACT

A large signal analysis of frequency mixing and frequency lockings based on the use of continuous fractions is developed. It is applied to a frequency locked loop and to an injection locked pulse oscillator. A new type of frequency instability due to the bad approximations of some quadratic algebraic frequency ratios is identified.

Keywords : frequency instability, arithmetic, non linearity, synchronization.

1. INTRODUCTION

Modelocking setups are widely used for frequency transfer or frequency stability analysis. In those devices nonlinearity operates the mixing of frequencies and the coherent coupling between themselves.

In this paper we study two practical modelocking schemes. Our first set up involves the common frequency locked loop. The reference signal and the voltage controlled oscillator (VCO) may be slaved thanks to the non linear action of the mixer. We study the mixing of frequencies in the open loop by using an arithmetical approach involving continous fractions associated to the reference/VCO frequency ratio. To account for the frequency lockings a discrete nonlinear phase mapping (Arnold map) associated to the closed loop is derived.

In the second set up we study a regenerative pulse delayed oscillator in which a carrier-to-envelope synchronization occurs. In such a device a carrier signal is split in wavepackets thanks to the action of a nonlinear GaAs modulator and slaved to a modulation frequency controlled in a loop using a surface acoustic wave delay line (SAW) and an envelope detector. Frequency fluctuations are controlled to a large extend from the fine structure of frequency ratios on the frequency-amplitude characteristic.

Finally we identify a source of frequency instabilities which is of an arithmetic origin, that is the error signal shows peaks at quadratic irrationalities like $(\sqrt{5}-1)/2$, $\sqrt{2}-1$, $(\sqrt{221}-11)/10$. This is explained thanks to the use of advanced mathematical techniques involving the hyperbolic group PSL(2,Z).

2. THE INTERMODULATION SPECTRUM IN THE OPEN LOOP

Frequency mixing is widely used in communication receivers, radars and plasma physics. In addition to applications such as signal processing and low noise FM demodulators it is often associated to a frequency-feedback loop or a phase locked loop (PLL). A large signal analysis of the device is very difficult to perform due to the significant number of possible intermodulation products and to the spurious signals being generated in the mixer or in others parts of the loop [1].

In essence a mixer should provide up or down conversion $\omega_0 \pm \omega_1$ from a carrier frequency ω_0 and a pump oscillator ω_1 . In PLL applications one generally uses a Schottky barrier diode balanced bridge in order to reduce the phase jitter of the pumped signal which is transferred via the mixer to the received signal, or the unwanted down converted products $q\omega_0 - p\omega_1$ which are present in the bandwidth of the loop.

We have found a new mathematical method to account for such higher order interaction products. Let $\nu = \frac{\omega_0}{\omega_1}$ the carrier to pump frequency ratio and $\mu = \frac{\omega_i}{\omega_1}$ the intermodulation to pump frequency ratio, these three frequencies may be accounted for by first expanding ν in continuous fractions, that is

$$\nu = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ldots + \cfrac{1}{a_{i\,\max} + \cfrac{1}{\ldots}}}}. \qquad (1)$$

The a_i's are positive integers which are easily obtained from the formulas :

$$a_0 = [\nu], \alpha_0 = \{\nu\},$$
$$\text{and for } i \geq 1 \quad a_i = [1/\alpha_{i-1}], \alpha_i = \{1/\alpha_{i-1}\} \qquad (2)$$

where $[x]$ denotes the integral part of x and $\{x\} = x - [x]$ is its fractional part.

In a second step the "error" frequency ratio μ is computed from the arithmetical expression :

$$\mu = q\,|\nu - p/q| \qquad (3)$$

in which the rational number p/q is obtained by truncating the fraction at some stage $i = imax$ in the continuous fraction expansion (1).

The physical meaning of such a calculation is best understood in terms of absolute frequencies, that is (3) may be rewritten as the down converted product

$$\omega_i = |q\omega_0 - p\omega_1|\,. \qquad (4)$$

The difficulty was to discover a method to account for the interactions between all frequencies in the whole intermodulation spectrum. It was circumvented by expressing this interaction from partitions $(a_1, a_2, ...a_i)$ with $a_i \leq a_{\max}$, with a_{max} an upper bound resulting from the practical set-up and its resolution capability and bandwidth.

In Fig.(1), we chose $a_{max} = 4$ for products with p and q odd integers and $a_{max} = 20$ for products with p or q even integers.

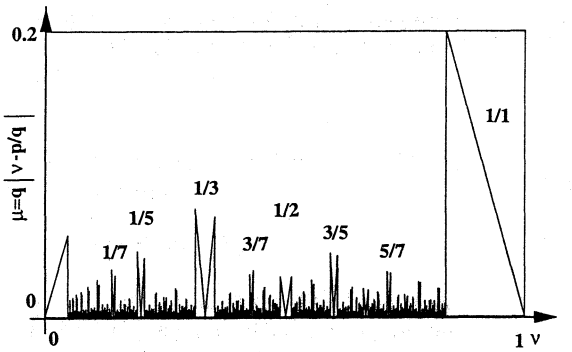


FIG. 1. The intermodulation spectrum as predicted from eq.(1)-(4) ($a_{max} = 4$ if p and q are odd ; $a_{max} = 20$ otherwise)

In order to check the validity of the mathematical approach we have performed the mixing of a fixed frequency pump oscillator of angular frequency $\omega_1 = 2\pi.10$ MHz with a carrier signal of variable frequency $0 < \omega_0 \leq \omega_1$ through a Schottky barrier diode balanced bridge. The down converted beat of the two oscillators at the output of the mixer was counted after passing through a low pass filter of cut-off frequency 100 kHz. The result is shown in Fig.2.

We see that the overall features of the spectrum are compatible with those computed in Fig. 1, although we did not try to get a good fit. It should be observed that the basin of attraction of main products extends far beyond the bandwidth of the loop filter : this result gives credit to the arithmetical approach above. It also confirms our previous hypotheses concerning the relationship between synchronization experiments and the number theory approach [2].

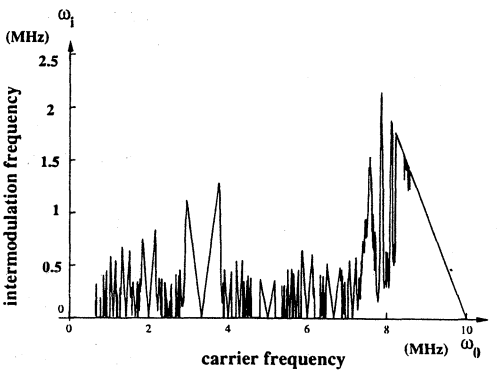


FIG. 2. The experimental intermodulation spectrum

3. THE INTERMODULATION SPECTRUM IN THE CLOSED LOOP

The method can be improved to account for the closed loop configuration. In such a case the frequency ratio results from the synchronization between the carrier and the pump signals. The output of the mixer is written from the cross product of input signals $u_0(t)$ (carrier) and $u_1(t)$ (pump) which is given by :

$$s(t) = \sum_{(p,q)} \beta^{(p,q)} u_0^q(t) u_1^p(t) \qquad (5)$$

where $\beta^{(p,q)}$ is the sensitivity of the mixer at the desired interaction (p, q). After low pass filtering of the signal $s(t)$ the remaining intermodulation products are of the form :

$$s_{filt.}(t) = \sum_{(p,q)} \mu^{(p,q)} \sin\{q\theta(t) - p\psi(t)\} \qquad (6)$$

where $\dot{\theta}(t) = \omega_0$ and $\dot{\psi}(t) = \omega_1 + \alpha\, s_{filt.}(t)$ are respectively the instantaneous frequencies of the carrier oscillator and of the voltage controled oscillator. The phase shift at the (p, q) product is written as :

$$\phi^{(p,q)} = q\theta(t) - p\psi(t) \qquad (7)$$

so that a first order differential system follows :

$$\dot{\phi}^{(p,q)} = \omega_i - p \sum_{(r,s)} K^{(r,s)} \sin \phi^{(r,s)} \qquad (8)$$

with ω_i is given in (4) and $K^{(r,s)} = \alpha\mu^{(r,s)}$ is the open loop gain of the intermodulation product (r, s).

After simple algebra this is rewritten in the uncoupled form as :

$$\dot{\phi}^{(p,q)}(t) = \omega_i - p \sum_{(r,s)} K^{(r,s)} \sin\left\{\frac{s}{p}\phi^{(p,q)}(t) + \frac{\omega_1}{q}(pr - qs)t\right\} \qquad (9)$$

In the simplest case the interaction of one product with the others is neglected so that $r/s = q/p$. The system is governed by :

$$\dot{\phi}\ (t) + K \sin\phi(t) = \Delta\omega \tag{10}$$

where $\Delta\omega = \omega_0 - \omega_1$ and $K = K^{(1,1)}$.

This equation is integrable by quadrature as follows :

$$\int \frac{d\phi}{u - sin\phi} = \int K dt \tag{11}$$

with $u = \frac{\Delta\omega}{K}$. Using elliptic integrals [3], the exact solution of Eq.(10) is :

$$u < 1\ ,\ \phi_A(t) = 2\arctan\left(\frac{1}{u} - \sqrt{\left(\frac{1}{u}\right)^2 - 1}\tanh\left(\sqrt{\left(\frac{1}{u}\right)^2 - 1}\frac{\Delta\omega}{2}t\right)\right) \tag{12}$$

with the property

$$\lim_{t\to\infty} \phi_A(t) = \arcsin(\frac{\Delta\omega}{K}) \tag{13}$$

$$u > 1\ ,\ \phi_B(t) = 2\arctan\left(\frac{1}{u} + \sqrt{1 - \left(\frac{1}{u}\right)^2}\tan\left(\sqrt{1 - \left(\frac{1}{u}\right)^2}\frac{\Delta\omega}{2}t\right)\right) \tag{14}$$

The error signal which is given by $u_{err}(t) = \sin\phi(t)$ is a continuous voltage in (12) and an alternative voltage in the (14) with the frequency

$$\omega_i = K\sqrt{\Delta\omega^2/K^2 - 1} \tag{15}$$

quickly converging to $\Delta\omega$ as $\Delta\omega >> K$.

In general the pump signal acts as a periodic perturbation for the conventional PLL. The modelocked states can be described as satisfying the condition $\dot{\phi}$= *cte* over each cycle provided the counting time is large in comparison to the period.

Let us illustrate the method for the fundamental product $p = q = 1$ under the perturbing action by subharmonics of the pump signal of the form $(r, s) = (r, 1)$. Assuming $\dot{\phi}$= *cte* $= \Delta\phi/\Delta t$ with the period $\Delta t = 2\pi/\omega_1$ and $\Delta\phi = \phi(t+\Delta t) - \phi(t) = \phi_{n+1} - \phi_n$ (n integer) the following mapping results :

$$\phi_{n+1} = \phi_n + 2\pi\nu - c\ \sin\phi_n \tag{16}$$

with $c = 2\pi K/\omega_1$ and $K = \sum_r K^{(r,1)}$ as the strength parameter. This mapping is a well known model for modelockings [4]. They are computed from an average frequency ratio as follows :

$$\nu_{av} = \lim_{n\to\infty} (\phi_{n+1} - \phi_n)/2\pi n \tag{17}$$

and the resulting error signal is obtained by using ν_{av} in (1) instead of the bare value ν.

The curve ν_{av} versus ν is a devil's staircase with steps attached to each rational value of ν. In particular the width of the fundamental step is in order c/π , that is $c\omega_0/\pi = 2K$ in terms of absolute frequencies. This is in agreement with eq.(15) and experiments. In Fig. 3 the nonlinear coupling has been amplified in order to let the locking at subharmonics visible.

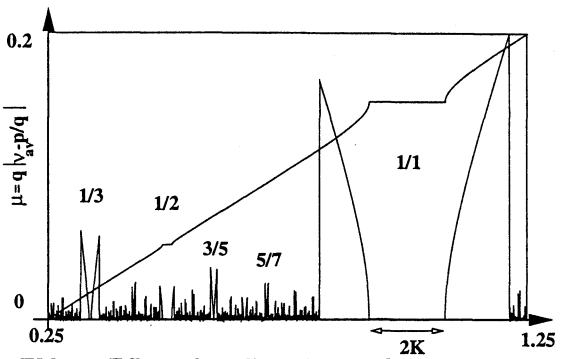


FIG. 3. Effect of nonlinearity on the intermodulation spectrum of Fig.(1) (strengh parameter: $c = 0.5$. The monotonous increasing curve is the devil's staircase ν_{av} versus ν.

4. THE REGENERATIVE PULSE OSCILLATOR

In our second set up the mixing of frequencies is operated in a GaAs modulator [5,6]. The source to drain electrons flow at the carrier frequency in a μm scale GaAs field effect transistor is controlled from a periodic rectangular shape negative voltage applied to the gate. If the applied voltage is strongly negative electrons enters the hot regime, the drift velocity saturates and eventually reaches the region of negative differential resistivity (Gunn effect). In that regime electrons are pinched off and the carrier signal is switched off. On the contrary if the gate voltage is closed to zero the electrons flow gets amplified by channeling from source to drain. A low frequency electronic oscillator has been built by slaving the low frequency gate voltage to the output of the modulator through a transmission delay line.

In terms of the theory of bands the scattering processes in the high electric field leads to the depopulation of the conduction valley (Γ band), where electrons enjoyed a low effective mass and high mobility in favour of on upper valley (X or L bands) where they suffer a large effective mass and lower mobility. They will of course attempt to return to the lowest band and a sustained regime may be obtained between the field-assisted upward transitions

and spontaneous downward transitions. Under appropriate conditions this pumping scheme leads to a self oscillating regime at the gate modulating frequency.

This bas been obtained by using a surface acoustic wave delay line or an electric transmission line in the feedback loop. Well defined plateaus of remarkable accuracy ($\sim 10^{-11}$) have been observed in the plots of the gate frequency versus the frequency or the amplitude of the carrier signal. They correspond to well defined resonances between short and long time scales at integer or rational frequencies. As in the a.c. Josephson effect and the quantum Hall effect the cooperativity of electrons is used to produce visible quantized effects above the thermal noise. But in contrast to the above effects the present experiment may be performed without magnetic field and at room temperature. Details concerning the experimental set-up and the data processing are given in ref. [5]. An exhaustive description of the physical phenomenon being involved in the frequency quantization is not yet at hand. But we succeeded in deriving a non linear mapping which accounts for the main features of this carrier-to-envelope synchronization.

The new advantage of such a set up with respect to the standard frequency locked loop is that the carrier/pump frequency ratio ν is strongly dependant on the amplitude u of the injection carrier signal. In the high amplitude regime an approximate linear dependance on the characteristic ν versus u is clearly identified (Fig.4).

Steps are ordered according to a Farey rule (0/1, ..., p/q, ..., (p+p')/(q+q'), ..., p'/q', ..., 1/1) where p, q, p' and q' are integers. This law gives strong credit to a mathematical approach based on an Arnold map. It should be remembered that such a Farey tree dependance was observed in many natural phenomenon including Josephson superconducting junctions.

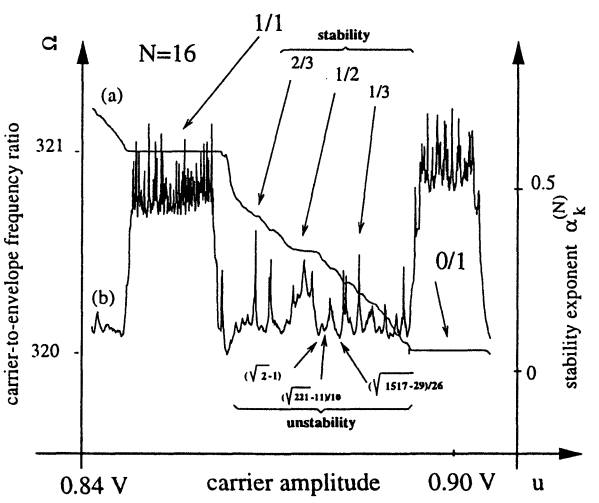


FIG. 4. identification

5. AN UNSTABILITY OF ARITHMETICAL ORIGIN

The approach developed at the beginning of the paper was inspired by results of number theory concerning diophantine approximations of real numbers [7]. Let us rewrite eq.(3) to exhibit the order of approximation of the real number ν :

$$\mu_i = q_i \, |\nu - p_i/q_i| \, . \tag{18}$$

The so-called Markoff constant is defined as :

$$A(\nu) = \lim_{i \to \infty} q_i \mu_i, \tag{19}$$

and is such that $A(\nu) \leq \frac{1}{\sqrt{5}}$ with equality if $\nu = [0; \bar{1}] = \frac{\sqrt{5}-1}{2}$ (the golden mean) with $a_i = 1$ whatever $i > 1$.

In 1879, Markoff improved the result by showing that there exists a discrete set of values $A_i = inf[A(\nu)]$ accumulating at $A_{min} = 1/3$. Such numbers are the so-called Markoff spectrum ; they correspond to badly approximated numbers ν_i, the so-called Markoff irrationalities. The discrete spectrum is given by [7] :

$$A_i = \frac{n_i}{\sqrt{9n_i^2 - 4}} \quad \text{with } n_i = 1, 2, 5, 13, 29, 34, 89, ... \tag{20}$$

and the n_i satisfy the arithmetical law

$$a^2 + b^2 + c^2 = 3abc \tag{21}$$

Results of the classical Markoff theory are illustated in Fig.(5).

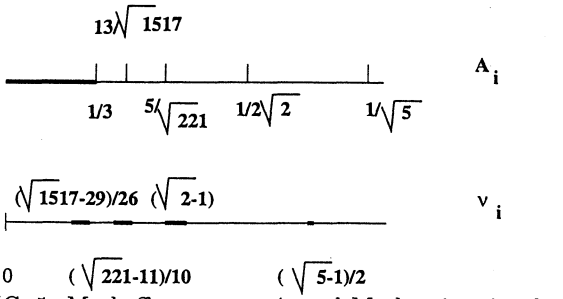


FIG. 5. Markoff spectrum A_i and Markov irrationalities ν_i

Markoff numbers accumulate at 1/3 and Markoff irrationalities (except for the first one) condensate around the tree numbers 0.414, 0.386, and 0.382. A generalisation of Markoff theory based on the group PSL(2,Z) was performed by S. Perrine [7].

Using the numerical approach of the paper it is easy to demonstrate the existence of such badly approximated numbers. By truncating the expansion

(1) to $imax = 7$, Fig.(6) shows that instabilities are growing at specific locations. The peaks at Markoff irrationalities have been identified. However the instability spectrum is extremely rich and nobody yet succeded in predicting it completely.

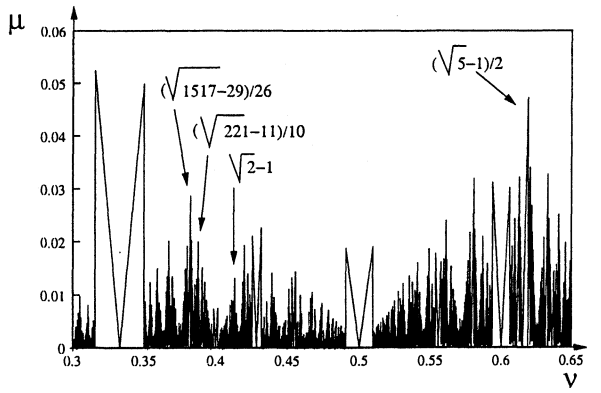


FIG. 6. $imax = 7$

6. CONCLUSION

An efficient algorithm of computing down converted intermodulation products from a mixer has been discovered. It was successfully applied to an electronic feedback loop. In addition the phenomemon of carrier-to-envelope synchronization was experimentaly demonstrated in a regenerative pulse oscillator built from the mixing of frequencies in a GaAs modulator. It confirms the usefulness of number theory in predicting the fine structure of frequency lockings. Finally an instability of an arithmetical origin (Markov spectrum) is demonstrated numerically in agreement with our previous experiments [6]. This creates an opening for future prospects into the study of spurious effects such as troncature versus nonlinear effects and the low frequency jitter noise.

[1] MASS S. A. *Microwave mixers.* Artech House, Inc, Norwood, 1986.

[2] M. PLANAT. *Synchronization, topology and oscillators.* Special issue of Annals of Telecommunications, France, **51**(7-8), 391, (1996).

[3] M. ABRAMOWITZ and I. A. STEGUN. *Handbook of mathematical functions.* Dover publications, INC, New York, 1970.

[4] T. BOHR, P. BAK, and M.H. JENSEN. *Transition to chaos by interaction of resonances in dissipative systems.* Phys. Rev. A, **30**, 1960, (1984).

[5] M. PLANAT, F. LARDET-VIEUDRIN, G. MARTIN, S. DOS SANTOS, and G. MARIANNEAU. *Frequency stability enhancement from carrier-envelope resonance in a surface acoustic wave delay line oscillator.* J. Appl. Phys., **80**, 2509, (1996).

[6] F. LARDET-VIEUDRIN, S. DOS SANTOS, and M. PLANAT. *Blocage par injection de modes sous-harmoniques d'un oscillateur utilisant une ligne à retard à ondes de surface.* Special issue of Annals of Telecommunications, France, **51**(7-8), 330, (1996).

[7] S. PERRINE. *Approximations diophantiennes (Théorie de Markoff).* PhD thesis, Université de Metz, 1988.

Nonparametric Theory Grouped Time and Frequency Standards

E. Erjomin
Science Centre of Metrology (Military Standards) -
Kharkov - Ukraine

We assume that $F_1,..,F_i,..,F_N$ - unknown's means of actual frequency standard's frequency in the group (N-general quantity standards in the group). We have ranged this means : $\{F_i\} \Rightarrow \{F_{(r)}\}$, where r - rank, $F_{(N)}=\max\{F_i\}$, $F_{(1)}=\min\{F_i\}$. Next, we have solve task : transformation F_i to F_i so that

$$|F_i - \min\{F_{(r)}\}| = \max\{|F_k - F_i|\}, \quad i,k=1,2,..,n; \quad i \neq k,$$

besides $F_i = \frac{1}{N}\sum_{i=1}^{N} F_i$.

Was found that

$$|F_i - F_{(1)}| = \max\{F_{(r)} - F_{(p)}\} = \frac{1}{N}\{\sum_{i=1}^{N}[F_{(N+1-2)} - F_{(2)}] + 2\sum_{i=1}^{K}[F_{(k)} - F_{(1)}]\}.$$

We have to find rank's of mean's $\{F_i\}$ by comparison difference between F_k and F_i (with the help of the simple device) and to provide for all $\{F_i\}$ difference between F_i and average mean of group time and frequency standard (F) with "equel errors" deviations. Besides we may not use measuring device's for putting into practice group standard.

LOW NOISE 10-5GHz REGENERATIVE FREQUENCY DIVIDER

Y. Gruson, J. Groslambert, V. Giordano

Laboratoire de Physique et de Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté Besançon
32, avenue de l'Observatoire- 25044 Besançon Cedex - France

1. ABSTRACT

Nowadays, in microwave domain, complex synthesis chains are generally used as reference source for phase noise measurements. In fact due to the high level of flicker frequency noise in conventional microwave sources, low frequency multiplied quartz oscillators appear as the best solution to obtain the lowest noise floor at low Fourier frequencies. The advent of ultra-low noise microwave sources based on sapphire resonators could change this situation. With this new technology, reference source will operate directly at high frequency, ie in X-band region. For any lower frequency oscillator to test, high purity demodulation signals will be obtain by frequency division and mixing with low noise RF synthesizer. Then, it will be necessary to develop low-noise frequency dividers. Regenerative frequency divider constitutes a good candidate for this purpose.
In order to test the potentialities of a such system, we have implemented a 10-5 GHz regenerative frequency divider. In this paper, we present the technical details and performances of this frequency divider.

Keywords: Microwave regenerative frequency dividers, phase noise measurement set-up.

2. INTRODUCTION

Ultra low phase noise X-band source has been demonstrated with the use of sapphire resonators. Phase noise spectrum $S_\varphi(f)$ as low as - 150 dBc @ 1 kHz has been obtained [1,2]. With this new type of microwave source, it is possible to surpass the phase noise performances of the classical synthesis chains based on ultra-stable quartz oscillators. Then X-band sapphire resonator oscillator could constitue an excellent reference for a phase noise measurement set-up. Lets suppose the sapphire reference oscillator operates at 10 GHz. To measure the phase noise of any microwave source operating at a frequency F, we can use a succession of frequency dividers by two and a RF synthetizer operating at a frequency where its phase-noise is minimum. As an exemple, lets suppose that the RF synthetizer presents its better phase-noise performance between 80 and 160 MHz. The frequency F could be expressed in the form:

$$F = a_0 F_0 + a_1 \frac{F_0}{2} + \ldots + a_n \frac{F_0}{2^n} + \Delta F$$

where F_0 is the reference frequency (10 GHz), ΔF is of the order of 150 MHz and $a_i = 0$ or 1.
The figure 1 describes the scheme that could be use for microwave oscillator phase-noise measurement. This circuit is equivalent to a source with a output frequency given by:

$$nF_{REF}/2^6$$

with $1 \le n \le 64$.

The performance of a such configuration will be mainly limited by the reference source for low Fourier frequencies and by the RF synthetizer phase noise for f higher than 1 kHz (see figure 2) .

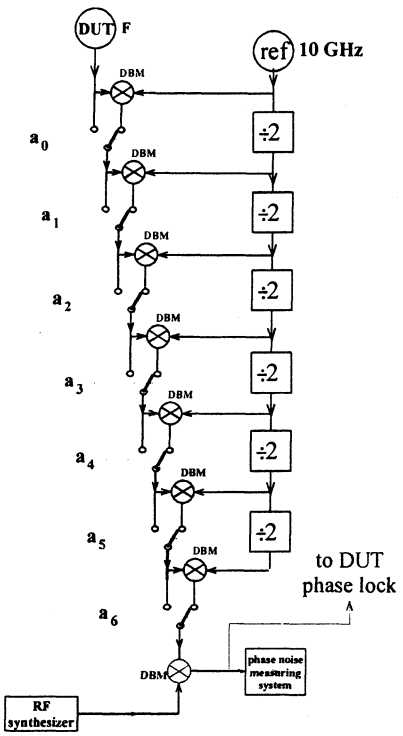


Figure 1. Phase noise measurement with X-band reference. (Actual circuit will contain additional low-noise amplifiers)

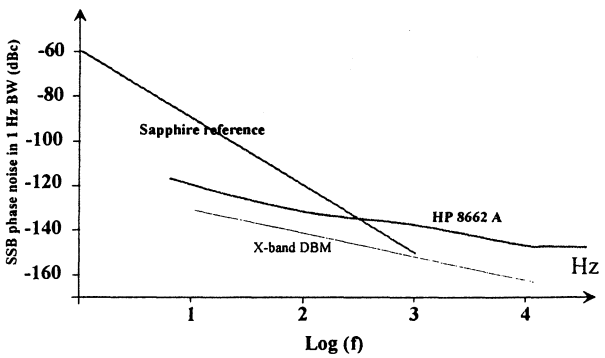


Figure 2. Expected performances of the microwave phase measurement set-up at 10 GHz.

Such a noise floor will be obtained only if we dispose of low-noise frequency dividers. At frequencies less than 1 GHz there is no difficulty to find dividers with added phase noise better then -120 dBc @ 1 Hz [3]. Nevertheless the added phase noise due to the first microwave frequency divider has

to be considered. There is not a lot of available information about phase noise level of microwave frequency dividers. That's why we decided to design and test microwave 10-5 GHz regenerative divider. In this paper, after a brief recall of regenerative frequency divider principles, we give the technical details of our conception. Phase noise levels of 2 equivalent 10-5 GHz frequency dividers were measured. The obtained results show that the regenerative frequency divider procures good phase noise performances compatible with our application.

3. REGENERATIVE DIVIDER PRINCIPLE.

The principle of regenerative frequency divider by N+1 is explicited in figure 3 :

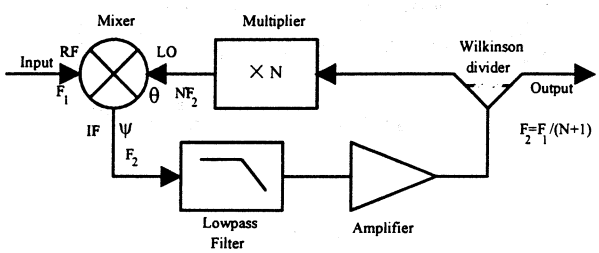


Figure 3. Principle of regenerative frequency divider

The input signal (frequency F_1) and the LO signal (frequency F_{LO}) are mixed. At the mixer IF port, we obtain many harmonics as $pF_1 \pm qF_{LO}$. The low pass filter select the lowest one:

$$F_2 = F_1 - F_{LO}$$

As $F_{LO} = N.F_2$, at the output of the circuit we obtain a signal at the frequency F_2 which is :

$$F_2 = F_1 / (N+1)$$

Regenerative frequency dividers operating in the RF band present low added phase noise [4]. The divider performance is mainly limited by the intrinsic phase noise due to the loop amplifier. It has been demonstrated that the output noise spectrum $S_\varphi(f)$ is related to the internal noise spectrum $S_\gamma(f)$ (mainly due to amplifier noise) by the relation [5]:

$$S_\varphi(f) = \left|\frac{1}{1-G}\right|^2 S_\gamma(f)$$

where G is defined as the phase gain of the mixer. This parameter depends on the phase and the level of the different harmonics created by mixing. It is generally of the order of -1 but varies with the level and the phase of the signals present at the mixer input ports.

We tested (10-5GHz) regenerative dividers (N=1). There is no multiplier in the circuit. We constructed first an "on table" prototype with commercially avalaible components. This first approach allowed to demonstrate the possibility of such frequency dividers. A more compact circuit was elaborated in a second time which give about the same phase noise performances.

4. "ON TABLE " PROTOTYPE

The figure 4 represents the "on table " prototype constructed with some commercially available components (loop amplifier , mixer and phase shifter) . The low pass filter and the power divider were home made.

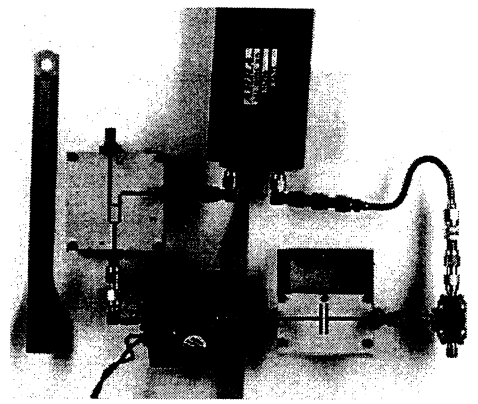


Figure 4. The "on table" prototype of frequency divider

The caracteristics of the different components are given in the following table :

loop amplifier	MENLO MLA-60120	Gain :25 dB Output power 18 dBm (1 dB compression)
mixer	MITEQ TB0218LW2	RF/LO 2 to 22 GHz IF 0.5 to 10 GHz
phase shifter	ARRA	
lowpass filter	LPMO	f_c= 6.2 GHz
power divider	LPMO	Wilkinson

Table 1 : Components characteristics

Mixer Phase noise measurement:
The following procedure has been used to measure the mixer added phase noise :

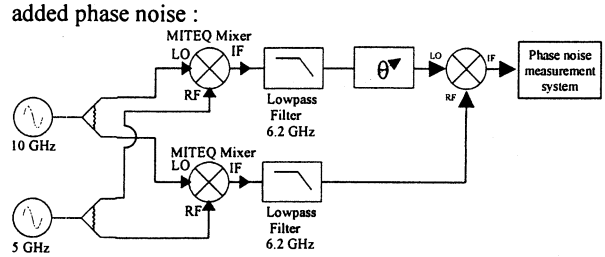


Figure 5. Mixer phase noise measurement procedure

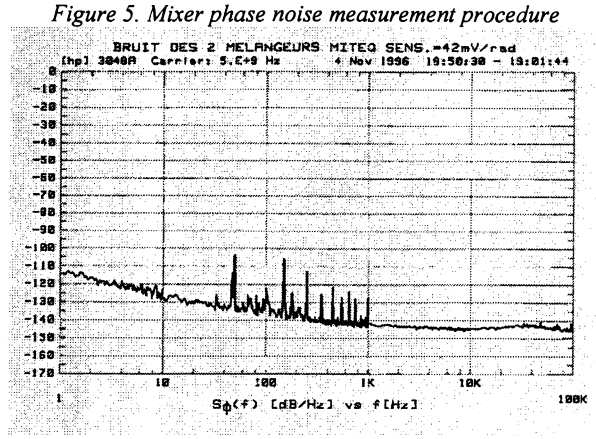


Figure 6 . Added phase noise of Miteq Mixer

The added phase noise of one mixer is -118dB/Hz @ 1Hz.

Amplifier phase noise:

A standard procedure has be used to evaluate the added phase noise of the loop amplifier. We measured -105dBc@1Hz.

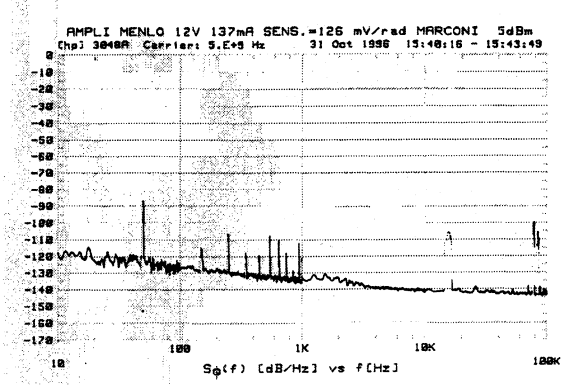


Figure 7. Commercial amplifier phase noise

Divider phase noise

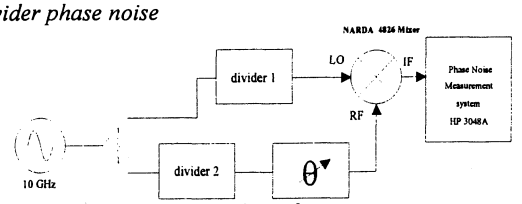


Figure 8. Frequency dividers phase noise measurement.

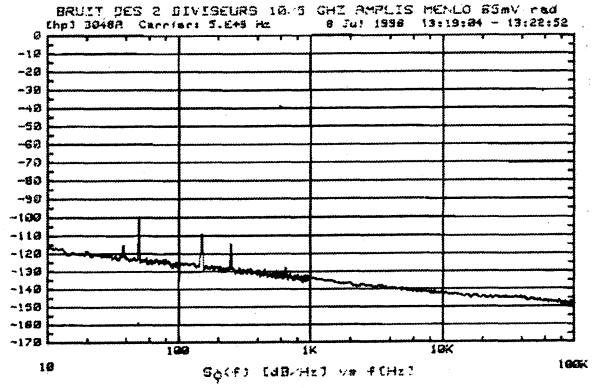


Figure 9. Frequency dividers phase noise

The added phase noise of one divider is -108dB/Hz @ 1Hz which is comparable to the added phase noise of the loop amplifier. In the figure 10 we compare the divider phase to the expected performances of the phase noise measurement set-up. One can see that the regenerative divider is compatible with a such noise floor.

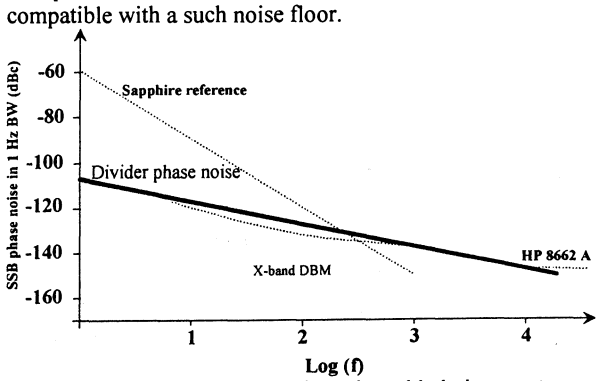


Figure 10. Comparaison of divider added phase noise and noise floor.

5. COMPACT VERSION OF THE DIVIDER

To facilitate the implementation of the divider in our laboratory instrumentation system, we design a compact version of the 10-5 GHz divider (see figure 11). The overall dimensions are : 11 x 10 x 2 cm^3.

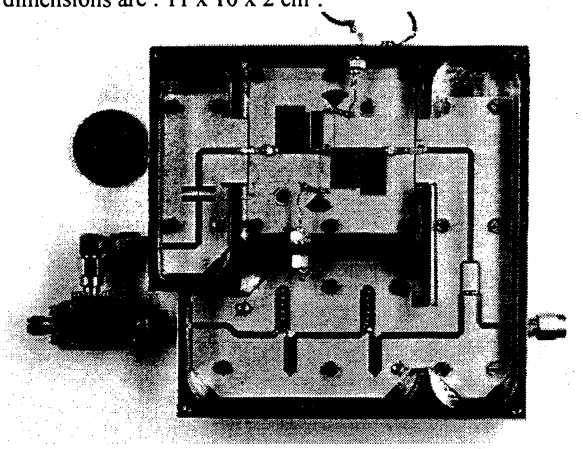


Figure 11. Compact version of the divider

One find the same components as previoulsy. The loop amplifier is single stage amplifier designed with a HP MESFET ATF25570. Its gain is 13 dB. The output power is 17 dBm (1 dB compression). The phase noise of this amplifier was measured and the result is given in figure 12 :

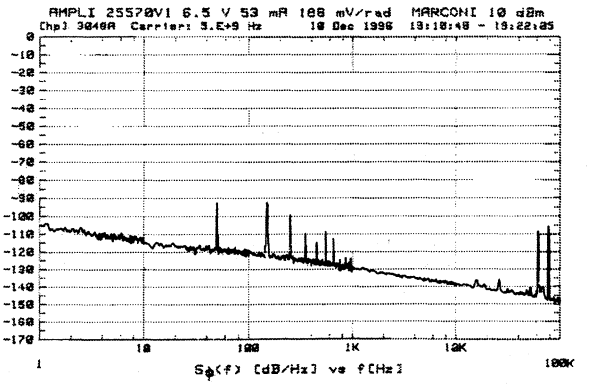


Figure 12. Phase noise of the ATF 25570 single stage loop amplifier.

It should be noted that better phase noise performances could be obtained by using a bipolar transistor in place of the MESFET. Nevertheless, we have to find a transistor which presents sufficiant gain at 5 GHz.

To adjust the phase in the loop, a tunable length 50Ω line was used rather than a varactor phase shifter. This last solution allows easier adjustement but could induce extra noise.

Finally two equivalent dividers were constructed and tested. The figure 13 represents the added phase noise of the two dividers:

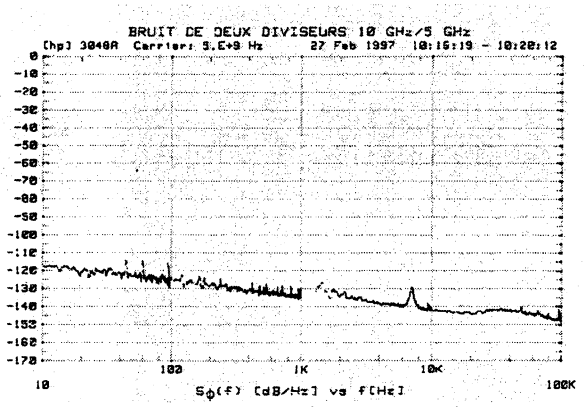


Figure 13. Phase noise of two compact dividers

The compact version presents about the same performances as the previous circuit. At the output of the divider we observe harmonics whose level is relatively low:

5 GHz : 10 dBm
10 GHz : -23 dBm
15 GHz : -33 dBm

6. CONCLUSION

We have demonstated the possibility to realize low phase noise microwave frequency dividers.
The observed noise spectrum presents a 1/f noise component with a level of- 108dB/Hz @ 1Hz. The obtained phase noise performances are compatible with the expected noise floor of microwave ultra-low noise sapphire oscillator.

7. REFERENCES

[1] D.G. Santiago, G.J. Dick, Microwave frequency discriminator with a cooled sapphire resonator for ultra-low noise phase noise. *Proceedings of the 1992 IEEE Frequency Control Symposium,* pp 176.182.

[2] E.N. Ivanov, M.E. Tobar, R.A. Woode. Advanced Phase Noise Suppression Technique for next Generation of Ultra-Low Noise Microwave Oscillators. *Proceedings of the 1995 IEEE Frequency Control Symposium,* pp.314-320.

[3] F.L. Walls and C.M Felton, Low noise frequency synthesis. *Proceedings of the 41st Annual Frequency Control Symposium 1987,* pp 512-518.

[4] J. Groslambert, M. Olivier, E. Rubiola, High Spectral Purity Frequency sources using Low Noise Regenerative Frequency Dividers, *Proceedings of the 45th Annual Frequency Control Symposium 1991.*

[5] E. Rubiola, M. Olivier, J. Groslambert, Phase Noise of Regenerative Frequency Dividers, *Proceedings of the 5th European Frequency and Time Forum 1991, pp 115-122*

SOME APPROACH TO COMPUTATION OF ADEV, TDEV AND MTIE

Michal Kasznia

Institute of Electronics and Telecommunications
Poznan University of Technology
Piotrowo 3A, 60-965 Poznan, Poland
email: mkasznia@et.put.poznan.pl

ABSTRACT

In the paper some saving time procedures of computations of ADEV, TDEV and MTIE are presented. In the first section author introduces the problem and presents the main reason for these considerations. In the next part the rules of measuring of synchronization signal parameters are presented. Then the procedures of ADEV and TDEV calculation and some modifications of theirs are also proposed and desribed. In section 4 the procedure of MTIE calculation and its modification is described. The results of computations with and without simplification are presented and compared for all parameters. In the end of the paper author comments proposed modification and obtained results.

Keywords: time error (TE), maximum time interval error (MTIE), Allan deviation (ADEV), time deviation (TDEV)

1. INTRODUCTION

The reason for these considerations was the design of software for measuring system SP-2000. The system designed in Institute of Electronics and Telecommunications Poznan University of Technology enables measuring synchronization signal parameters in PDH and SDH networks. The software of the system makes computations of Allan deviation (ADEV), time deviation (TDEV) and maximum time interval error (MTIE) possible. These parameters are basic values characterized the quality of synchronization signal. The results of calculation are presented as a function of observation interval τ in the form of plot in logaritmic scale of time and value (Fig. 1). User interface of SP-2000 software enables graphic comparison of the obtained results with the standards on the plots presented on the screen. In Figure 1 an example plot of TDEV compared with the standards for synchronization suply unit (SSU) and SDH equipment clock (SEC) is presented.
Because of complicated formulas describing the parameters and very long data sequences needed to calculate the parameters, time used for computation is very long. It is inconvenient for a user, who has no access to a fast workstation in the node of telecommunication network or in the central office. Thus the main problem is to decrease the time used for computation the parameters.

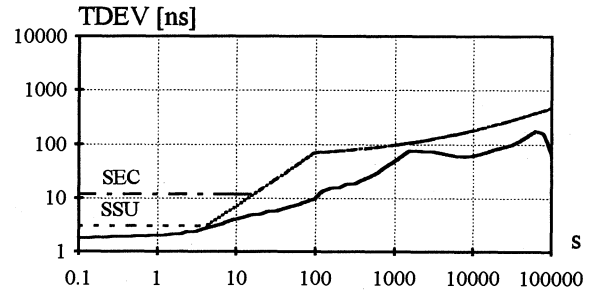


Fig. 1. Example plot of TDEV in logaritmic scale of time and value

2. TIME ERROR MEASUREMENT

The main characteristic of time (synchronization) signal is time error (TE) (Ref. 2). Using time error values other parameters characterizing time signal are calculated. In order to calculate ADEV, TDEV and MTIE time error is measured first. The results of measurement are taken with interval τ_0 and stored in the data file. In the telecommunication standards (Ref. 2) some conditions for the measurement of time error are specified. Time error shall be measured using an anti-aliasing filter with the cut-off frequency f_0. The minimum observation interval τ_{min} must be three times greater then maximum sampling time τ_0 and the anti-aliasing filter cut-off frequency f_0 must be equal $f_0=1/\tau_{min}$. For observation intervals of 0.1 s to 1000 s the maximum sampling interval must be $\tau_0=1/30$ s with the filter frequency $f_0=10$ Hz. For observation intervals of 10 s to 100000 s the maximum sampling interval must be $\tau_0=3.3$ s with the filter frequency $f_0=0.1$ Hz. The minimum measurement period T for time deviation calculation shall be twelve times longer then the observation interval τ.

In the paper two cases will be analized. In the first case the time error was measured with the interval τ_0=1/30 s in the period T=4 h. The results of ADEV, TDEV and MTIE calculations are presented for observation intervals of 0.1 s to 1000 s. The results of time error measurement for this case are presented in Figure 2. In the second case time error was measured with the interval τ_0=3.3 s in the period of four days. The value of this period doesn't follow the requirement given above, but it is sufficient to obtain some results of TDEV for the observation intervals of 10 s to 100000 s and to analize the calculation algorithm. The results of TE are presented in Figure 3. The analized data for both cases were obtained by means of SP-2000 mesuring system.

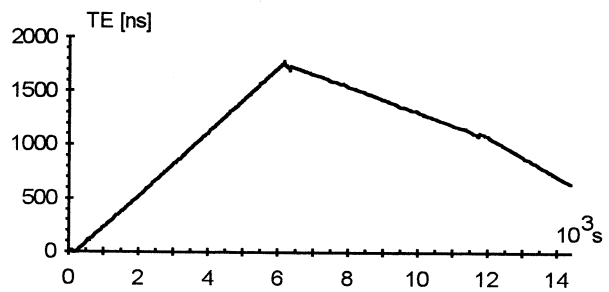


Fig. 2. Results of time error measurement in the period of 4 hours with sampling interval τ_0=1/30 s

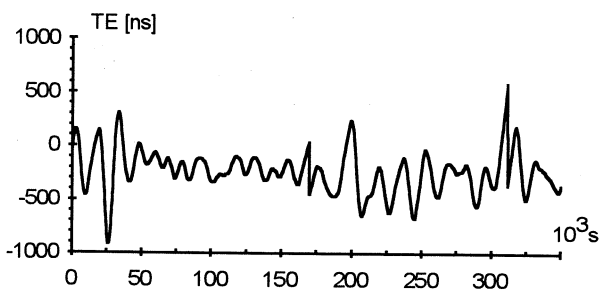


Fig. 3. Results of time error measurement in the period of 4 days with sampling interval τ_0=3.3 s

3. ADEV AND TDEV CALCULATION

3.1. ADEV and TDEV calculation procedures

In telecommunication standards (Ref. 2) the formulas for Allan deviation ADEV and time deviation TDEV are defined. Allan deviation can be estimated by:

$$\hat{ADEV}(\tau)=\sqrt{\frac{1}{2n^2\tau_0^2(N-2n)}\sum_{i=1}^{N-2n}(x_{i+2n}-2x_{i+n}+x_i)^2} \quad (1)$$

where:
$\{x_i\}$ - sequence of N samples of time error function x(t) taken with interval τ_0;
$\tau=n\tau_0$ - observation interval.
Time deviation can be estimated by:

$$\hat{TDEV}(\tau)=\sqrt{\frac{1}{6n^2(N-3n+1)}\sum_{j=1}^{N-3n+1}\left[\sum_{i=j}^{j+n-1}(x_{i+2n}-2x_{i+n}+x_i)\right]^2} \quad (2)$$

where:
$\{x_i\}$ - sequence of N samples of time error function x(t) taken with interval τ_0;
$\tau=n\tau_0$ - observation interval.

In order to reduce the calculation time of the sum in (2), the estimator will be presented in the form:

$$\hat{TDEV}(n\tau_0)=\sqrt{c(n)\sum_{j=1}^{N-3n+1}S_j^2} \quad (3)$$

where:

$$c(n)=\frac{1}{6n^2(N-3n+1)}$$

$$S_j=S_{j-1}-x_{j-1}+3x_{j+n-1}-3x_{j+2n-1}+x_{j+3n-1}$$

$$S_1=\sum_{i=1}^{n}(x_{i+2n}-2x_{i+n}+x_i)$$

Both formulas in (2) and (3) are identical.
It was mentioned, that the results of computation are presented as plots in logaritmic scale of time. The logaritmic plot is devided into 6 segments, as in Figure 1. The value n, that set the size of observation interval τ, is taken from:

$$n=k\cdot 10^{\frac{i}{m}}$$

where i - the number of points on the chart, m - number of points for one segment, k - number of sampling intervals τ_0 per minimum observation interval τ_{min}, $\tau_{min}=k\tau_0$. If the samples of time error are taken with the sampling interval τ_0=1/30 s and the minimum observation interval is τ_{min}=0.1 s, than for the plot devided into 6 segments with 25 points, the observation interval τ for ADEV and TDEV computation is given as:

$$\tau=n\tau_0=3\cdot 10^{\frac{i}{25}}\tau_0$$

where i = 1..150.
The conditions described above where applied for computations carried out and described in this paper.

3.2. Modification of ADEV and TDEV procedures

In order to speed-up the computation of ADEV and TDEV the procedures given by (1) and (2) were modificated [1]. The calculations of these parameters are divided into 6 calculation segments (like in Fig. 1), respectively: 0.1-1 s, 1-10 s, 10-100 s, 100-1000 s, 1000-10000 s,

10000-100000 s. Dependig on the length of measuring period the number of segments can be smaller. The results of time error measurement are stored in data file. For every calculation segment the set of data and the interval between samples are modified. In the first two segments: (0.1-1) and (1-10), the data are used for calculation without any change. The data interval τ'_0 used for computation is equal to sampling interval τ_0. Shifting to the next segments the data interval τ'_0 is magnified ten times. In the third segment (10-100) it is equal $\tau'_0=10\tau_0$. A new value in the next set is the average value of ten values taken from the previous data set. This algorithm for sampling interval τ_0=1/30 s is presented in Figure 4.

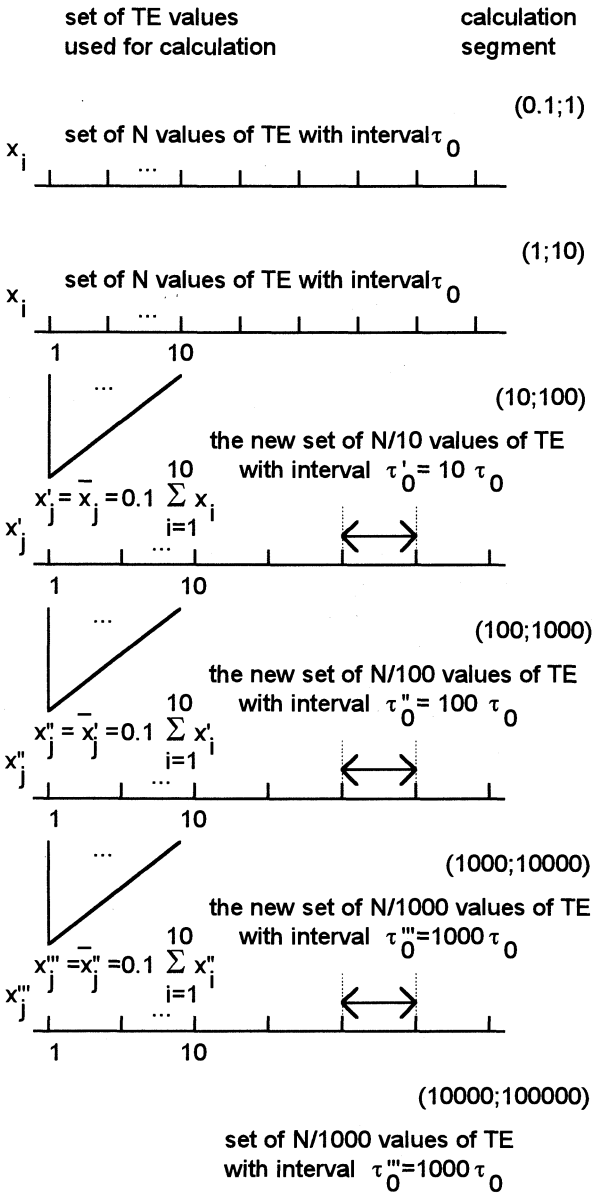


Fig. 4. Modification of the set of data used for ADEV and TDEV computation

In general, for all values of τ_0, the first calculation segment is ($3\tau_0$-$30\tau_0$). The modifications starts for observation interval $\tau=300\tau_0$. These modifications are provided for every segment, except of the last one (for τ_0=1/30 s it is 10000-100000 s). This action results in shorter data sequence used for calculation and smaller computation time in effect. In Table 1 the relations between the number of segment and the number of data used for TDEV calculation are presented. In this case the maximum observation interval is 1000 s and the sampling interval is τ_0=1/30 s.

segment number	observation interval τ [s]	data interval τ'_0 [s]	amount of data N
1	0.1-1	0.033	$3.6\ 10^5$
2	1-10	0.033	$3.6\ 10^5$
3	10-100	0.33	$3.6\ 10^4$
4	100-1000	3.3	$3.6\ 10^3$

Table 1. Segment, data interval τ'_0 and number of data N relations

3.3. Results of calculations

For both cases of time error measurement presented in Figure 2 and 3 the computations are provided. In Figure 5 the results of Allan deviation computations for first case with (ADEVm) and without (ADEV) modifications are presented. In Figure 6 and 7 the error and relative error between results is shown. Note that there is no error for first two calculation segments. The error has the greatest value for the begining of segments (about of 10% for third and fourth segment) and is falling down with observation interval.

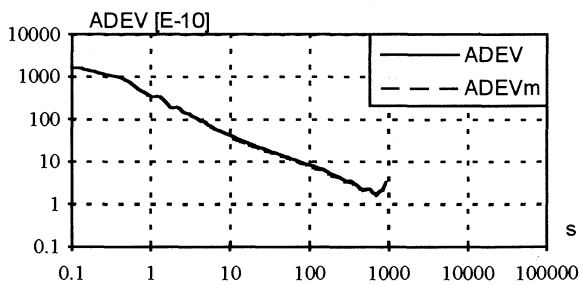


Fig. 5. Results of ADEV computation with and without modifications for the first case of TE measurement

In Figure 7-8 the results of time deviation computation and the plots of error and relative error are presented. The computations with (TDEVm) and without (TDEV) modifications are provided for τ_0=1/30 s. The error is between -0.5

and 0.5 ns. The relative error is smaller then 1% and much smaller then in ADEV.

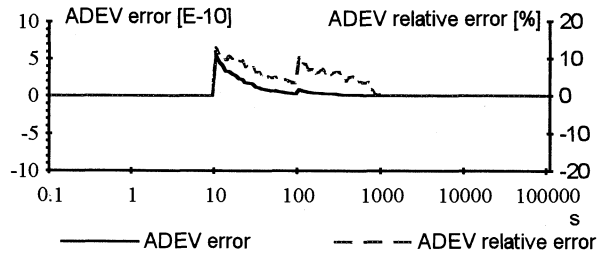


Fig. 6. Error and relative error of ADEV calculation

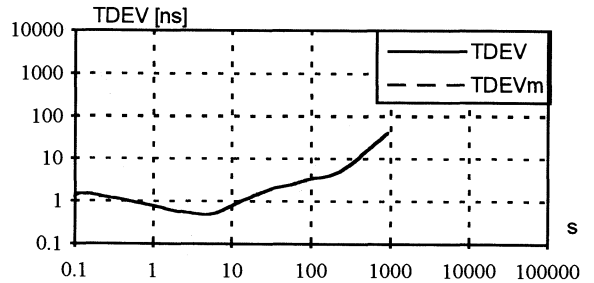


Fig. 7. Results of TDEV computation with and without modifications for the first case of TE measurement

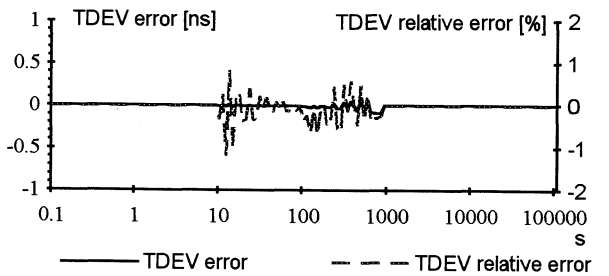


Fig. 8. Error and relative error of TDEV calculation

In Figure 9-10 and 11-12 the results and error of computations of ADEV and TDEV are presented respectively. These characteristics are calculated for the sampling interval τ_0=3.3 s. Similarly as in the first case, the relative error ist greater for ADEV.

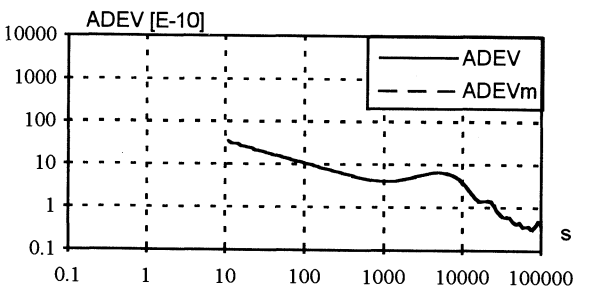


Fig. 9. Results of ADEV computation with and without modifications for the second case of TE measurement

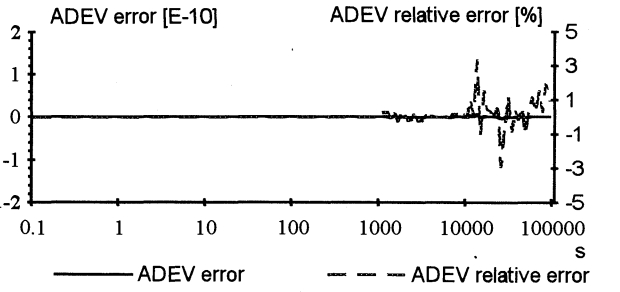


Fig. 10. Error and relative error of ADEV calculation

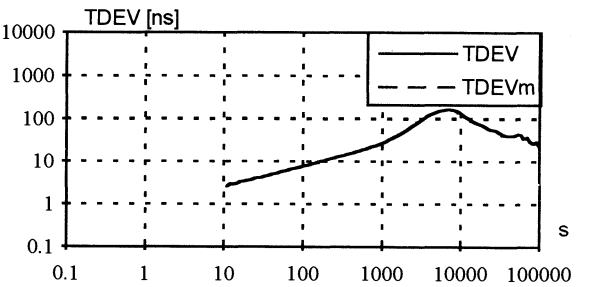


Fig. 11. Results of TDEV computation with and without modifications for the second case of TE measurement

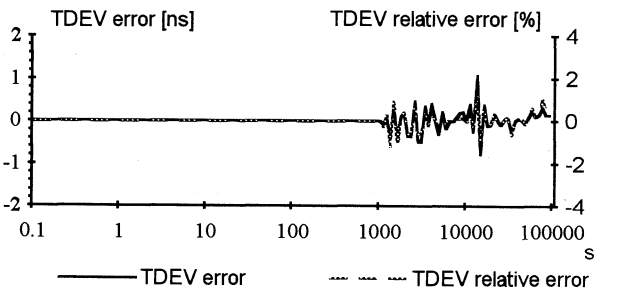


Fig. 12. Error and relative error of TDEV calculation

4. MTIE CALCULATION

4.1. MTIE calculation procedure and its modification

Maksimum time interval error is defined as maksimum peak-to-peak value of time error function in all observation intervals τ enclosed in measurement period T. MTIE can be estimated by:

$$\hat{MTIE}(n\tau_0) = \max_{1\le k\le N-n}\left(\max_{k\le i\le k+n} x_i - \min_{k\le i\le k+n} x_i\right) \qquad (4)$$

where:
$\{x_i\}$ - sequence of N equally spaced samples of time error function x(t) taken with interval τ_0;
$\tau=n\tau_0$ - observation interval.

In order to calculate MTIE for observation interval $\tau=n\tau_0$, a "window" with width of n samples must

be created and in this window the peak-to-peak value must be found (Fig. 13). Then the window is shifted trough the data sequence and the operation is repeated. For *N* samples in the sequence this operation must be repeated *N-n* times.

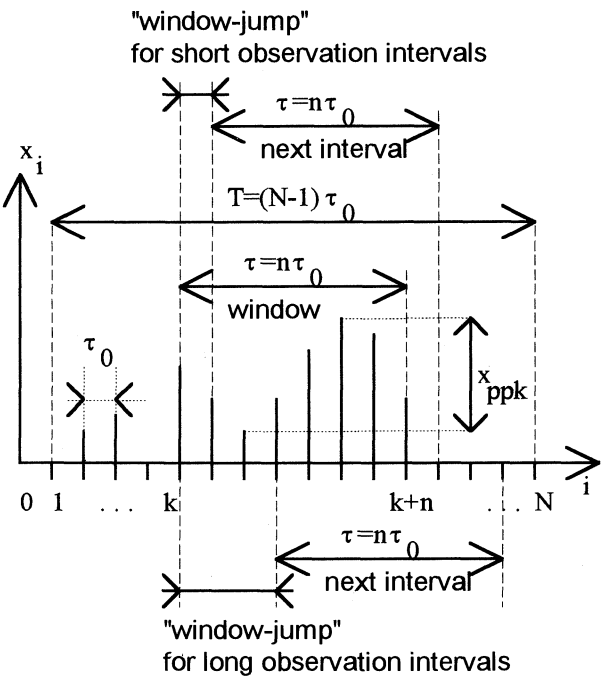


Fig. 13. Method for MTIE calculation with modification

segment number	observation interval τ [s]	window-jump [s]
1	0.1-1	0.033
2	1-10	0.33
3	10-100	0.33
4	100-1000	3.3
5	1000-10000	3.3
6	10000-100000	3.3

Table 2. Relation between observation interval and window-jump

The calculation complexity is growing rapidly with incresing *n*, therefore the following modification is suggested. For observation intervals longer than $30\tau_0$, the "window-jump" over the data sequence is greater than for short intervals (Fig. 13). It is performed with the shift $10\tau_0$ instead of τ_0. Therefore for long intervals $\tau=n\tau_0$, the number of intervals for analysis is ten times smaller. The next increase of "window-jump" is performed for observation intervals longer then $3000\tau_0$. In this case the value of shift is $100\tau_0$. For sampling interval τ_0=1/30 s the relations between observation intervals and window-jump are presented in the Table 2.

4.2. Results of calculations

The results of MTIE computation with (MTIEm) and without (MTIE) modifications for the sampling interval τ_0=1/30 s are presented in Figure 14. Figure 15 shows the error and relative error of computation. The error comes up to 9 ns and the relative error to 20%.

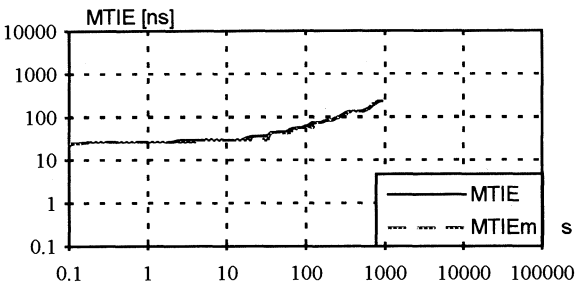


Fig. 14. Results of MTIE calculation for the sampling interval τ_0=1/30 s

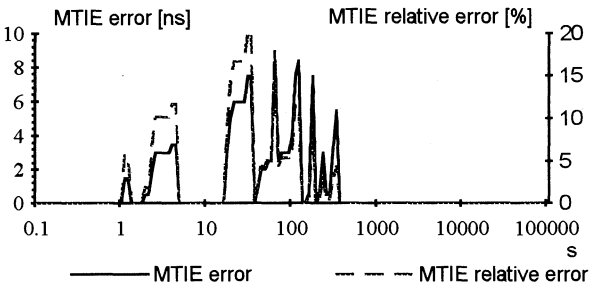


Fig. 15. Error and relative error of MTIE calculation

5. CONCLUSIONS

Presented modifications of ADEV, TDEV and MTIE computation allow to save time used for calculations. The differences between the obtained results appears to be reasonable for ADEV and TDEV. It is suggested to start modified MTIE computation from the third segment.

This work was performed in the frame of project TB-44-485/97-DS.

REFERENCES

1. A. Dobrogowski, M. Jessa, M. Kasznia, K. Lange, "Measuring and calculation of synchronization signal parameters in telecommunication network", Poznan Telecommunications Workshop PWT'96, conference materials pp. 132-139, Poznan 11-12 December 1996. (in Polish)

2. European Telecommunication Standard. Transmission and Multiplexing; Generic requirements for synchronization networks. ETSI 1995.

256 Mbps REAL-TIME VLBI SYSTEM USING A HIGH-SPEED ATM NETWORK

Hitoshi KIUCHI*, Michito IMAE*, Tetsuro KONDO*, Mamoru SEKIDO*, Shin'ichi HAMA*
Takashi HOSHINO**, Hisao UOSE**, and Takashi YAMAMOTO**

*Communications Research Laboratory, Japan
**NTT Telecommunication Network Laboratory Group, Japan

1. ABSTRACT

The communications Research Laboratory and NTT Telecommunication Network Laboratory Group have developed a new VLBI system, called the KSP (Key Stone Project) real-time VLBI system, using a high-speed ATM (Asynchronous Transfer Mode) network. Research into digital data transmission using optical fiber links has made remarkable progress in recent years. In real-time VLBI, the cross-correlation processing and data observation are carried out simultaneously, and it takes about three hours to analysis the crustal deformation data after the VLBI observation is completed. This system is one of the most significant technologies in the KSP, and will give us data on crustal deformation in the Tokyo Metropolitan Area quickly.

Keywords: Real-time VLBI, ATM (Asynchronous Transfer Mode) network

2. INTRODUCTION

"VLBI" stands for Very Long Baseline Interferometry (Ref. 1), and is a type of radio interferometry. VLBI is based on techniques of space communications, the atomic frequency standard, and ionospheric observation. In VLBI, two or more antennas independently receive a radio wave coming from a celestial radio source. The received signals are recorded on magnetic tape together with a timing signal from the atomic standard. After shipment of the magnetic tapes from the observation station to the correlation site, data is reproduced and cross-correlated to obtain the signal delays of the two stations' received signals. From the obtained delays we can determine the precise distance between the antennas. VLBI is a very effective method for measuring the motion of the earth's plates. The plate tectonics theory was corroborated using this method. In KSP (Key Stone Project), metropolitan area crustal deformation measurement project, we have two VLBI systems; a tape-based and a real-time VLBI systems. The tape-based system is already operated in daily observation, of which accuracy of 3 mm is achieved. In the current tape-based VLBI system, the observed data is recorded on magnetic tapes at the observing site, and the tapes are transported to the correlation site by mail, and the analysis is done the next day. So it takes one day to obtain a measured value of crustal deformation. The major cause of this delay is data (tape) transportation, which takes over 18 hours. This problem is solved by using an ATM network (Ref. 2), which has a transmission capability of up to 2.4 Gbps in a basic-optical fiber link. The observed VLBI data is transmitted through the ATM communication network instead of being recorded onto the magnetic tapes. The KSP has four VLBI stations connected by an ATM network and the data transmitted from these remote observing stations to the correlation site in Tokyo is processed in real time.

Real-time VLBI is expected to be one of the breakthrough technologies for improving VLBI performance.

3. ATM NETWORK

A block diagram of the real-time VLBI is shown in Figure 1. ATM transmits information in fixed length packets called cells. A cell is composed of 53 bytes of data in total: a 5-byte header and a 48-byte payload. The signals input to the ATM transmitter are written in the payload of the ATM cell in arrival order (Fig. 2). The header of the cell, which shows its destination, is attached when the payload becomes full, and the cell is output to the 2.488-Gbps transmission line. A VPI (Virtual Path Identifier) shows the destination of the virtual path in the header. It is necessary to collect the data observed at the four stations at the correlation site. Signals from the each observation station are multiplexed by the cross-connect switch located in ATM network, and the multiplexed signal is carried to the receiver along one transmission line. The virtual path for each cell is selected by the VPI on the cell.

It is possible to transmit data of various rates through the same virtual path. In ATM, the transmission rate (capacity) can be chosen freely. There is no need for changes in the data transfer rate to be controlled on the network side.

In the receiver, the multiplex signal is separated into the data for each station and restored in the receiver as a digital signal after the delay fluctuation is absorbed. The real-time VLBI system also has a function that removes delay and delay fluctuations of the transmission system in the receiver. The delay sustained by each cell depends upon the cross-connect switching timing. Therefore, the cell interval is not preserved in the ATM network. The arrival interval of the cells differs (fluctuated) from the initial interval. This is the cell delay variation (CDV), and is dependent on the number of pieces of ATM equipment, data rate, transmission line accommodation rate, and the traffic characteristics in the virtual path along which the cell passes. The cell delay fluctuation is absorbed by the memory in the receiver, and reassembled so that to the

data is regularly spaced and equal to that on the transmitter side.

Cell loss or mis-delivery may occur in the transmission of signals through the ATM network. In VLBI, bit-make and bit-slip are more likely to be fatal than bit-error. In the receiver, the mistaken delivery and cell loss that lead to bit-make and bit-slip are detected as fatal errors. If cell loss occurs, the lost bits are inserted with the same number of bits, and if a cell is mistakenly delivered, it is removed.

4. VLBI correlation processing system for real-time VLBI

The K-4/KSP VLBI data acquisition system (Ref. 3) consists of a reference distributor, IF distributor, local oscillator, video converter, input interface, and ATM network.

The local oscillator synthesizes the local frequency signal for the video converter. By the action of this video converter, windows in the IF-signal (500-1000 MHz) input are converted into video signals (32 MHz). The frequency conversion is achieved by the image rejection mixer using single-sideband conversion. The input interface unit samples the video signal from the video converter, and sends the digital data to the data recorder or the ATM transmitter together with the time data, which is phase locked to an external time reference.

The input interface unit quantizes the 16-channel (max.) video signal, and produces a data train of up to 256 Mbps (max.).

The rate at which data is output to the recorder or the ATM transmitter can be selected from five rates between 16 and 256 Mbps. Time-code insertion can be selected to be uniformly spaced or at uniform time intervals, or no time-code can be selected. In real-time VLBI, uniformly spaced time code is used. This is a requirement of the ATM side.

In real-time VLBI, multiple signals are separated by the virtual path isolating function of the receiver into the signals from each observation station. The transmission path delay between the observation station and the correlation site is different for each station and must be absorbed. The data from each station is time synchronized in the receiver (Fig. 3). Time stamps composed of year, day, time, minute, and second information and a SYNC code, which is used for the time code recognition, are inserted in the data at regular intervals (every 64 Mbit). The data is output to the correlator after the timing has been synchronized, so the output data for each station is correct up to the time on the time stamp.

The data for an identical time taken at different stations differs in arrival time by the difference of the transmission path length. To absorb the transmission path delay, the signal begins to accumulate in the buffer memory from the time when the time stamp is received. The accumulated period in the buffer memory is long when the transmission path is short, and is short when the transmission path is long. The readout of the buffer memory starts immediately after the time stamps from all observation stations have arrived. Thus the timing can be synchronized. The correlation processing system of the real-time VLBI is shown in Fig. 4.

In ATM, the transmission rate (capacity) can be chosen freely. There is no need for changes in the data transfer rate to be controlled on the network side.

The obtained fringe (correlated pattern) in fine delay search function is shown in Fig. 5.

An input/output signal of the ATM transmission system is unified to the D-1 standard (Ref. 4), and is adapted to have the same data format as the tape-based K-4/ KSP system. Adoption of the identical data format enables the tape-based correlation processing system to be used as the real-time system.

5. conclusion

It is possible to adopt tape-based VLBI and real-time VLBI with identical formats, by using the same data-acquisition and correlation systems. A tape-based 4-station 6-baseline KSP system is used for daily crustal measurement. Good fringes (correlated patterns) are obtained in the real-time KSP system, and we are making automatic correlation control software. In real-time VLBI, the cross-correlation processing and data observation are carried out simultaneously, and it takes about three hours to analysis the crustal deformation data after the VLBI observation is completed. This system is one of the most significant technologies in the KSP, and will give us data on crustal deformation in the Tokyo Metropolitan Area quickly.

Reference

[1] A.E.E. Rogers,"Coherence limits for very-long-baseline interferometry," IEEE Trans. Instrum. Meas., vol. IM-30, no. 4, pp. 283-286, Dec. 1981.

[2] K. Sato, S.Ohta and I. Tokizawa, "Broad-band ATM network architecture based on virtual paths," IEEE Trans. Commun. vol.38, No.8, Aug. 1990.

[3] H. Kiuchi, S. Hama, J. Amagai, Y. Abe, Y. Sugimoto, and N. Kawaguchi,"K-3 and K-4 VLBI data acquisition terminals," J. Commun. Res. Lab., vol. 38, no. 3, pp. 435-457, Nov. 1991.

[4] ANSI (American National Standard Institute) X3B.61 ID-1 format

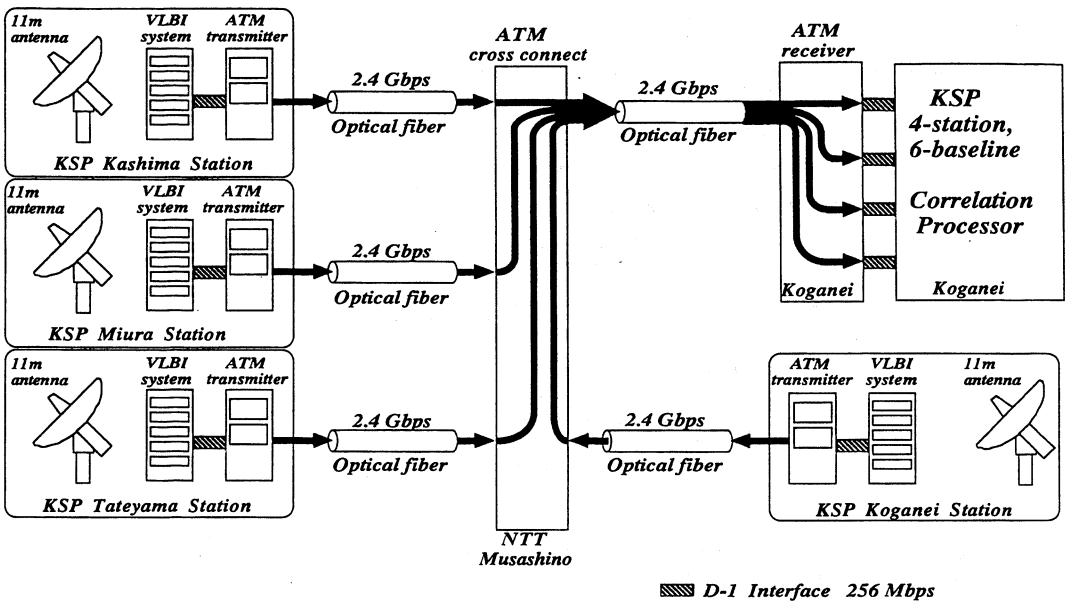


Figure 1 A block diagram of the KSP real-time VLBI system

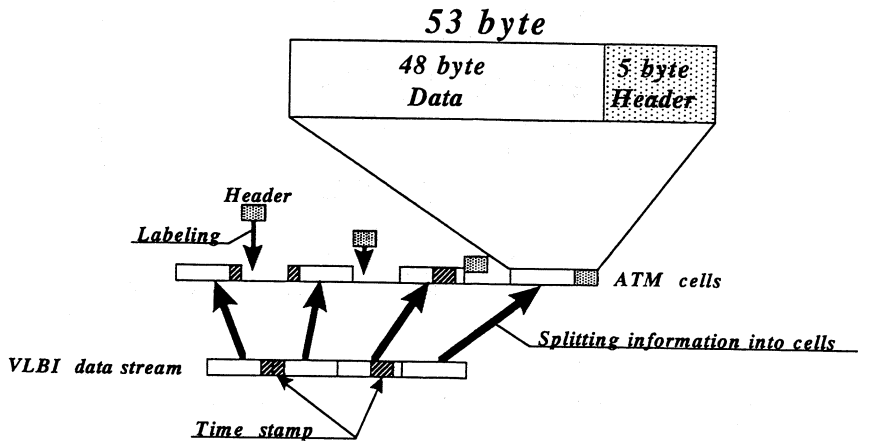


Figure 2 ATM cell

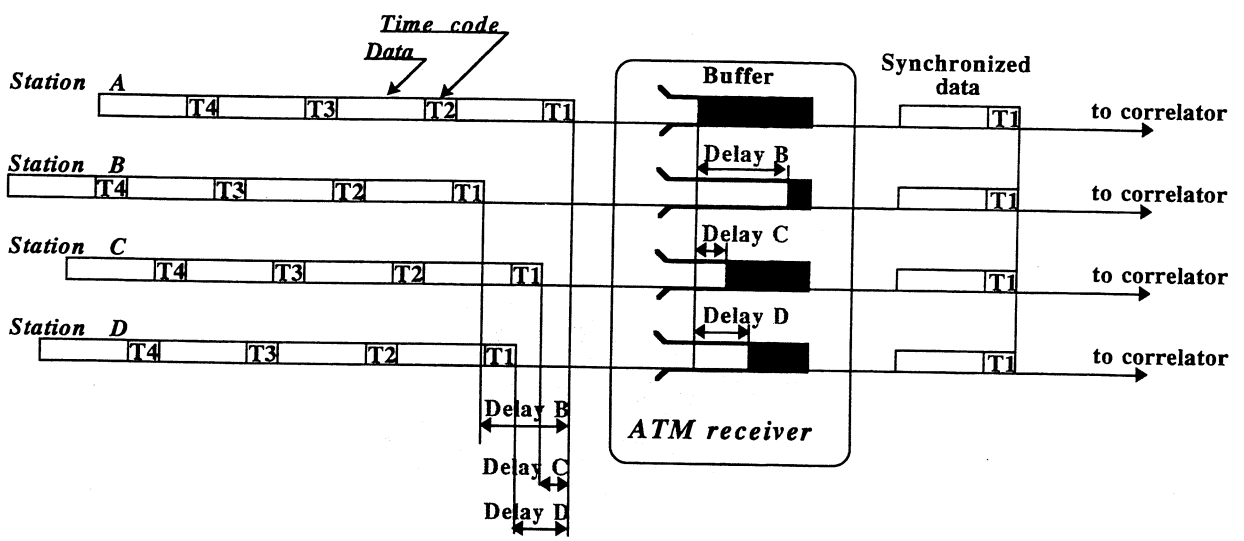


Figure 3 Data synchronization

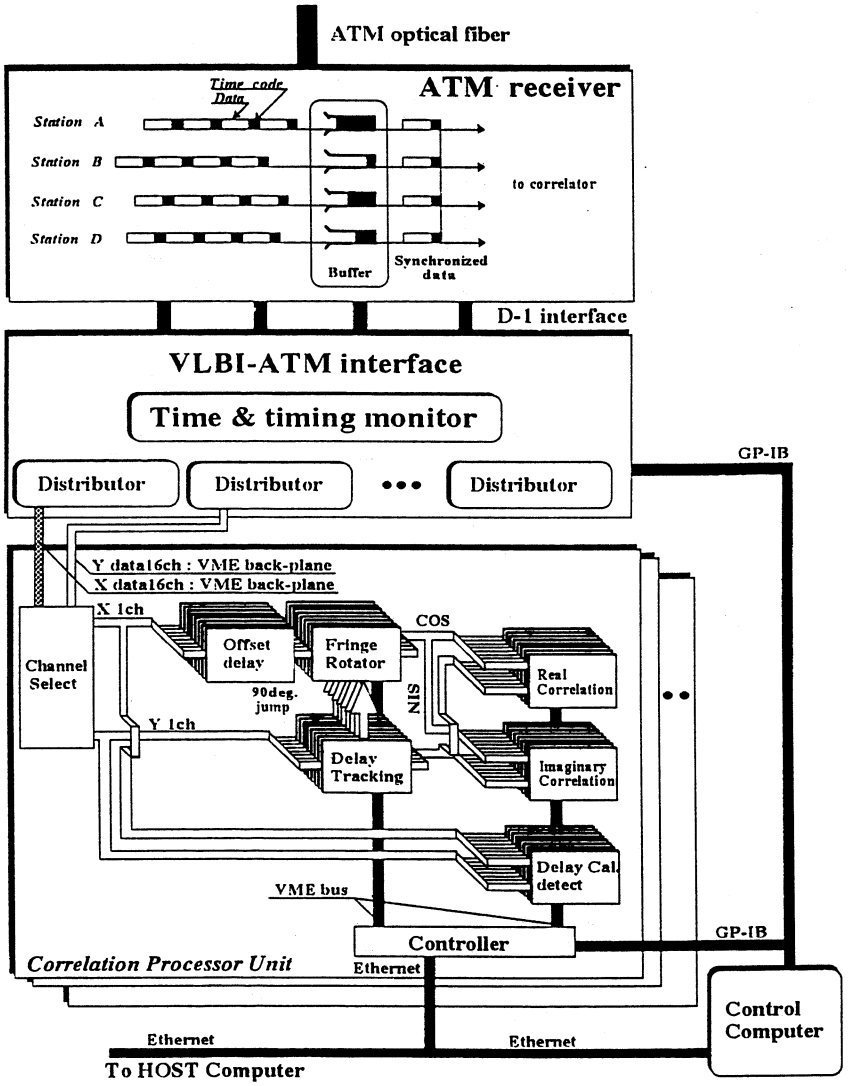


Figure 4 A block diagram of the KSP real-time correlation processing system

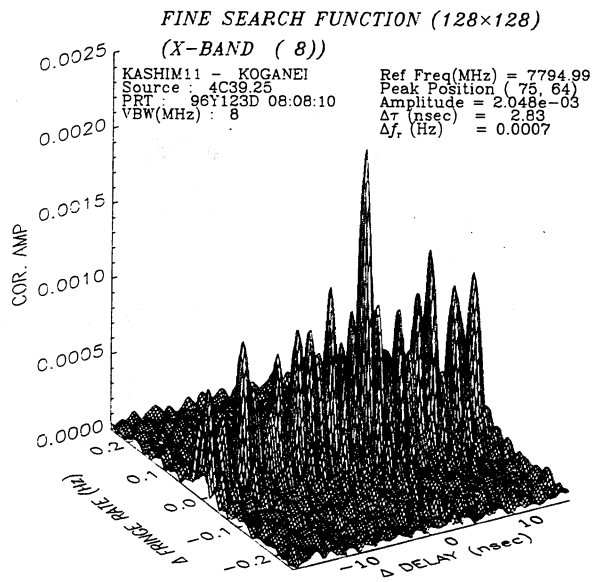


Figure 5 Obtained fringe (in fine delay function) ; data rate 256 Mbps

ESTIMATION OF THE UNCERTAINTY OF A MEAN FREQUENCY MEASUREMENT

F. Vernotte, M. Vincent
Observatoire de Besançon
41 bis, avenue de l'Observatoire - B. P. 1615 - 25010 BESANÇON Cédex - France
Phone : (33) 03.81.66.69.22 - Fax : (33) 03.81.66.69.44 - E-mail : Francois@Obs-Besancon.Fr

Abstract

It is well known that, in the case of a white frequency noise, the standard deviation of a sequence of mean frequency measurements decreases as the inverse of the square root of the number of samples. Consequently, the uncertainty of one mean frequency measurement decreases as the inverse of the square root of the average time. Thus, knowing the level of the white frequency noise and the average time, it is easy to estimate the uncertainty of one mean frequency measurement. However, for the other types of noise (from f^{-2} to f^{+2} frequency noises), such an estimation may be much more difficult, particularly in the case of non-stationary noises (f^{-1} and f^{-2} frequency noises).
Using the transfer function of the true variance, we solved this problem for all the types of noises. The details of this approach will be described in this paper. This method yields an estimation of the uncertainty of a mean frequency measurement knowing the different frequency noise levels, the duration of the average, and the number of mean frequency samples.

Keywords : *frequency measurement, uncertainty, low frequency noises, non-stationarity, low cut-off frequency.*

1. MEAN FREQUENCY MEASUREMENT

In order to characterize the frequency of an oscillator, we perform a sequence of N consecutive frequency measurements f_k : $\{f_1, f_2, \ldots, f_N\}$.
Each f_k is the mean frequency over an integration time τ_0 :

$$f_k = \frac{1}{\tau_0} \int_{t_k}^{t_k + \tau_0} f(t)\, dt$$

f(t) is the instantaneous frequency of the oscillator.

What is the uncertainty domain about the estimation of f_0, the nominal frequency of this oscillator ?
How this uncertainty domain is linked to the standard deviation of the f_k's ?

Let us define the instantaneous normalized frequency deviation samples $\overline{y_k}$:

$$\overline{y_k} = \frac{f_k - f_0}{f_0}$$

$\overline{y_k}$ is a dimensionless centered random variable.

We assume that its power spectral density $S_y(f)$ is modeled by a sum of power laws :

$$S_y(f) = \sum_{\alpha=-2}^{+2} h_\alpha f^\alpha$$

What relationship exists between these noise levels h_α and the uncertainty of f_0 ?

2. THE TRUE VARIANCE

The transfer function of the true variance is :

$$|H_y(f)|^2 = \frac{\sin^2(\pi \tau f)}{\pi^2 \tau^2 f^2}$$

where $\tau = N \tau_0$ is the duration of the sequence.
The result of the true variance applied to a sequence of which the spectral density is $S_y(f)$, may be calculated as :

$$\sigma_T^2(\tau) = \int_{f_l}^{f_h} S_y(f)\, |H_y(f)|^2\, df$$

f_l and f_h are the low and high cut-off frequencies.

Table I shows that the true variance only

$S_y(f)$	$h_{-2} f^{-2}$	$h_{-1} f^{-1}$	$h_0 f^0$	$h_{+1} f^{+1}$	$h_{+2} f^{+2}$
True variance	$h_{-2}\left(\frac{2}{f_l} - \frac{2\pi^2\tau}{3}\right)$	$h_{-1}\left(\frac{3}{2} - C - \ln(2 f_l \pi \tau)\right)$	$\frac{h_0}{2\tau}$	$h_{+1} \frac{C + \ln(2 f_h \pi \tau)}{2\pi^2\tau^2}$	$h_{+2} \frac{f_h}{2\pi^2\tau^2}$

Table I : Response of the true variance for the different types of noise. These responses are approximated except for the white frequency noise.

converges for the white frequency noise. This result is proportional to $1/\tau$: the standard deviation of the $\overline{y_k}$'s decreases as $\sqrt{N}$.

For high frequency noises, the result depends on f_h but remains close to the result of the Allan variance.
For low frequency noises, not only the result depends on f_l but we observed that the measured true variance of simulated f^{-2} frequency noise sequences may be completely different from the value given by table I.
What is the physical meaning of the low cut-off frequency f_l ?
What is the difference between the theoretical variance and the experimental variance ?

3. CASE OF THE LOW FREQUENCY NOISES

3. 1. Bias and random fluctuations

Figure 1 shows a sequence of N = 1024 $\overline{y_k}$'s samples of f^{-2} frequency noise. The low cut-off frequency of this noise is close to the inverse of the duration of this sequence. Let us notice y_0, the average of the $\overline{y_k}$'s samples and σ their standard deviation :
$N \times \tau_0 \approx 1/f_l \Rightarrow y_0 \approx 0$ **and** $\sigma = 68$
(representative of the $\overline{y_k}$'s fluctuations).

An enlargement of this figure is shown below. In this case, N' = 100, thus :
$N' \times \tau_0 << 1/f_l \Rightarrow y_0 = 80$ **and** $\sigma = 11$
(not representative of the fluctuations of the whole $\overline{y_k}$'s sequence).

A f^{-2} frequency noise is characterized by very long term fluctuations (period about $1/f_l$) with high amplitude. Considering a short sequence, these fluctuations are seen as a constant (y_0) and the standard deviation is underestimated. But, for different subsequences, this constant term y_0 is a random variable.

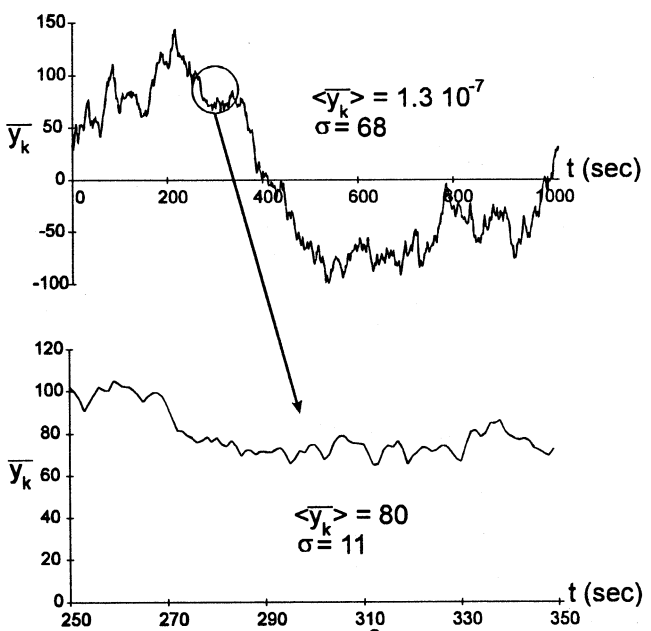


Figure 1 : Sequence of f^{-2} frequency noise.

How does this effect influence the mean frequency measurement ?
Let us consider that this sequence of 100 samples was provided by an oscillator with an unknown nominal frequency f_0.
Each f_k measurement may be written as :

$$f_k = f_0 (1 + \overline{y_k}).$$

The mean frequency is then : $< f_k > = f_0 (1 + y_0)$
and the standard deviation : $\sigma_f = f_0 \sigma$.
The use of the mean frequency as a measurement of the nominal frequency yields a bias.
The standard deviation underestimates the measurement uncertainties.

How could we model the statistics of the $\overline{y_k}$'s samples ?
The $\overline{y_k}$'s samples may be rewritten as :

$$\overline{y_k} = y_0 + \varepsilon_k$$

where ε_k is a centered random variable representing the fluctuations over the studied sequence. Obviously, the statistics of the $\overline{y_k}$'s samples does not depend on $\tau = N \tau_0$, the duration of the sequence. On the other hand, the statistics of y_0 and ε_k depends on τ :

$\tau << 1/f_b \Rightarrow \varepsilon_k \to 0 \quad \overline{y_k} \equiv y_0$

$\tau \geq 1/f_b \Rightarrow y_0 \to 0 \quad \overline{y_k} \equiv \varepsilon_k$.

The formula given in table I is the variance of y_0.

3. 2. Geometrical interpretation

Let us consider the vectors :

$$\vec{Y} = \begin{pmatrix} \overline{y_1} \\ \dots \\ \overline{y_N} \end{pmatrix}, \quad \vec{Y_0} = \begin{pmatrix} y_0 \\ \dots \\ y_0 \end{pmatrix} \quad \text{and} \quad \vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \dots \\ \varepsilon_N \end{pmatrix}$$

These vectors are linked by the relationship :

$$\vec{Y} = \vec{Y_0} + \vec{\varepsilon}$$

Moreover, the square of the modulus of $\vec{Y}$ is :

$$\|\vec{Y}\|^2 = \|\vec{Y_0} + \vec{\varepsilon}\|^2 = \sum_{k=1}^{N} (y_0 + \varepsilon_k)^2$$

$$\|\vec{Y}\|^2 = N y_0^2 + \sum_{k=1}^{N} \varepsilon_k^2 + 2 y_0 \sum_{k=1}^{N} \varepsilon_k$$

Since ε_k is centered, it comes

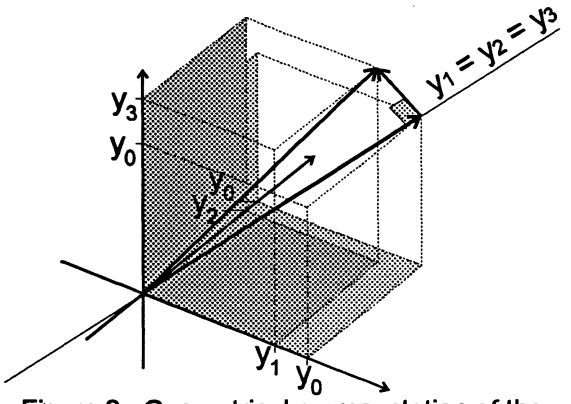


Figure 2 : Geometrical representation of the vectors $\vec{Y}$, $\vec{Y_0}$ and $\vec{\varepsilon}$ for 3 dimensions.

$$\|\vec{Y}\|^2 = N y_0^2 + \sum_{k=1}^{N} \varepsilon_k^2$$

$$\|\vec{Y}\|^2 = \|\vec{Y_0}\|^2 + \|\vec{\varepsilon}\|^2$$

Thus, $\vec{Y_0}$ **and** $\vec{\varepsilon}$ **are orthogonal :** $\vec{Y_0}$ **is the projection of** $\vec{Y}$ **onto the straight line defined by** $y_1 = y_2 = ... = y_N$ (see figure 2).

3. 3. Case of f^{-2} frequency noise

Since $y_0 = \overline{y_k} - \varepsilon_k$,

$$\sigma_{y0}^2(\tau) = \frac{2}{f_b} - \frac{2\pi^2\tau}{3} = \sigma_{yk}^2 - \sigma_{\varepsilon k}^2(\tau)$$

with σ_{yk}^2 independent of τ and $\sigma_{\varepsilon k}^2(\tau)$ depending on τ.
By identification, we obtain :

$$\sigma_{yk}^2(\tau) = \frac{2}{f_b} \text{ and } \sigma_{\varepsilon k}^2(\tau) = \frac{2\pi^2\tau}{3}$$

(for $\tau << 1/f_b$).

Thus, the standard deviation of such a noise, $\sigma_{\varepsilon k}(\tau)$, increases as $\sqrt{\tau}$. But, $\sigma_{\varepsilon k}(\tau)$ is not an estimator of the measurement uncertainties. The uncertainty is due to the bias term y_0 : the measured mean frequency $\langle f_k \rangle$ is equal to the nominal frequency f_0 plus or minus the bias term $\frac{y_0 \langle f_k \rangle}{1 + y_0} \approx y_0 \langle f_k \rangle$.
The uncertainty domain of the mean frequency measurement is then :

$$\sigma_{f0}(\tau) = \sqrt{h_{-2}\left(\frac{2}{f_b} - \frac{2\pi^2\tau}{3}\right)} \langle f_k \rangle$$

(for $\tau << 1/f_b$).

It decreases weakly as τ increases.

3. 4. Case of f^{-1} frequency noise

$$\sigma_{y0}^2(\tau) = K_0 - \ln(f_b \tau) = \sigma_{yk}^2 - \sigma_{\varepsilon k}^2(\tau).$$

Using a Monte-Carlo simulation, we obtained :

$$\sigma_{yk}^2(\tau) = K_1 - \ln(f_b)$$

and $\sigma_{\varepsilon k}^2(\tau) = K_2 + \ln(\tau)$

with $K_1 = -0.45 \pm 0.06$ and $K_2 = -0.16 \pm 0.05$.
Thus, the uncertainty domain of the mean frequency measurement is then :

$$\sigma_{f0}(\tau) = \sqrt{h_{-1}\left(\frac{3}{2} - C - \ln(2 f_b \pi \tau)\right)} \langle f_k \rangle$$

(for $\tau << 1/f_b$).

3. 5. Case of the other types of noise

Because of their stationarity, no low cut-off frequency is needed to consider the case of the other types of noise. This means that no bias will affect the mean frequency measurement. Let us denote $\Phi_\alpha(\tau)$ the response of the true variance for a f^α frequency noise with a level equal to one (see table I). The measured mean frequency $\langle f_k \rangle$ is a random variable, centered around the nominal frequency f_0 with a standard deviation :

$$\sigma_{f0}(\tau) = \sqrt{h_\alpha \Phi_\alpha(\tau)} \langle f_k \rangle$$

Thus, for an oscillator affected by different types of noises, the knowledge of the different noise levels allows us to estimate the uncertainty of the determination of the nominal frequency :

$$f_0 = \langle f_k \rangle \pm \langle f_k \rangle \sqrt{\sum_{\alpha=-2}^{+2} h_\alpha \Phi_\alpha(\tau)}$$

(see $\Phi_\alpha(\tau)$ **in table I).**

4. CONCLUSION : PHYSICAL MEANING OF THE LOW CUT-OFF FREQUENCY

In the case of a low frequency noise, the results above show that the standard deviation of the mean frequency depends explicitly on the low cut-off frequency.

What is the low cut-off frequency of an oscillator : the age of the oscillator, its life time, the age of the Universe, ... ?

Let us consider a f^{-2} frequency noise. It's derivative, the aging z(t), is then a white noise. The instantaneous normalized frequency deviation y(t) is then :

$$y(t) = \int_{t_0}^{t} z(\theta)\, d\theta$$

where t_0 is the instant of setting up the oscillator. If y(t) is defined as the integration of z(t), no low cut-off frequency is needed to insure the finite power of the signal.
y(t) is then a centered random variable of which standard deviation will increases with time, i. e. with the age of the oscillator. There is an analogy between the age of the oscillator and the inverse of the cut-off frequency.
In table I, one could replace the low cut-off frequency f_l by the inverse of the age Θ of the oscillator :

$$f_I = \frac{1}{\Theta}$$

What is the real frequency of an oscillator during a short (relatively to the age of the oscillator) sequence : its nominal frequency or the mean frequency during that sequence ?

The answer depends also on the model we chose. Considering the power law model of the power spectral density, the frequency is the nominal frequency of the oscillator, because the use of this model implies to define the parameters of this model (the nominal frequency, the noise levels, etc...) in the frequency domain : these quantities are then time independent. Thus the uncertainty we calculated above means : since its setting up, the frequency of the oscillator was included in this uncertainty domain.

However, it seems logical to consider the frequency of an oscillator only during the short sequence : we define the frequency as a quantity which depends on time.
In this assumption, the frequency of the oscillator during a sequence of duration τ is the measured mean frequency and its standard deviation is $\sigma_{\varepsilon k}(\tau)$ (see above).

For a f^{-2} frequency noise as well as a f^{-1} frequency noise, this standard deviation increases with the duration of the sequence :
the longer the duration of the sequence, the larger the uncertainty domain !
This is due to the long term fluctuations.

5. REFERENCES

[1] F. Vernotte, "Stabilité temporelle des oscillateurs : nouvelles variances, leurs propriétés, leurs applications", PhD thesis, Université de Franche-Comté, Besançon, 1991.

[2] G. Zalamansky, "Recherche d'un bruit de fond d'ondes gravitationnelles dans les résidus de chronométrage des pulsars milliseconde", PhD thesis, Université de Paris VII, Paris, 1996.

[3] J. Rutman, "Characterization of phase and frequency instabilities in precision frequency sources : fifteen years of progress", Proceedings of the IEEE, vol. 66, N# 9, September 1978, pp. 1048-1075.

[4] G. Parreins, "Techniques statistiques, moyens rationels de choix et de décision", Dunod technique, Paris, 1974.

On the dissipation effect of the random transitions perturbations

N.Yahyabey (1) and S. Hassani (2)

(1) Université François Rabelais de Tours, Faculté des sciences, Parc de grandmont, 37200 Tours, France.

(2) Centre de developpement des techniques nucléaires, 2 Bld Frantz Fanon, Alger, Algeria.

A proposition that handles and computes the r^{th} order spectrum of non Fourier-transformable functions because of perturbations affecting their periodicities is analysed. By means of a method based on the concept of the overall period, the computed r^{th} order spectrum shows a dependence on the Fourier transform of the distribution function of the perturbations. For absolutely continuous distribution functions with respect to Lebesgue measure, the average amplitude spectrum (first order spectrum) is reduced and the power spectral density (second order spectrum), subject to dissipation, exhibits neither a continuous background nor a continuum spectral component. The reduction factor is linked to the Fourier transform of the probability density function. This important property can be carried out in perturbation characterization process. Also, the random transitions perturbations induce no enhancements of spectral lines widths.

An **experimental** test of this theoretical proposition is presented. It consists of a sinusoidal oscillator coupled to a triggering system operating with two threshold detection levels. The gaussian noise signal present at the control input of the triggering system is used to define randomly the two threshold detection levels for the perturbation of the square wave transitions. Experimental evidence of the reduction effect with no enhancement of existant spectral line widths is shown and the reduction factor depends on the Fourier transform of the probability density function of the random variables measuring the fluctuations at the nominal commutation instants. Also, the measured spectral lines are subject to power dissipation. The obtained experimental results are in agreement with the proposition theoretical predictions.

MULTICHANNEL DISCRIMINATORS FOR HIGH ACCURANCY FREQUENCY AND TIME EVALUATION IN LARGE ENERGY RANGE OF EFFECTIVE SIGNAL

Kirill V. Zaichenko
STATE ACADEMY OF AEROSPACE INSTRUMENTATION, ST. PETERSBURG, RUSSIA

Multichannel evaluation procedures are being inherent result of optimum measures synthesis for arbitrary parameter in conditions of extended by this parameter informative signal. In order to realize optimum procedures in time and frequency areas we have to obtain several readings of informative process in order to achive the highest accurancy in estimation the assigned a priori sector of this informative process signal release. It is known that the optimum procedure demands plausible function maximum calculation by its differentiation. Optimum procedure synthesis indicated that the admitted algorithms can be found in ratio aspect of two polynomials which are being informative process readings weight sums. Number of parameter readings choise has to be concerted with system resolution ability. For example - optimum procedure for three-channel measure is being the ratio of a difference between two lateral readings and a weight sum of all three readings. Denominator of this ratio performs a standardization function ensures so parameter appreciation independence in large scale of signal-noize energy ratio. As the synthesis result were obtained the admitted algorithms of three- and five-channel discriminators for frequency and time measures. In time area evaluation is proceeded for time delay in regard to some supporting value while at frequency area - for average frequency value of informative process spectrum density. The analysis of admitted schemes for time and frequency discriminators based upon the multichannel principle registered their accurancy superiority in comparison with commonly used at present two-channel discriminators. Proposed discriminators advantage is manifested particulary strong in violent variations of signal-noize ratio that is urgent for radionics and radar systems which are characterized by a large energy range of effective signal.

Revision of long-term stability theory of Rb Frequency standard

Agap'ev B.D., Khutorschikov V.I.

St.Petersburg State Technical University

The history of the light shift research was analyzed. It was shown that for last 30 years progress in the long term instability reduction of Rb frequency standarts is small.

This fact may be partially conditioned on the disregard of the random character of the pumping light intensity fluctuations and not account of their correlation properties.

In the present work fluctuation properties of the pumping light are consequentially accounted and the light fluctuations transformation into the frequency fluctuations of the atomic transitions are studied.

It is shown that conditions of reduction of the light shifts due to the light radiation fluctuations qualitativel differ from the basic recomendations. Besides usual recomendations may result to the increase of the long term instability of atomic standart frequency.

Development of a Digital Cesium Clock for Space at the U.S. Naval Research Laboratory

J. White*, W. M. Golding**, E. Powers*, J. Brad*, F. Danzy*, A. Frank**, R. Beard*,

*U.S. Naval Research Laboratory, Washington, DC
**SFA, Largo, MD

1. ABSTRACT

The Naval Research Laboratory has been involved in the development of space-qualified atomic clocks since the early 1970's, providing support to the Global Positioning System (GPS). The goal of the current effort is to bring some aspects of the digital cesium technology that are readily available in commercial atomic clocks (Refs. 1,2) to space. Digital cesium clock technology is particularly valuable in space applications because it can provide enhanced manufacturability, improved operability, and decreased environmental sensitivity. The progress and future direction of this effort are described.

Keywords: digital atomic clocks, GPS cesium clocks

2. BACKGROUND

GPS began flying atomic clocks in space in 1974 with the Naval Research Laboratory's Navigation Technology Satellite I (Ref. 3), the first launch for the Tri-service GPS program. Today, GPS has the operational 24-satellite constellation in orbit and each satellite carries 4 atomic clocks. The GPS experience has provided considerable insight into the most important features required of space clocks. While clock stability is always an essential performance factor in a GPS clock, reliability and diagnostic capability are very important to the operation of the overall GPS system.

For example, some of the GPS Block II cesium clocks have shown excellent frequency stability while others have exhibited performance problems (Ref. 4). The biggest problem that the GPS operators face with these clocks is determining whether on-orbit anomalies are an indication of failure, degradation, or impending failure. Few of these units have experienced catastrophic failure, but rather performance anomalies which could affect system performance or the ability to maintain it. In orbit, the clock's RF signal output cannot be seen directly. The GPS signal is based on the clock but is affected by synthesizers, transmitters, signal propagation and receivers. The telemetry monitors provided to the GPS control segment are essentially those available on the front panel of a 1960's vintage commercial clock. Those monitors can provide insight into the clock's health but are limited in scope and availability. While a laboratory user can combine continuous monitor data with high-resolution phase data and has the opportunity to open the covers to make internal diagnostic measurements, the space operator has very limited data. Even during the satellite vehicle integration and testing phases, the manufacturer must work with limited data to verify that a clock is healthy and should be launched. Temperature sensitivity is another major variable in the Block II cesiums. The clocks were built to a specification of 1×10^{-13} per degree C. With a typical 2-degree peak-to-peak swing in baseplate temperature for a GPS clock, the temperature effect is a significant contributor to the one-day stability.

The GPS program office has also had concerns about the long-term viability of clock suppliers, since the cesium suppliers have been small businesses. In order to minimize the dependence on any particular manufacturer, it is desirable to create an open clock architecture that, if necessary, could be supported with critical clock components from more than one vendor.

The digital cesium clocks now being manufactured for ground or commercial use solve many of these problems. They have demonstrated improved stability and a decreased need for operator interaction. A benefit of the digital servo is to provide a more comprehensive set of monitors that allows the user to have a better understanding of the unit's state-of-health. Digital techniques can be produced more consistently since the reliance on critical and difficult to adjust analog circuitry is much reduced. Due to their advanced digital servo scheme and a re-designed beam tube, the Hewlett Packard 5071A has demonstrated nearly zero coefficients for temperature and magnetic field effects (Refs. 5,6). Because of the capabilities of digital servos, digital clocks also have the potential to provide improved reliability.

3. DESIGN GOALS

Based on the GPS Block II operational experience, the NRL goals for space-based digital cesium clocks are:

3.1 Frequency Stability

The new clocks must meet or exceed the level of frequency stability attained by the present generation of GPS cesium clocks. These clocks have met or exceeded their design requirements for stability, and the GPS system is operating within its specification using them (Ref. 4). The inherent accuracy, drift free behavior and long term stability obtainable with beam tube based cesium clocks are important to the ability of GPS to maintain system time, navigation performance and transfer accurate time. These qualities must be maintained in the next generation of space cesium clocks.

3.2 Manufacturability

The new generation of clocks must be easier to build. Analog style clocks have been difficult and time consuming to manufacture. The difficulty of the setup procedure requires a high level of training and produces a wide level of variation in performance from unit to unit. Precise measurements are required, and each clock must be observed and tested for long periods to verify correct operation. Because of the low-level signals involved, great care must be taken to prevent grounding problems and leakage of critical signals between input and output signal paths. The analog modulation used in the servos requires exceptional symmetry to eliminate frequency offsets and the associated degradation in stability. Another problem area has been the generation of a clean microwave spectrum. Traditional techniques for producing the interrogation signal also produce large amplitude sidebands within 20 MHz of the carrier. Strong sidebands and other spurious signals in the interrogation signal cause the

ESTIMATION OF THE UNCERTAINTY OF A MEAN FREQUENCY MEASUREMENT

F. Vernotte, M. Vincent
Observatoire de Besançon
41 bis, avenue de l'Observatoire - B. P. 1615 - 25010 BESANÇON Cédex - France
Phone : (33) 03.81.66.69.22 - Fax : (33) 03.81.66.69.44 - E-mail : Francois@Obs-Besancon.Fr

Abstract

It is well known that, in the case of a white frequency noise, the standard deviation of a sequence of mean frequency measurements decreases as the inverse of the square root of the number of samples. Consequently, the uncertainty of one mean frequency measurement decreases as the inverse of the square root of the average time. Thus, knowing the level of the white frequency noise and the average time, it is easy to estimate the uncertainty of one mean frequency measurement. However, for the other types of noise (from f^{-2} to f^{+2} frequency noises), such an estimation may be much more difficult, particularly in the case of non-stationary noises (f^{-1} and f^{-2} frequency noises).
Using the transfer function of the true variance, we solved this problem for all the types of noises. The details of this approach will be described in this paper. This method yields an estimation of the uncertainty of a mean frequency measurement knowing the different frequency noise levels, the duration of the average, and the number of mean frequency samples.

Keywords : *frequency measurement, uncertainty, low frequency noises, non-stationarity, low cut-off frequency.*

1. MEAN FREQUENCY MEASUREMENT

In order to characterize the frequency of an oscillator, we perform a sequence of N consecutive frequency measurements f_k : $\{f_1, f_2, \ldots, f_N\}$.
Each f_k is the mean frequency over an integration time τ_0 :

$$f_k = \frac{1}{\tau_0} \int_{t_k}^{t_k + \tau_0} f(t)\, dt$$

f(t) is the instantaneous frequency of the oscillator.

What is the uncertainty domain about the estimation of f_0, the nominal frequency of this oscillator ?
How this uncertainty domain is linked to the standard deviation of the f_k's ?

Let us define the instantaneous normalized frequency deviation samples $\overline{y_k}$:

$$\overline{y_k} = \frac{f_k - f_0}{f_0}$$

$\overline{y_k}$ is a dimensionless centered random variable.

We assume that its power spectral density $S_y(f)$ is modeled by a sum of power laws :

$$S_y(f) = \sum_{\alpha=-2}^{+2} h_\alpha f^\alpha$$

What relationship exists between these noise levels h_α and the uncertainty of f_0 ?

2. THE TRUE VARIANCE

The transfer function of the true variance is :

$$|H_y(f)|^2 = \frac{\sin^2(\pi \tau f)}{\pi^2 \tau^2 f^2}$$

where $\tau = N \tau_0$ is the duration of the sequence.
The result of the true variance applied to a sequence of which the spectral density is $S_y(f)$, may be calculated as :

$$\sigma_T^2(\tau) = \int_{f_l}^{f_h} S_y(f)\, |H_y(f)|^2\, df$$

f_l and f_h are the low and high cut-off frequencies.

Table I shows that the true variance only

$S_y(f)$	$h_{-2} f^{-2}$	$h_{-1} f^{-1}$	$h_0 f^0$	$h_{+1} f^{+1}$	$h_{+2} f^{+2}$
True variance	$h_{-2}\left(\frac{2}{f_l} - \frac{2\pi^2\tau}{3}\right)$	$h_{-1}\left(\frac{3}{2} - C - \ln(2 f_l \pi \tau)\right)$	$\frac{h_0}{2\tau}$	$h_{+1} \frac{C + \ln(2 f_h \pi \tau)}{2\pi^2\tau^2}$	$h_{+2} \frac{f_h}{2\pi^2\tau^2}$

Table I : Response of the true variance for the different types of noise. These responses are approximated except for the white frequency noise.

converges for the white frequency noise. This result is proportional to $1/\tau$: the standard deviation of the $\overline{y_k}$'s decreases as $\sqrt{N}$.

For high frequency noises, the result depends on f_h but remains close to the result of the Allan variance.
For low frequency noises, not only the result depends on f_l but we observed that the measured true variance of simulated f^{-2} frequency noise sequences may be completely different from the value given by table I.
What is the physical meaning of the low cut-off frequency f_l ?
What is the difference between the theoretical variance and the experimental variance ?

3. CASE OF THE LOW FREQUENCY NOISES

3. 1. Bias and random fluctuations

Figure 1 shows a sequence of N = 1024 $\overline{y_k}$'s samples of f^{-2} frequency noise. The low cut-off frequency of this noise is close to the inverse of the duration of this sequence. Let us notice y_0, the average of the $\overline{y_k}$'s samples and σ their standard deviation :
$N \times \tau_0 \approx 1/f_l \Rightarrow y_0 \approx 0$ **and** $\sigma = 68$
(representative of the $\overline{y_k}$'s fluctuations).

An enlargement of this figure is shown below. In this case, N' = 100, thus :
$N' \times \tau_0 << 1/f_l \Rightarrow y_0 = 80$ **and** $\sigma = 11$
(not representative of the fluctuations of the whole $\overline{y_k}$'s sequence).

A f^{-2} frequency noise is characterized by very long term fluctuations (period about $1/f_l$) with high amplitude. Considering a short sequence, these fluctuations are seen as a constant (y_0) and the standard deviation is underestimated. But, for different subsequences, this constant term y_0 is a random variable.

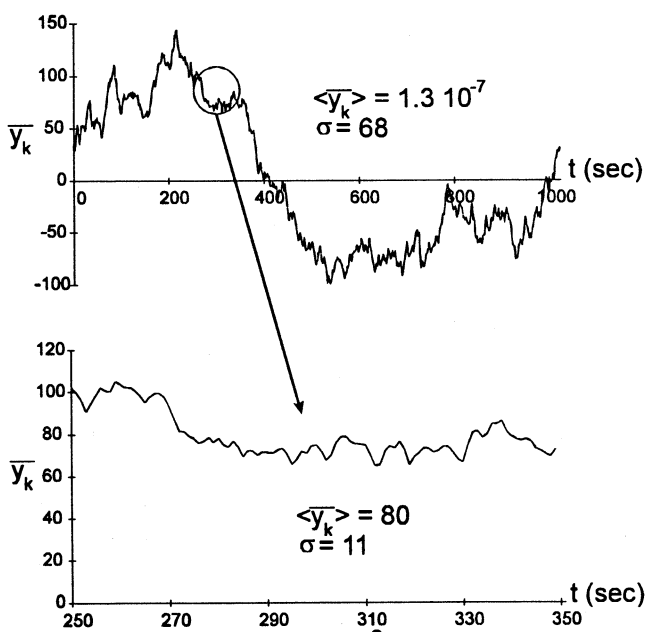


Figure 1 : Sequence of f^{-2} frequency noise.

How does this effect influence the mean frequency measurement ?
Let us consider that this sequence of 100 samples was provided by an oscillator with an unknown nominal frequency f_0.
Each f_k measurement may be written as :

$$f_k = f_0 (1 + \overline{y_k}).$$

The mean frequency is then : $< f_k > = f_0 (1 + y_0)$
and the standard deviation : $\sigma_f = f_0 \sigma$.
The use of the mean frequency as a measurement of the nominal frequency yields a bias.
The standard deviation underestimates the measurement uncertainties.

How could we model the statistics of the $\overline{y_k}$'s samples ?
The $\overline{y_k}$'s samples may be rewritten as :

$$\overline{y_k} = y_0 + \varepsilon_k$$

where ε_k is a centered random variable representing the fluctuations over the studied sequence. Obviously, the statistics of the $\overline{y_k}$'s samples does not depend on $\tau = N \tau_0$, the duration of the sequence. On the other hand, the statistics of y_0 and ε_k depends on τ :

$\tau << 1/f_b \Rightarrow \varepsilon_k \to 0 \quad \overline{y_k} \equiv y_0$

$\tau \geq 1/f_b \Rightarrow y_0 \to 0 \quad \overline{y_k} \equiv \varepsilon_k$.

The formula given in table I is the variance of y_0.

3. 2. Geometrical interpretation

Let us consider the vectors :

$$\vec{Y} = \begin{pmatrix} \overline{y_1} \\ \dots \\ \overline{y_N} \end{pmatrix}, \quad \vec{Y_0} = \begin{pmatrix} y_0 \\ \dots \\ y_0 \end{pmatrix} \quad \text{and} \quad \vec{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \dots \\ \varepsilon_N \end{pmatrix}$$

These vectors are linked by the relationship :

$$\vec{Y} = \vec{Y_0} + \vec{\varepsilon}$$

Moreover, the square of the modulus of $\vec{Y}$ is :

$$\|\vec{Y}\|^2 = \|\vec{Y_0} + \vec{\varepsilon}\|^2 = \sum_{k=1}^{N} (y_0 + \varepsilon_k)^2$$

$$\|\vec{Y}\|^2 = N y_0^2 + \sum_{k=1}^{N} \varepsilon_k^2 + 2 y_0 \sum_{k=1}^{N} \varepsilon_k$$

Since ε_k is centered, it comes

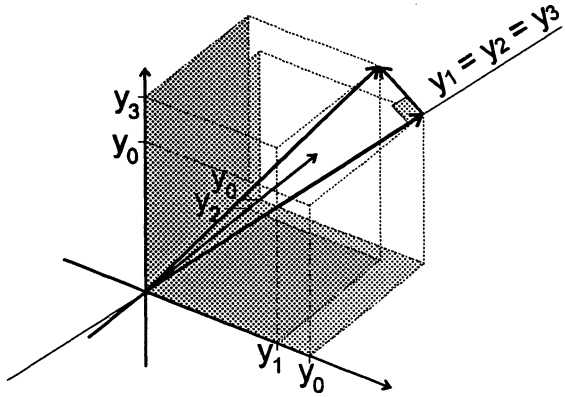


Figure 2 : Geometrical representation of the vectors $\vec{Y}$, $\vec{Y_0}$ and $\vec{\varepsilon}$ for 3 dimensions.

$$\|\vec{Y}\|^2 = N y_0^2 + \sum_{k=1}^{N} \varepsilon_k^2$$

$$\|\vec{Y}\|^2 = \|\vec{Y_0}\|^2 + \|\vec{\varepsilon}\|^2$$

Thus, $\vec{Y_0}$ **and** $\vec{\varepsilon}$ **are orthogonal :** $\vec{Y_0}$ **is the projection of** $\vec{Y}$ **onto the straight line defined by** $y_1 = y_2 = ... = y_N$ (see figure 2).

3. 3. Case of f^{-2} frequency noise

Since $y_0 = \overline{y_k} - \varepsilon_k$,

$$\sigma_{y0}^2(\tau) = \frac{2}{f_b} - \frac{2\pi^2\tau}{3} = \sigma_{yk}^2 - \sigma_{\varepsilon k}^2(\tau)$$

with σ_{yk}^2 independent of τ and $\sigma_{\varepsilon k}^2(\tau)$ depending on τ.
By identification, we obtain :

$$\sigma_{yk}^2(\tau) = \frac{2}{f_b} \text{ and } \sigma_{\varepsilon k}^2(\tau) = \frac{2\pi^2\tau}{3}$$

(for $\tau << 1/f_b$).

Thus, the standard deviation of such a noise, $\sigma_{\varepsilon k}(\tau)$, increases as $\sqrt{\tau}$. But, $\sigma_{\varepsilon k}(\tau)$ is not an estimator of the measurement uncertainties. The uncertainty is due to the bias term y_0 : the measured mean frequency $\langle f_k \rangle$ is equal to the nominal frequency f_0 plus or minus the bias term $\frac{y_0 \langle f_k \rangle}{1 + y_0} \approx y_0 \langle f_k \rangle$.
The uncertainty domain of the mean frequency measurement is then :

$$\sigma_{f0}(\tau) = \sqrt{h_{-2}\left(\frac{2}{f_b} - \frac{2\pi^2\tau}{3}\right)} \langle f_k \rangle$$

(for $\tau << 1/f_b$).

It decreases weakly as τ increases.

3. 4. Case of f^{-1} frequency noise

$$\sigma_{y0}^2(\tau) = K_0 - \ln(f_b\,\tau) = \sigma_{yk}^2 - \sigma_{\varepsilon k}^2(\tau).$$

Using a Monte-Carlo simulation, we obtained :

$$\sigma_{yk}^2(\tau) = K_1 - \ln(f_b)$$

and $\sigma_{\varepsilon k}^2(\tau) = K_2 + \ln(\tau)$

with $K_1 = -0.45 \pm 0.06$ and $K_2 = -0.16 \pm 0.05$.
Thus, the uncertainty domain of the mean frequency measurement is then :

$$\sigma_{f0}(\tau) = \sqrt{h_{-1}\left(\frac{3}{2} - C - \ln(2 f_b \pi \tau)\right)} \langle f_k \rangle$$

(for $\tau << 1/f_b$).

3. 5. Case of the other types of noise

Because of their stationarity, no low cut-off frequency is needed to consider the case of the other types of noise. This means that no bias will affect the mean frequency measurement. Let us denote $\Phi_\alpha(\tau)$ the response of the true variance for a f^α frequency noise with a level equal to one (see table I). The measured mean frequency $\langle f_k \rangle$ is a random variable, centered around the nominal frequency f_0 with a standard deviation :

$$\sigma_{f0}(\tau) = \sqrt{h_\alpha \Phi_\alpha(\tau)} \langle f_k \rangle$$

Thus, for an oscillator affected by different types of noises, the knowledge of the different noise levels allows us to estimate the uncertainty of the determination of the nominal frequency :

$$f_0 = \langle f_k \rangle \pm \langle f_k \rangle \sqrt{\sum_{\alpha=-2}^{+2} h_\alpha \Phi_\alpha(\tau)}$$

(see $\Phi_\alpha(\tau)$ **in table I).**

4. CONCLUSION : PHYSICAL MEANING OF THE LOW CUT-OFF FREQUENCY

In the case of a low frequency noise, the results above show that the standard deviation of the mean frequency depends explicitly on the low cut-off frequency.

What is the low cut-off frequency of an oscillator : the age of the oscillator, its life time, the age of the Universe, ... ?

Let us consider a f^{-2} frequency noise. It's derivative, the aging z(t), is then a white noise. The instantaneous normalized frequency deviation y(t) is then :

$$y(t) = \int_{t_0}^{t} z(\theta)\, d\theta$$

where t_0 is the instant of setting up the oscillator. If y(t) is defined as the integration of z(t), no low cut-off frequency is needed to insure the finite power of the signal.
y(t) is then a centered random variable of which standard deviation will increases with time, i. e. with the age of the oscillator. There is an analogy between the age of the oscillator and the inverse of the cut-off frequency.
In table I, one could replace the low cut-off frequency f_l by the inverse of the age Θ of the oscillator :

$$f_l = \frac{1}{\Theta}$$

What is the real frequency of an oscillator during a short (relatively to the age of the oscillator) sequence : its nominal frequency or the mean frequency during that sequence ?

The answer depends also on the model we chose. Considering the power law model of the power spectral density, the frequency is the nominal frequency of the oscillator, because the use of this model implies to define the parameters of this model (the nominal frequency, the noise levels, etc...) in the frequency domain : these quantities are then time independent. Thus the uncertainty we calculated above means : since its setting up, the frequency of the oscillator was included in this uncertainty domain.

However, it seems logical to consider the frequency of an oscillator only during the short sequence : we define the frequency as a quantity which depends on time.
In this assumption, the frequency of the oscillator during a sequence of duration τ is the measured mean frequency and its standard deviation is $\sigma_{\varepsilon k}(\tau)$ (see above).

For a f^{-2} frequency noise as well as a f^{-1} frequency noise, this standard deviation increases with the duration of the sequence :
the longer the duration of the sequence, the larger the uncertainty domain !
This is due to the long term fluctuations.

5. REFERENCES

[1] F. Vernotte, "Stabilité temporelle des oscillateurs : nouvelles variances, leurs propriétés, leurs applications", PhD thesis, Université de Franche-Comté, Besançon, 1991.

[2] G. Zalamansky, "Recherche d'un bruit de fond d'ondes gravitationnelles dans les résidus de chronométrage des pulsars milliseconde", PhD thesis, Université de Paris VII, Paris, 1996.

[3] J. Rutman, "Characterization of phase and frequency instabilities in precision frequency sources : fifteen years of progress", Proceedings of the IEEE, vol. 66, N# 9, September 1978, pp. 1048-1075.

[4] G. Parreins, "Techniques statistiques, moyens rationels de choix et de décision", Dunod technique, Paris, 1974.

On the dissipation effect of the random transitions perturbations

N.Yahyabey (1) and S. Hassani (2)

(1) Université François Rabelais de Tours, Faculté des sciences, Parc de grandmont, 37200 Tours, France.

(2) Centre de developpement des techniques nucléaires, 2 Bld Frantz Fanon, Alger, Algeria.

A proposition that handles and computes the r^{th} order spectrum of non Fourier-transformable functions because of perturbations affecting their periodicities is analysed. By means of a method based on the concept of the overall period, the computed r^{th} order spectrum shows a dependence on the Fourier transform of the distribution function of the perturbations. For absolutely continuous distribution functions with respect to Lebesgue measure, the average amplitude spectrum (first order spectrum) is reduced and the power spectral density (second order spectrum), subject to dissipation, exhibits neither a continuous background nor a continuum spectral component. The reduction factor is linked to the Fourier transform of the probability density function. This important property can be carried out in perturbation characterization process. Also, the random transitions perturbations induce no enhancements of spectral lines widths.

An **experimental** test of this theoretical proposition is presented. It consists of a sinusoidal oscillator coupled to a triggering system operating with two threshold detection levels. The gaussian noise signal present at the control input of the triggering system is used to define randomly the two threshold detection levels for the perturbation of the square wave transitions. Experimental evidence of the reduction effect with no enhancement of existant spectral line widths is shown and the reduction factor depends on the Fourier transform of the probability density function of the random variables measuring the fluctuations at the nominal commutation instants. Also, the measured spectral lines are subject to power dissipation. The obtained experimental results are in agreement with the proposition theoretical predictions.

MULTICHANNEL DISCRIMINATORS FOR HIGH ACCURANCY FREQUENCY AND TIME EVALUATION IN LARGE ENERGY RANGE OF EFFECTIVE SIGNAL

Kirill V. Zaichenko
STATE ACADEMY OF AEROSPACE INSTRUMENTATION, ST. PETERSBURG, RUSSIA

Multichannel evaluation procedures are being inherent result of optimum measures synthesis for arbitrary parameter in conditions of extended by this parameter informative signal. In order to realize optimum procedures in time and frequency areas we have to obtain several readings of informative process in order to achive the highest accurancy in estimation the assigned a priori sector of this informative process signal release. It is known that the optimum procedure demands plausible function maximum calculation by its differentiation. Optimum procedure synthesis indicated that the admitted algorithms can be found in ratio aspect of two polynomials which are being informative process readings weight sums. Number of parameter readings choise has to be concerted with system resolution ability. For example - optimum procedure for three-channel measure is being the ratio of a difference between two lateral readings and a weight sum of all three readings. Denominator of this ratio performs a standardization function ensures so parameter appreciation independence in large scale of signal-noize energy ratio. As the synthesis result were obtained the admitted algorithms of three- and five-channel discriminators for frequency and time measures. In time area evaluation is proceeded for time delay in regard to some supporting value while at frequency area - for average frequency value of informative process spectrum density. The analysis of admitted schemes for time and frequency discriminators based upon the multichannel principle registered their accurancy superiority in comparison with commonly used at present two-channel discriminators. Proposed discriminators advantage is manifested particulary strong in violent variations of signal-noize ratio that is urgent for radionics and radar systems which are characterized by a large energy range of effective signal.

Revision of long-term stability theory of Rb Frequency standard

Agap'ev B.D., Khutorschikov V.I.

St.Petersburg State Technical University

The history of the light shift research was analyzed. It was shown that for last 30 years progress in the long term instability reduction of Rb frequency standarts is small.

This fact may be partially conditioned on the disregard of the random character of the pumping light intensity fluctuations and not account of their correlation properties.

In the present work fluctuation properties of the pumping light are consequentially accounted and the light fluctuations transformation into the frequency fluctuations of the atomic transitions are studied.

It is shown that conditions of reduction of the light shifts due to the light radiation fluctuations qualitativel differ from the basic recomendations. Besides usual recomendations may result to the increase of the long term instability of atomic standart frequency.

Development of a Digital Cesium Clock for Space at the U.S. Naval Research Laboratory

J. White*, W. M. Golding**, E. Powers*, J. Brad*, F. Danzy*, A. Frank**, R. Beard*,

*U.S. Naval Research Laboratory, Washington, DC
**SFA, Largo, MD

1. ABSTRACT

The Naval Research Laboratory has been involved in the development of space-qualified atomic clocks since the early 1970's, providing support to the Global Positioning System (GPS). The goal of the current effort is to bring some aspects of the digital cesium technology that are readily available in commercial atomic clocks (Refs. 1,2) to space. Digital cesium clock technology is particularly valuable in space applications because it can provide enhanced manufacturability, improved operability, and decreased environmental sensitivity. The progress and future direction of this effort are described.

Keywords: digital atomic clocks, GPS cesium clocks

2. BACKGROUND

GPS began flying atomic clocks in space in 1974 with the Naval Research Laboratory's Navigation Technology Satellite I (Ref. 3), the first launch for the Tri-service GPS program. Today, GPS has the operational 24-satellite constellation in orbit and each satellite carries 4 atomic clocks. The GPS experience has provided considerable insight into the most important features required of space clocks. While clock stability is always an essential performance factor in a GPS clock, reliability and diagnostic capability are very important to the operation of the overall GPS system.

For example, some of the GPS Block II cesium clocks have shown excellent frequency stability while others have exhibited performance problems (Ref. 4). The biggest problem that the GPS operators face with these clocks is determining whether on-orbit anomalies are an indication of failure, degradation, or impending failure. Few of these units have experienced catastrophic failure, but rather performance anomalies which could affect system performance or the ability to maintain it. In orbit, the clock's RF signal output cannot be seen directly. The GPS signal is based on the clock but is affected by synthesizers, transmitters, signal propagation and receivers. The telemetry monitors provided to the GPS control segment are essentially those available on the front panel of a 1960's vintage commercial clock. Those monitors can provide insight into the clock's health but are limited in scope and availability. While a laboratory user can combine continuous monitor data with high-resolution phase data and has the opportunity to open the covers to make internal diagnostic measurements, the space operator has very limited data. Even during the satellite vehicle integration and testing phases, the manufacturer must work with limited data to verify that a clock is healthy and should be launched. Temperature sensitivity is another major variable in the Block II cesiums. The clocks were built to a specification of 1×10^{-13} per degree C. With a typical 2-degree peak-to-peak swing in baseplate temperature for a GPS clock, the temperature effect is a significant contributor to the one-day stability.

The GPS program office has also had concerns about the long-term viability of clock suppliers, since the cesium suppliers have been small businesses. In order to minimize the dependence on any particular manufacturer, it is desirable to create an open clock architecture that, if necessary, could be supported with critical clock components from more than one vendor.

The digital cesium clocks now being manufactured for ground or commercial use solve many of these problems. They have demonstrated improved stability and a decreased need for operator interaction. A benefit of the digital servo is to provide a more comprehensive set of monitors that allows the user to have a better understanding of the unit's state-of-health. Digital techniques can be produced more consistently since the reliance on critical and difficult to adjust analog circuitry is much reduced. Due to their advanced digital servo scheme and a re-designed beam tube, the Hewlett Packard 5071A has demonstrated nearly zero coefficients for temperature and magnetic field effects (Refs. 5,6). Because of the capabilities of digital servos, digital clocks also have the potential to provide improved reliability.

3. DESIGN GOALS

Based on the GPS Block II operational experience, the NRL goals for space-based digital cesium clocks are:

3.1 Frequency Stability

The new clocks must meet or exceed the level of frequency stability attained by the present generation of GPS cesium clocks. These clocks have met or exceeded their design requirements for stability, and the GPS system is operating within its specification using them (Ref. 4). The inherent accuracy, drift free behavior and long term stability obtainable with beam tube based cesium clocks are important to the ability of GPS to maintain system time, navigation performance and transfer accurate time. These qualities must be maintained in the next generation of space cesium clocks.

3.2 Manufacturability

The new generation of clocks must be easier to build. Analog style clocks have been difficult and time consuming to manufacture. The difficulty of the setup procedure requires a high level of training and produces a wide level of variation in performance from unit to unit. Precise measurements are required, and each clock must be observed and tested for long periods to verify correct operation. Because of the low-level signals involved, great care must be taken to prevent grounding problems and leakage of critical signals between input and output signal paths. The analog modulation used in the servos requires exceptional symmetry to eliminate frequency offsets and the associated degradation in stability. Another problem area has been the generation of a clean microwave spectrum. Traditional techniques for producing the interrogation signal also produce large amplitude sidebands within 20 MHz of the carrier. Strong sidebands and other spurious signals in the interrogation signal cause the

apparent atomic frequency to be shifted through the AC Zeeman effect, producing frequency offsets that are dependent upon the stability of this spectrum. Many electronic effects will cause the sideband levels and the level of the interrogation signal to vary. This can only lead to a degradation of stability and a increased temperature coefficient. The digital servo approach can reduce or eliminate many manufacturing problems. Since the analog FM or PM modulation method is no longer in use, the need to align the modulation and detection circuitry is gone. The discrete frequency hopping approach, coupled with a Dielectric Resonator Oscillator (DRO) or Surface Acoustic Wave (SAW) oscillator in the Radio Frequency (RF) chain, greatly reduces the problems associated with spurious signals being applied to the beam tube and lessens the problems associated with the design and construction of the RF multiplier chain.

3.3 Environmental performance

Both temperature and magnetic fields change in the on-orbit environment, and there is benefit to reducing the level of these effects. The goals are to reduce the temperature coefficient to less than $5x10^{-14}$ degree C and the coefficient of external magnetic fields to less than $1x10^{-13}$/Gauss. Due to higher than expected temperature coefficients, a base plate thermal controller (+/- 0.1 C) has been added in the GPS Block IIR satellite design to improve the temperature control of both the rubidium and cesium clocks. This improved temperature control was needed to allow the satellite to meet specifications but at the cost of tens of watts of precious satellite power. The key to meeting the temperature coefficient requirement is to reduce the magnitude of residual frequency offsets and the factors that cause changes in them. These factors are obviously beam tube design dependent, and one challenge of the current effort is to reduce the variations from one manufacturer to another.

An important factor is the sensitivity to interrogation power levels. Digital servos can stabilize the interrogation power level by maximizing the beam current versus power response of the tube. The magnetic field effects can be reduced by use of the digital servo to measure and stabilize the Zeeman splitting by interrogating the field dependent transitions and adjusting the C-field. The servo loop gain can be held constant over the life of the tube by adjusting the electron multiplier gain, preamplifier gain or digital filter coefficients. This helps reduce offsets and offset variations related to the servo system and increases the effective useful life of a clock relative to the earlier analog designs.

3.4 Testability

Testing the clock during vehicle integration is time consuming and difficult. This new clock will provide a clearer picture of the clock's true state of health than previous designs. It will include more than the traditional oven and signal level monitors. Under computer control, the beam tube state may be measured internally and output to the user. This can be thought of as an automatically updated manufacturer's specification sheet for a particular tube. The flexibility of the digital architecture allows many tube parameters to be measured. These measurements can be made either by running an extensive measurement and characterization routine or by defining frames within the controllers time sequence for specific measurements to be performed.

For example, the measured values of peak, valley and background beam currents as well as signal-to-noise estimates obtained from fluctuations in the sampled data can be used to track the figure of merit and predicted stability. During the test sequence, estimates of the velocity distribution can be made allowing accurate calculation and tracking of velocity-dependent frequency shifts such as second-order Doppler. Using the field dependent transitions, the homogeneity of the C-field can be measured and tracked. Variations in the amplitude of the cross-field transitions can also be tracked. Schemes to measure and monitor most aspects of beam tube performance can be developed using the flexibility of the digital servo system. The ability to remotely monitor these internal diagnostics can help determine whether a clock is a reliable flight candidate at any time prior to launch. After launch, these diagnostic routines and remotely available data can help the operators make an informed decision about clock performance and potential longevity. Graceful transitions to backup clocks can be made as necessary with minimum disruption to satellite operation and performance. These monitors or data may be available in the normal data stream and will not require opening the clock or removing it from the space vehicle.

3.5 On-orbit monitoring

In general, the same remote diagnostic capabilities that facilitate testing will greatly improve on-orbit state of health monitoring, operational readiness and lifetime estimates. This capability is more critical on-orbit since the ability to directly look at the clock's output signal is no longer available. In particular, the ability to see the critical beam tube performance factors can help the ground operators maintain the system performance with greater reliability and confidence.

4. NRL EXPERIMENTAL WORK

These goals are now being experimentally investigated in the laboratory. A digital space cesium clock testbed, shown in Figure 1, has been built using a combination of a prototype synthesizer, computer controlled test equipment and a variety of beam tubes from older space clocks and commercial units.

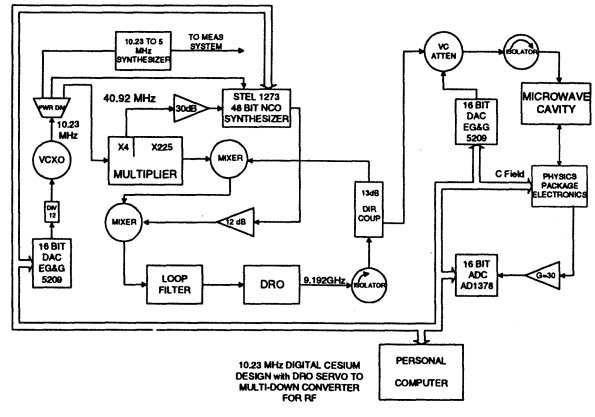


Figure 1. A block diagram of the basic NRL testbed system showing the system architecture. Tubes from various manufacturers can be easily substituted for one another.

The use of commercial laboratory instruments allows the focus of the efforts to be on realizing the project goals and determining hardware requirements while deferring the building of the actual space hardware until later in the program. Areas of particular interest have been the selection and optimization of the beam tube and servo algorithms, the elimination of frequency offsets, the use of beam tubes from multiple sources, and the determination of how to best use the diagnostic capabilities of this system in a space clock.

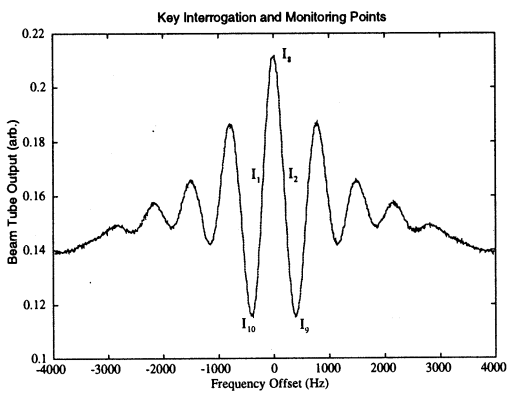


Figure 2. Various tones within the central Ramsey fringes that are used in the system. Other tones outside the central fringes are also used.

The typical tone modulation frequencies are shown in Figure 2 superimposed on a beam tube Ramsey response curve. This selection allows location of the center of the primary response, measurement of effects of external microwave fields, microwave power level optimization, and measurement of the beam tube figure of merit. The sequence and duration of the tones over a measurement cycle are described in Table 1. The symbols in Table 1 are: I_1, the first beam current measurement etc., W the Ramsey linewidth, F_Z the Zeeman splitting of the transition, P_n the interrogating power level, and Δ the offset used for the power leveling servo. The sequence shown describes a basic servo scheme and tones are easily added to allow various measurements.

Table 1. Simplified Interrogation Tone Sequence Describing the Basic Servo and Monitor Points

Sequence Position	*Tone Frequency Offset*	*Tone Power Level*	*Tone Duration (msec)*
I_1	-W/2	P_n	200
I_2	+W/2	P_n	200
I_3	-(F_Z-W/2)	P_n	200
I_4	+(F_Z+W/2)	P_n	200
I_5	-W/2	P_n+Δ	200
I_6	+W/2	P_n-Δ	200
I_7	-F_Z/2	P_n	100
I_8	0	P_n	100
I_9	+W	P_n	100
I_{10}	-W	P_n	100

Another important area in the research has been the effect of spurious signals in the microwave excitation signal applied to the beam tube. Traditional microwave multipliers have shown strong spurious signals. The effect of these on the frequency of the clock can be very significant. A stable spurious signal will produce an offset in the frequency of the clock whose magnitude is proportional to the power in the spur and inversely proportional to the offset from the cesium transition frequency. Instabilities in the clock's output frequency arise when these spurious signals change amplitude. The most common cause for changes in spurious signal levels is temperature. As line lengths and component values change, the phase along critical microwave paths also change, leading to variations in the spectrum of the interrogation signal. To reduce these effects, the NRL design uses a microwave oscillator in a clean-up loop configuration. The measured spectrum of a DRO slaved to the interrogation signal at 9.192 GHz, Figure 3, shows no spurious signals within 70 dB of the carrier. SAW oscillators can also be used to produce comparable results and may have other system dependent advantages.

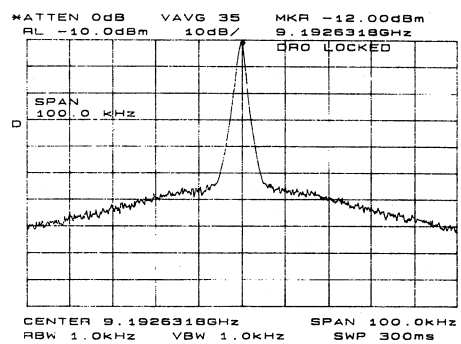


Figure 3. Measured spectrum of the 9192 MHz interrogation signal. Factional frequency offsets due to spurios signals are estimated to be in the low 15's.

Because one of our program goals is to minimize the dependence on beam tubes from a single manufacturer, five beam tube types from four different tube manufacturers were operated in the lab testbed. These tubes showed widely varying sets of basic parameters, Table 2, and ranged widely in age and general health. In each case, the digital servo was able to provide good to excellent performance. and an example of the performance achievable is shown in Figure 4.

The digital servo provides more comprehensive information about the internal workings of the clock than prior analog designs. Information routinely available from the system includes: signal to noise ratio, Zeeman frequency, RF power level, VXCO control voltage, C-field homogeneity, figure of merit, stability estimates, velocity distribution, flop to background ratios and so forth.

Table 2. The stability measurement results using the NRL testbed electronics with beam tubes from various manufacturers

Tube Source	*Line width (Hz)*	*DC Signal to Noise*	*Figure of Merit (Ref. 7)*	*Measured Allan Deviation*
1A	425	580	1.4	$5x10^{-11}/\sqrt{\tau}$
1B	494	1100	2.2	$1x10^{-11}/\sqrt{\tau}$
2	366	7850	21.5	$7x10^{-12}/\sqrt{\tau}$
3	473	1790	3.8	$2x10^{-11}/\sqrt{\tau}$
4	500	100	0.2	$8x10^{-11}/\sqrt{\tau}$

In addition to the work with the NRL testbed, monitor and internal parameter data has been collected from that available in digital commercial clocks. This data is being analyzed to look for significant trends. Taken together, this database is providing a broad picture of the internal workings of digital clocks. The current state of the clock, shown from these monitors, is being used to determine the information that a ground control system would need for successful long-term operation in space.

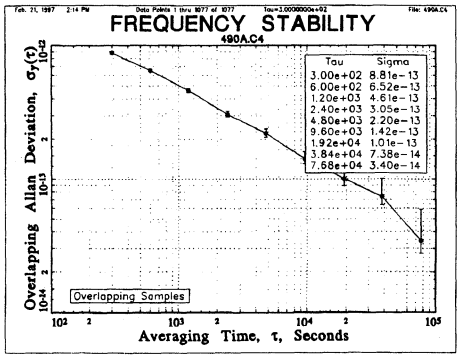


Figure 4. A sample stability measurement using the NRL testbed electronics.

5. FUTURE WORK

Future areas to be investigated are diagnostic collection and interpretation techniques that allow for useful state of health estimation, and flexible techniques for measuring and compensating for environmental effects. The effectiveness of optimal control techniques will be studied. Test regimen will be developed that can be applied during the integration of the clocks into satellite systems to verify the state-of-health of the clocks through sub-system, system and launch readiness testing of the complete satellite. On-orbit operational telemetry and state-of-health monitoring techniques for reliable, long-life operation of clocks for GPS will ultimately result.

6. CONCLUSIONS

The NRL digital space cesium clock effort has shown the feasibility of a space-qualified cesium clock employing the digital design concepts of today's best commercial cesium clocks. In addition, we have shown that such a clock can be built using beam tubes from more than one vendor, leading to a more open architecture. The capacity of the digital servo to give a much more complete picture of the state of the clock to the operator will provide space clocks that perform better, operate more simply and last longer.

7. REFERENCES

1) Karlquist, R., "A New RF Architecture for Cesium Frequency Standards," *Proceedings of the 1992 IEEE Frequency Control Symposium.*

2) Cutler, L.S. and Giffard, R.P., "Architecture and Algorithms for New Cesium Beam Frequency standard Electronics," *Proceedings of the 1992 IEEE Frequency Control Symposium.*

3) Bartholomew, C.A., "Navigation Technology Satellite I," *Proceedings of the Sixth Annual Precise Time and Time Interval (PTTI) Planning Meeting*, December 1974.

4) McCaskill T.B., Reid W.G. and Buisson J.A., "On-orbit Frequency Stability Analysis of GPS Block I and II Navstar Clocks," *23rd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, 1991.

5) Kusters, J.A. and Johnson J.L., "A New Cesium Beam Frequency Standard Performance Data," *Proceedings of the 1992 IEEE Frequency Control Symposium.*

6) Cutler L.S. et al., "Frequency Pulling in Cesium Beam Frequency Standards Due to Delta M=±1 (Sigma) Transitions," *Proceedings of the 45th Annual Symposium on Frequency Control,* pp. 544-553, 1991.

7) Vanier, J. and Audoin, C. *The Quantum Physics of Atomic Frequency Standards*. Bristol and Philadelphia: Adam Hilger, 1989.

IMPROVEMENT OF THE SHORT TERM STABILITY OF THE LPTF CESIUM BEAM FREQUENCY STANDARD

A. Makdissi, J. P. Berthet* and E. de Clercq

BNM-Laboratoire Primaire du Temps et des Fréquences
Observatoire de Paris
61 Avenue de l' Observatoire-75014 Paris-France
*present address : Virgo, Laser Optics, LAL, Université Paris-Sud, 91 405 Orsay, France

Abstract. The short term stability of our optically pumped Cs beam frequency standard has been greatly improved by using two atomic transitions for optical pumping. The present stability is $\sigma_y(\tau) = 3.8\ 10^{-13}\ \tau^{-1/2}$.

Keywords : frequency stability, Cs frequency standard, optical pumping.

1. INTRODUCTION

In the present optically pumped cesium beam frequency standards a single atomic transition is used for pumping the atoms from one hyperfine level to the other one in the ground state. The atomic population is then distributed among all the Zeeman sub-levels of this hyperfine level and fifteen percent of the atoms at best are used in the clock transition. It is known (Ref. 1) that the clock signal can be increased by using simultaneously a second transition depleting all the Zeeman sub-levels except the m = 0 sub-level of the final state. This has already been experimentally observed (Refs. 2,3), but up to now the signal increase has not been followed by the same improvement on the Signal-to-Noise ratio. Indeed, the Signal-to-Noise ratio, or the short term stability, can also be improved if narrow linewidth lasers are used.
We compare here the results obtained with narrowed lasers in the two cases : pumping with one transition and pumping with two transitions.

2. TECHNICAL FEATURES

The optically pumped Cs beam primary standard JPO has already been described elsewhere (Ref.4). The main features are a Ramsey cavity of one meter length and a width of the central resonance equal to 100 Hz.

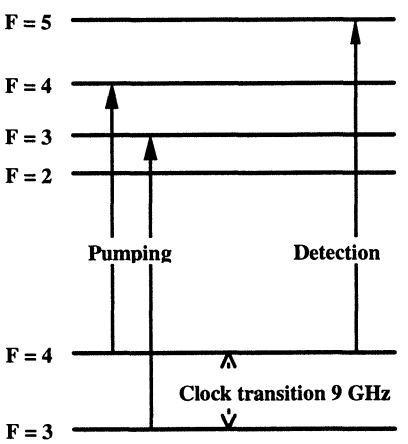


Fig. 1 *Cs energy levels*

In our new experimental set-up the optical pumping is performed by two extended-cavity diode lasers (their linewidth is less than 100 kHz). They are frequency locked to the D2 line by means of the saturated absorption spectrum in a Cs cell. Fig. 1 shows the involved energy levels. The first laser is frequency locked to the 4-4 σ polarised transition and the second laser to the 3-3 π polarised transition. The powers of the 4-4 and 3-3 laser beams in the pumping region are respectively 1mW and 0.5mW. The detection is performed with the 4-5 transition, a part of the 4-4 laser beam is frequency shifted by an acousto-optic modulator for this purpose.

3. SIGNAL

We compare two pumping configurations : one transition (4-4 σ polarized) and two transitions (4-4 σ and 3-3 π).

3.1 Pumping with one transition

In the first configuration all the Zeeman sub-levels of the (F=4) state are depleted and the final atomic population is distributed over the (F=3, m) Zeeman sublevels. The corresponding MicroWave spectrum is shown on Fig.2. The seven MW σ transitions ($\Delta F=1$, $\Delta m=0$) in the ground state are clearly visible. The dotted line shows the signal when the MW power is off. The pumping is nearly complete. The ratio of unpumped atoms in the (F=4) level is less than 1.10^{-4} of the initial population of this level. The Signal-to-Noise ratio measured in 1 Hz bandwidth at the top of the central Ramsey fringe is $1.6\ 10^4$.

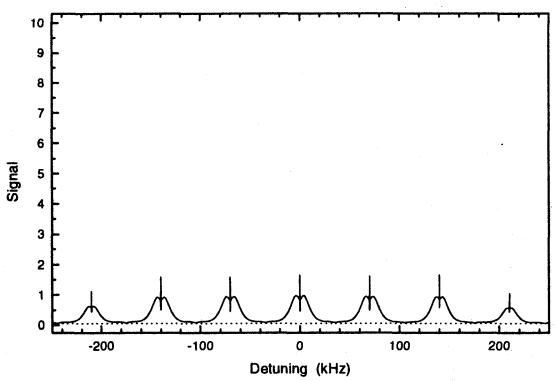


Fig. 2 *One transition pumping, microwave spectrum*

Fig. 3 shows the population in percent of each (F=3, m) Zeeman sublevel. The population is calculated from the signal at resonance at the optimum MW power for each one.

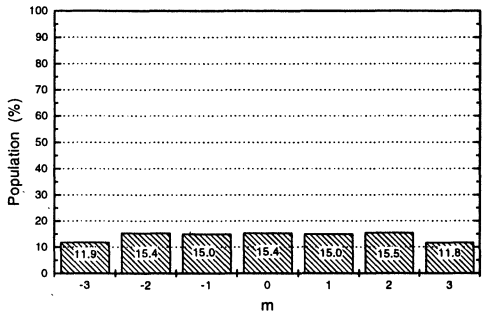


Fig. 3 *One transition pumping, distribution of the population over the levels (F=3, m)*

3.2 Pumping with two transitions

As the (F=3, m=0)-(F'=3, m=0) transition is forbidden by selection rules, when we add this second laser beam for pumping the (F=3, $m \neq 0$) Zeeman sublevels will be depleted, but not the (F=3, m=0) sublevel. Then, after a few cycles absorption-spontaneous emission the atoms are accumulated in the (F=3, m=0) sublevel.
The MW spectrum recorded with the same gain and offset as in Fig. 2 is shown on Fig. 4. The amplitude of the 0-0 clock transition is multiplied by 6. The ratio of unpumped atoms in the (F=4) level is 1.5 10^{-4} . The Signal-to-Noise measured at the top of the resonance is 3 10^4.

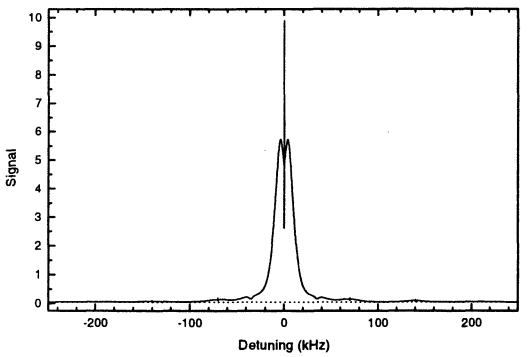


Fig. 4 *Two transitions pumping, microwave spectrum*

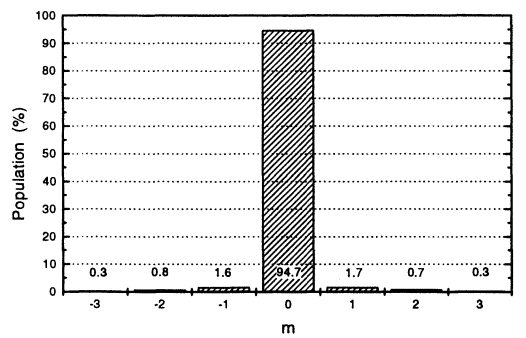


Fig. 5 *Two transitions pumping, distribution of the population*

The pumping is almost total. Fig. 5 shows the relative populations of (F=3) Zeeman sublevels. Almost 95% of the initial population is gathered in the (F=3, m=0) sublevel

3.3 Comparison of velocity distributions

We have verified that the atomic velocity distribution is not modified by a two transitions pumping. It can be simply checked by comparison of the Ramsey patterns recorded at the same high MW power. Fig. 6 shows the Ramsey patterns recorded at 25 times the optimum power for the two pumping configurations. The record of the one transition configuration is fitted on the other-one by adjusting the gain and the offset.

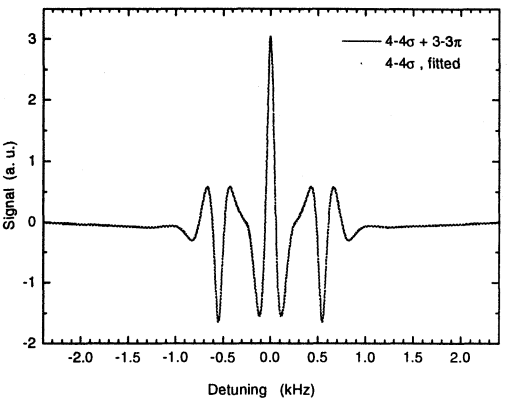


Fig. 6 *Comparison of Ramsey fringes at high MW power*

The two curves fit very well. Then the velocity distribution is the same in the two cases.

4. STABILITY

The relative frequency stability of our standard is measured with respect to an H maser. Fig. 7 shows the Allan standard deviation measured with the two configurations.

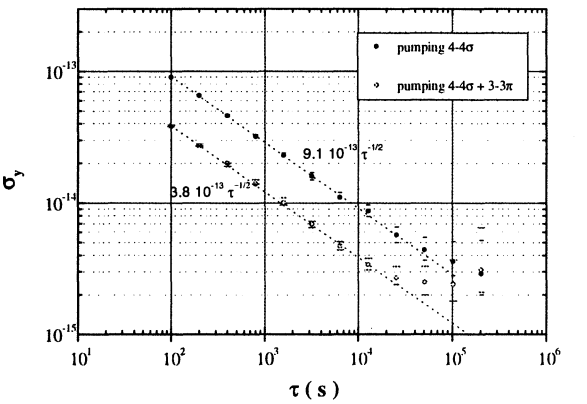


Fig. 7 *Allan standard deviation Cs beam - H maser*

With one transition pumping the short term stability is $\sigma_y(\tau) = 9.1\ 10^{-13}\ \tau^{-1/2}$. With the two transitions pumping configuration the measured stability is $\sigma_y(\tau) = 3.8\ 10^{-13}\ \tau^{-1/2}$. They are limited for long integration times by the stability of the H maser.

5. CONCLUSION

We have shown that using two transitions for optical pumping can significantly improve the short term stability of optically pumped cesium beam frequency standards. With this pumping configuration the present stability of our standard is $\sigma_y(\tau) = 3.8\ 10^{-13}\ \tau^{-1/2}$. To our knowledge this result is the best obtained, until now, on a Cs beam frequency standard.

REFERENCES

[1] H. J. Gerritsen and G. Nieuhuis, *Appl. Phys. Lett.*, **26**, pp. 347-349, 1975.

[2] G. Avila, E. de Clercq, M. de Labachellerie and P. Cérez, *IEE Trans. Instrum. Meas.*, **IM-34**, pp. 139-143, 1985.

[3] N. Dimarcq, V. Giordano, C. Jacques, G. Théobald and P. Cérez, Proc. *5th European Frequency and Time Forum*, Besançon, pp. 187-189, 1991.

[4] G. D. Rovera, E. de Clercq, and A. Clairon, *IEE Trans. Ultrason. Ferroelec. Freq. Contr.*, vol. 41, pp. 245-249, 1994.

IMPROVED 3-D MOLASSES FOR TALL Cs FOUNTAINS

J.-S. Boulanger, M.-C. Gagné and R.J. Douglas

INMS, National Research Council of Canada
Ottawa, Canada K1A 0R6
(613) 993-5698, Fax (613) 952-1394
Email: jean-simon.boulanger@nrc.ca

Optical molasses was studied in 3-dimensional static and 2-dimensional moving configurations for tall cesium fountains at NRC[1][2]. The friction coefficient of the molasses, which limits the acceleration of atoms for a cesium fountain, was measured by the temperature damping technique[3] changing suddenly the laser intensity. We see two time constants, one very short ($\sim 2-4\mu s$), and one longer, varying from 20 to 200μs. We associate the short time constant with atoms which are tightly linked to the molasses and have a large friction coefficient. The long time constant, associated with atoms weakly linked to the molasses, accounts for the observed spreading of accelerated atoms. Laser modulation was found to affect this composite friction coefficient.

Laser cooled, optical molasses, cesium, cesium fountains, friction, acceleration, laser modulation.

1 INTRODUCTION

In optical molasses a friction coefficient determines the maximum acceleration that can be given to the atoms during the launching phase of an atomic fountain. The friction coefficient of 3-D molasses can be deduced by measuring the damping time constant τ_α of the temperature after a sudden change of the laser intensity [3].

Our first measurements of the friction coefficient [4] led us to conclude that we had stronger values ($\alpha \sim 2 \times 10^{-20}$kg/s) than previously reported[3] in the σ^+/σ^- configuration over the detuning frequency range(10 to 70 MHz). It is of the same order as the α coefficient measured in the lin $\perp$ lin configuration [3], which we have not yet studied on our system.

Looking for an explanation, it appeared that our use of an offset-locked master laser, and optical molasses laser beams free of unwanted modulation could have been the factor which lead to our improved results. However, there had also been other improvements, for example in the cat's-eye acousto-optic modulator combination - to allow for a large frequency change of the laser beams without any effect on the beam direction. A new series of measurements were made to investigate the effect of adding different modulation to the master laser, of types that would commonly be used for locking the master laser. These measurements showed that we had two different friction coefficients. The measurements were repeated over a wide range of power (from $\Omega^2/\Gamma^2 = 3$ to .3, where Ω is the Rabi frequency and Γ is the linewidth of 5 MHz for Cs D_2) and frequency detuning (from-2Γ to -14Γ).

These very good coefficients for our optical molasses permit acceleration of atoms to launch velocities of 7 m/s and more, even when limited to a power of 6 mW per beam. We first present the details of our molasses that allow strong friction coefficients and high launching velocities. The next section will discuss the results obtained for the friction coefficient, the effect of laser beam modulation and the launching of atoms.

2 EXPERIMENTAL SETUP

Our magneto-optic trap is an orthogonal system with three pairs of counterpropagating beams, two of which are at $\pm 45°$ from the vertical axis. The beams intersect in a cesium vapour cell[5] The laser beams are provided by diode lasers.

2.1 Unmodulated Lasers

An offset locking technique provides a stable unmodulated master laser beam for our experiments. A master laser (extended-cavity diode laser) is tuned to the F=4→F'=5 cycling transition of the cesium D_2 line at 852 nm. It is offset-locked to the saturated crossover absorption line F=4→F'=5/4 in a cesium cell with the proper offset (125.7 MHz) provided by double passage through an acousto-optic modulator (AOM) and cat's-eye combination, but leaving the main master laser beam unmodulated.

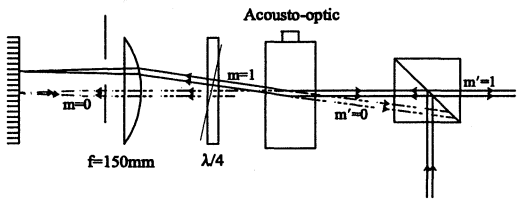


Figure 1: Schematic of the optics for the cat's-eye and acousto-optic modulator.

The 125.7 MHz frequency offset is generated using a direct digital synthesizer modulated at 33 kHz with a modulation depth of 5 MHz.

The master laser beam is split in two. One half is used for the detection and the other half is frequency shifted and amplified by injection into a more powerful (150 mW) diode laser, called the slave laser.

The injection is done through another cat's-eye/acousto-optic modulator mounted in such a way that the full range of the AOM (80 MHz ±30 MHz) can be used without any deviation of the offset beam and with no return of the zero order beam.

2.2 Improved Cat's-eye

In a cesium fountain design, a cat's-eye is commonly used with an acousto-optic modulator to offset the frequency of a laser without changing the beam's direction. The use of a traditional cat's-eye does help but a simple modification of the ensemble cat's-eye/AOM makes it very stable over the whole frequency response of the AOM.

A long focal length plano-concave lens (150mm) is used, with the focus at the center of the acousto-optic modulator (see Figure 1). Alignment of the mirror, which is ideally at one focal length behind the lens but can be closer, is done first, using the zero order beam. The lens is then inserted at a position such that the outcoming beam keeps its alignment. A variable aperture is then set between the lens and the mirror to select the order of the deflected beam and stop the zero order beam. This setup is also a very good switch: turning off the AOM drive eliminates the returned laser beam. In this way the AOM's eliminated the need for mechanical shutters.

2.3 Laser Beams in the Trap

The slave laser beam is split in three and each beam is independently controlled in frequency (and amplitude) by another AOM/cat's-eye combination. The frequency of each laser beam can be controlled with a resolution of better than one Hertz, not accounting for the 10 kHz linewidth of the master laser. Each beam is spatially filtered by a 50 μm pinhole and expanded to a 7.4 mm ($1/e^2$) beam. The intensity of each beam can be adjusted by controlling the amplitude of the RF frequency feeding the AOM's. The proper polarisation (σ^+/σ^-) is adjusted by $\lambda/4$ waveplates.

Another diode laser, with a 20 MHz linewidth, is locked to the F=3→F'=4 transition to 'repump' the atoms which are escaping the F=4→F'=5 cycling and trapping transition. A simpler locking technique is used, since this transition is not as critical as the cycling transition. The repumping beam is combined with the horizontal cycling beam through a polarizing cube.

2.4 Cesium Vapor Cell

The cesium is in a vacuum system made of stainless steel. The optical ports are anti-reflection coated with less than 0.5% loss in transmission. The vacuum, excluding cesium, is better than 1.3×10^{-7} Pa. Hydrogen accounts for most of the background pressure. Other gases are at levels 20 (or more) times lower than hydrogen.

Activated graphite tubes are used to isolate the trap chamber from the rest of the vacuum system, and act as a very efficient cesium getters - allowing us to control the cesium partial pressure in the trap by adjusting the temperature of the cesium source. Typically we operate at 10°C, equivalent to an equilibrium with a source at -15°C (2×10^{-6} Pa).

Two anti-Helmholtz coils of 95 turns of wire generate the magnetic field gradient needed for MOT operation. Each coil has an average radius of 4.7 cm and they are 8.5 cm apart. Under computer control, the current can be switched between any of two preset values (adjusted manually) and zero. It takes less than two milliseconds to stabilize to the new value. The range of preset values is from 0.6 A to 5.3 A. The resulting gradient of magnetic field at the zero field point ranges from 2 to 17.7 G/cm.

2.5 Control and Measurements

Direct digital synthesizers feed the AOM's all the RF signals needed, with phase continuous frequency changes. The frequencies can be ramped under computer control. The amplitude of the radio frequency signal can be adjusted manually to two different levels which can be switched under computer control. Shifting the frequency of the slave laser one way and the output of the correcting AOM to compensate can put to use the limited bandwidth of the AOM to fur-

ther play with the amplitude without changing the frequency of the laser beams.

A calibrated CCD camera is used to image the trap and provide an estimation of the number of atoms seen in the trap. The CCD camera was calibrated against a calibrated photodiode (at 852nm), using the same optics and the same target.

Under computer control, a homemade timing sequencer generates all the signals needed by the synthesizers, the CCD camera and the data acquisition system, with a timing precision better than 100 ns.

2.5.1 Detection

Calibrated photodiodes are used to monitor the trap and to detect the fluorescence of the atoms falling in the detection region, which is 20 cm below the trap and isolated by a graphite tube to absorb stray cesium atoms from the background pressure in the trap region.

Two horizontal counterpropagating laser beams fed directly from the master laser, tuned on the F=4→F'=5 cycling transition, are crossing the falling atoms. The fluorescence is detected by a calibrated photodiode.

2.5.2 Time of Flight

The time of arrival of atoms in the detection region depends on the atomic cloud velocity distribution and acceleration when the MOT or molasses are turned off. Recording this signal gives primary information on the temperature of the trap or molasses (see Figure 4, as an exemple). This method of measurement is called the time of flight (TOF) technique [6]. When no acceleration is given to the atoms, it also gives the relative position of the trap and the detection region very accurately (better than 0.1 mm).

3 Measurements

3.1 Friction Coefficient

The friction coefficient has been measured by the temperature damping technique [3], reducing suddenly the laser intensity by one half, or more. When we first reported our trap parameters [4], we still had to explain why the values of our friction coefficient were so high compared to values measured elsewhere [3].

The results obtained on our first attempts [4] had shown some curious behaviour which had been attributed to the inaccuracy of the temperature measurement, using the time of flight technique. It seemed that a single exponential fitting to the curve was not always the best model to use. Nevertheless, single exponential fitting gave good results.

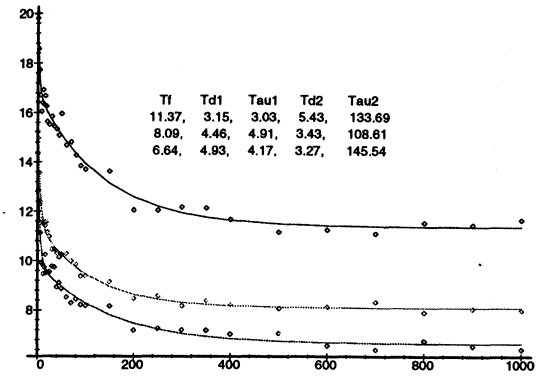


Figure 2: Temperature of molasses (μK)measured via the time of flight technique versus the supercooling time (μs). The parameters displayed are in the same order than the curves generating them. From top to bottom the modulation depth is 6.2 MHz, 3.2 MHz and 0 MHz. The modulation frequency is always 33 kHz.

More careful measurements, going up to 1 msec after the power switching, have shown that the fitting should really be done with two exponential curves, using two different time constants,τ_1, very short ($\sim 4\mu s$), and another one, τ_2, very long (up to $\sim 200\mu s$). Referring to the parameters in Figures 2 and 3, the equation of temperature versus cooling time after the power change is

$$T(t) = T_f + T_{d1} * exp^{-t/\tau_1} + T_{d2} * exp^{-t/\tau_2} \quad (1)$$

where T_f,T_{d1} and T_{d2} are respectively the final temperature and the temperature change associated with each time constant. Then the question arises: which one is the time constant associated with the friction coefficient? If we use the short time constant, the friction coefficient, α, is then of the order of $\sim 5 \times 10^{-20}$kg/s.

3.2 Modulation

We thought that the characteristics of our optical system could explain some of the improvements found with our system. We repeated the measurements using a more classic locking technique for the master laser [7], modulating directly the master laser. We used a 33kHz modulation rate, with

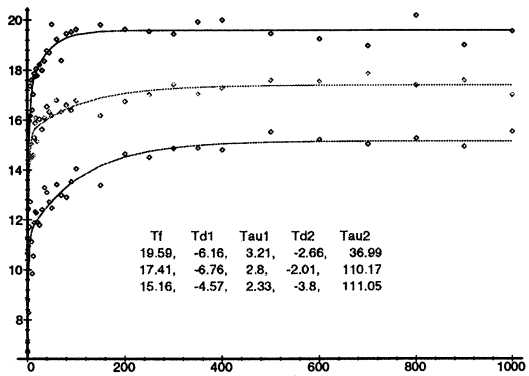


Figure 3: Temperature of molasses (μK)measured via the time of flight technique versus the less-cooling time (μs). The parameters displayed are in the same order as the curves generating them. From top to bottom the modulation depth is 6.2 MHz, 3.2 MHz and 0 MHz. The modulation frequency is 33 kHz.

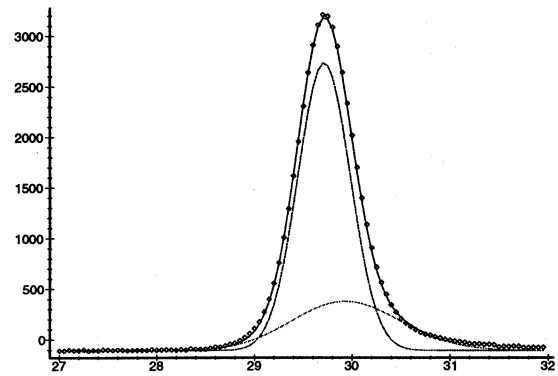


Figure 4: Time of flight measurements of atoms accelerated downwards at 6 m/s. The amplitude is the recorded signal in millivolts, and the time after launching is in milliseconds. It can clearly be seen that the experimental points can be fitted by the sum of two gaussian curves, one of cooled atoms at the nominal velocity, and another one of warmer atoms at a slightly lower velocity.

a peak-to-peak modulation depth of 3.1 MHz or 6.2 MHz.

The short time constant is virtually independent of the frequency offset or the modulation technique. See Figure 2. The long time constant ($\sim 20 \leftrightarrow 200\mu s$) appears to be inversely proportional to the modulation depth.

The number of atoms is also inversely proportional to the modulation depth. Unfortunately, we cannot use directly the measurement made in the detection area as a relative measure of the number of atoms from one modulation depth to the other since the detection beam is also modulated, reducing the detection efficiency. But the value of the fluorescence in the trap is also affected by that modulation and the ratio of the fluorescence of the two regions is a better estimate of the change in the real number of atoms. There is a real drop (as large as 40%) in the number of atoms at 6.2 MHz. The time constant is of the order of $20\mu s$.

The short time constant, which is in the range of $\sim 2 - 4\mu s$, appears to be due to atoms strongly linked with the molasses, while the long time constant would be atoms near the top of the quantum well created by the molasses. When the power is suddenly dropped (in less than one μs, the atoms already in the new well are kept there and further cooled by the Sisyphus process while the atoms near the top of the quantum well have to be 'recaptured' by the Sisyphus process. Thus the modulation of the laser would have an effect on the atoms at the top of the quantum wells. Because of the 20% or so shaking due to the modulation of the laser, the well would appear shallower than it is actually without modulation. This could explain the difference in temperature for the three molasses shown in the Figures 3 and 4.

In order to sort out the various hypotheses we repeated the measurements by increasing the laser power instead of reducing it. The puzzle is still there, altough the hypothesis of atoms in a shaking well explains the fact that the number of atoms do not increase with a power increase, while it is reduced with a reduction of power.

3.3 Launching

We are presently limited by the vertical extent of our experimental apparatus in the long-drift time analysis obtainable from upwards launching. The maximum upwards speed detectable is about 2.5m/s due to the low ceiling of our present vacuum system.

No such limit applies to downwards accelerations, towards the detection region. Atoms which slip out of the main acceleration group are also most easily detected with downwards acceleration. Velocities up to 10m/s were recorded, although at such speed we were losing some atoms as can be seen by distorted TOF curves. For acceleration of less than 2msec, speeds of 6m/s without much distorsion were obtained (see Figure 4). The distortion could be due to the occasionally distorted shape of the atomic cloud in the trap. Upwards accelerations with a taller vacuum chamber will resolve this matter.

The temperature did not seem to change over the

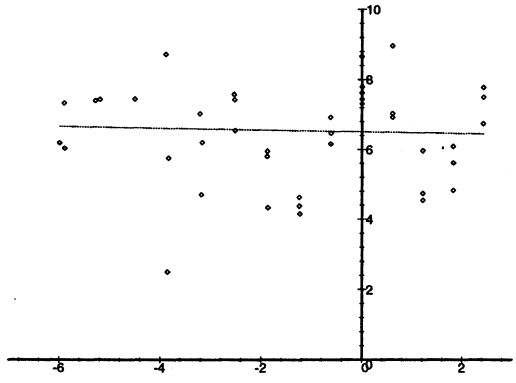


Figure 5: Temperature (μK)of the atoms versus the launching velocity (m/s). Because of our actual configuration, the maximum upwards velocity is limited to 2.5 m/s.

range of velocities investigated as far as we could measure it. Since the measurement of TOF is much less accurate for short TOF than long ones, the accuracy of the recorded temperature is not very good at high velocities (see Figure 5). There seem to be no increase for the upwards launches, up to 2.4 m/s, hence we can assume that the parameters used are also good for the downwards launches.

Some additional insight has been gleaned from the gradual failure of a synthesizer. The faulty synthesizer started to behave erratically only at the frequency used for the molasses, skipping some cycles intermittently, thus losing phase coherence with the other beam. The TOF obtained under accelerating conditions were distorted in funny ways, showing a small cooled peak of atoms being normally accelerated, superposed to a more or less broad pedestal showing a loose link to the molasses for most of the atoms. Even the simple test of measuring the temperature by just dropping the atomic cloud gave rise to strange results and higher temperatures of the molasses than usual. This incident could point out that the phase coherence of the molasses is very important for the temperature and launching of atoms.

4 CONCLUSION

Using frequency synthesizers with phase coherence also promotes high accelerations without significant heating. The use of offset locking technique for the master laser and the use of improved directivity in the AOM/cat's-eye configuration allows for simultaneous acceleration and cooling which will permit cesium fountains with Ramsey interaction times in the range between one and two seconds.

5 ACKNOWLEDGEMENTS

We wish to thank Patricia Mitcheler, Chantal Marois and Joanna Karczmarek for their summers' contributions. Many thanks to Louis Marmet for his help in processing some of the data. The technical and engeneering assistance of Bill Hoger, Bill Cazemier, Ray Pelletier and Stan Cundy are gratefully acknowledged.

References

[1] J.-S. Boulanger, D. Morris, R. J. Douglas. and M.-C. Gagné, **Hydrogen Masers and Cesium Fountains at NRC**, in *Proc. 25th Annual PTTI Applications and Planning Meeting*, pp. 345-356 (1993).

[2] R.J. Douglas and J.-S. Boulanger, **Local Oscillator Requirements for Timekeeping in the $10^{-14}\tau^{-1/2}$ era**, *Proceedings of the 1992 IEEE Frequency Control Symposium*, pp. 6-26 (1992).

[3] M. Drewsen, Ph. Laurent, A. Nadir, G. Santarelli, A. Clairon, Y. Castin, D.Grison, C. Salomon, **Investigation of sud-Doppler cooling effects in a cesium magneto-optical trap**, *Appl. Phys.* B 59, 283-298 (1994).

[4] J.-S. Boulanger, M.-C. Gagné and R. J. Douglas, **Cold Atoms and Cesium Fountains at NRC**, *Proceedings of the 1996 IEEE International Frequency Control Symposium, Hawaii*, pp. 1089-1096 (1996).

[5] C. Monroe, W. Swann, H. Robinson, and C. Wieman, **Very Cold Trapped Atoms in a Vapor Cell**,*PRL*, **65**, pp. 1571-1574 (1990).

[6] P. Lett, R. Watts, C. Westbrook, W. Phillips, P. Gould and H. Metcalf, **Observation of Atoms Laser cooled below the Doppler limit**, *PRL*, **61**, pp. 169-172 (1988).

[7] G. D. Rovera, G. Santarelli and A. Clairon, **A laser system stabilized on the caesium D2 line**, *Rev.Sci.Instrum.* 65,1502-1505 (1994).

PROGRESS IN DEVELOPMENT OF ACTIVE AND PASSIVE HYDROGEN FREQUENCY STANDARDS AT IEM «KVARZ»

N.A.Demidov, V.A.Logachev, A.S.Marakhov, V.A.Fedorov
Institute of Electronic Measurements "Kvarz"
176, Prospect Gagarina, Nizhny Novgorod, Russia

ABSTRACT

This paper reports on the latest achievements in development of active and passive hydrogen frequency standards at IEM «Kvarz».

Keywords: hydrogen maser, frequency standard, frequency stability

1. INTRODUCTION

Hydrogen maser frequency standards have been developed and produced by the Institute of Electronic Measurements (IEM) "Kvarz"for more than 30 years. Since then over 200 hydrogen masers have been produced which are used at Russian National Time Service, including National time and frequency reference, in VLBI (Very Long Baseline Interferometers) of Rissian Academy of Science and at the sites of observation and control of GLONASS system.

At present IEM «Kvarz» is producing the active hydrogen maser CH1-75 and the only in the world passive hydrogen maser CH1-76. Their characteristics and design are presented in [Ref.1].

This paper reports on the latest achievements in improving the frequency stability of the CH1-75 and in development of the new model of passive hydrogen maser featuring smaller size, weight and lower power consumption. It can operate in more severe enviromental conditions than the previous model CH1-76.

2. IMPROVEMENT OF ACTIVE HYDROGEN MASER CH1-75 FREQUENCY STABILITY

The active hydrogen maser CH1-75 has been produced by IEM "Kvarz" since 1991. It has an automatic cavity frequency control (ACFC) system, in which atom line quality modulation method is used.

The block-diagram of this system is shown in Fig.1, photo of the CH1-75 - in Fig.2.

The ACFC system operation is based on measurement of frequency difference of maser at two atom line quality values by means of frequency comparator and reversible counter and cavity frequency control versus a sign of this difference.

The count time of reversible counter T equals 100 s. After one tuning cycle the tuned and the reference masers change their places and cavity tuning of the second maser is carried out.

However, the chosen count time 100 s is non-optimum, which leads to the large value of ACFC system random error and increases a hydrogen maser frequency instability for measurement time intervals from 10 s to 1 day (see Fig.3).

The aim of our investigations is to minimize the frequency stability degradation.

As it follows from Fig.3, the optimal value for T (in the absence of systematic linear frequency drift of hydrogen maser due to cavity aging) is that, at which Allan variance σ_y^2 (2, T) has the minimum value, which is T=1 h.

In the presence of systematic drift of cavity frequency the ACFC system eliminates it, so the following condition must be fulfilled (taking into account that at alternative tuning of two masers the time interval between two successive tuning cycles of one maser is 4T):

$$4T < \Delta f_s / \alpha , \qquad (1)$$

where Δf_s is the frequency step of frequency correction, α is the linear frequency drift coefficient due to cavity aging.

Due to noise the amount of frequency corrections must be increased and it may be assumed, that the following condition for T must be fulfilled:

$$T < \Delta f_s / 12\alpha . \qquad (2)$$

In this case Δf_s value must be chosen small enough ($\Delta f_s \approx \sigma$ (2,T), so that frequency correction should not lead to notable frequency variations in comparison with the case, when cavity frequency is not corrected. If we take $\Delta f_s / f_m = 1.5 \cdot 10^{-15}$, $\alpha = 1 \cdot 10^{-14}$ per day, then $T = 10^3$ s (f_m is maser frequency).

The second factor, which affects ACFC system error, is an atom line quality modulation depth or R factor:

$$R = \Delta Q_L / Q_{Lmin} , \qquad (3)$$

where ΔQ_L is the atom line quality variation in modulation, Q_{Lmin} is the minimum atom line quality value achieved in modulation.

The maximum value of R factor achieved in modulation depends on modulation parameter type and depth.

We have investigated the atom line quality variation by atom beam density modulation and by bringing inhomogeneous magnetic field into a storage bulb.

In the beam density modulation Q_L value is limited up by decreasing maser power and down- by decreasing maser life time due to great beam density.

The experimental investigations have shown, that R factor value equal to 0.5-0.7 is achieved at a beam density variation by 3-4 times.

In the line quality modulation by bringing inhomogeneous magnetic field into a storage bulb R factor value is limited by maser power decrease and frequency shifts. The experiment has shown, that maximum R value, which is achieved in this method, is 0.3-0.5, and a maser tuning frequency by this method differs from that by the beam density modulation method by $(0.5\text{-}3)\cdot10^{-13}$ (for different masers).

The advantage of this tuning method is simplicity of modulation implementation.

R factor value may be increased by 1.5-2 times, if the combination of two line quality variation methods is used - increasing the beam density and bringing the inhomogeneous magnetic field into the bulb together.

The above described changes in ACFC system - the increase of count time of reversible counter and the combined method of atom line quality modulation - have been implemented in the hydrogen maser CH1-75 and frequency stability measurements have been made. The frequency instability $\sigma_y(2,\tau)$ equal $6\cdot10^{-15}$ and $2\cdot10^{-15}$ per 1 h was achieved at T=100 s and 1000 s respectively.

The results of two masers CH1-75 frequency instability (per 1 day) measurements for different modulation methods and for different count time intervals of the ACFC system reversible counter are shown in Table 1.

Table 1

The results of two CH1-75 masers (serial numbers 27 and 28) frequency instability (per day) measurements.

T, s	100	100	1000	1000	1000
Modulation type	magnetic field	beam	magnetic field	beam	magnetic field+ +beam
R_{27} R_{28}	0.50 0.31	0.52 0.46	0.50 0.31	0.52 0.46	0.84 0.64
Observation time, days	15	17	13	18	12
$\sigma_y(2,\tau)\cdot10^{-15}$ (τ=1d)	4.5	4.1	2.5	2.0	1.5

Measurements of frequency instability of CH1-75 masers (serial numbers 39 and 46) with modulation by magnetic field plus beam (R_{39} = 0.72, R_{40} = 1.02) have shown, that $\sigma_y(2,\tau) = 0.7\cdot10^{-15}$ per 1 day on the observation time interval 10 days.

Thus, frequency stability improvement per time interval 1h-1d by 3-4 times is achieved. The obtained results are shown in Fig.3 (curve 3).

3. DEVELOPMENT OF THE NEW MODEL OF PASSIVE HYDROGEN MASER

At present IEM «Kvarz» is producing the passive hydrogen maser CH1-76, which is successfully used in various laboratories both in Russia and in other countries. Further investigation in this area has resulted in the development of the new model of the passive hydrogen maser CH1-86.

The block diagram of the new model is similar to that of the CH1-76, however differences in design and engineering solutions have allowed to reduce dimensions, weight and consumption power with some improvement of metrological characteristics.

In the CH1-86 physics package the quarz cilindrical storage bulb with 0,45 l volume covered with fusible ftorpolimer is used.

The cavity with diameter of 115 mm (instead of 132 mm in the CH1-76), 3 magnetic screens made from 81 NMA permalloy with thickness of 0.5 mm, located inside a vacuum chamber, and the 4-th external magnetic screen from 79 NM permalloy with thickness of 1 mm are used. The influence of external magnetic field variations on the maser frequency is reduced due to reduction of size of shielded area ($1.5\cdot10^{-14}$/ Oersted instead of $2.5\cdot10^{-14}$/ Oersted in the CH1-76).

To increase the reliability of atomic hydrogen source the high frequency oscillator with automatic pulse excitation of discharge in the source bulb is used. This design developed for the CH1-86 is used at present in the serial maser CH1-76 as well.

In designing electronic part of the device the special attention was paid to reduction of temperature variations influence of electronic units on maser frequency. The basic units influencing frequency with temperature change are receiver, frequency multiplier and modulator. The Butterwort passband filter of the 3-rd order with a passband 25 MHz and with passband flatness 0.1 dB is used.

The spurious modulation in the modulator and frequency multiplier has been reduced to a level $-110 \div -120$ dB. All this has allowed to considerably reduce frequency shifts due to electronic units and temperature frequency coefficient, which equals $1 \cdot 10^{-14}$/grad.

The operating temperature range has been extended to 0-40 °C, the weight has been reduced to 38 kgs (instead of 55 kgs in the CH1-76), volume - to 53 l (instead of 75 l in the CH1-76). There is the built-in accumulator battery in the maser, whose capacity is sufficient for device operation during 1 hour.

The rather rigid design of the device allows to use it as a transportable clock. The photo of the CH1-76 and the CH1-86 masers is shown in Fig.4.

The frequency instability of the CH1-86 maser may be approximated by expression $8.5 \cdot 10^{-13}/\sqrt{\tau} + 6 \cdot 10^{-15}$ (see Fig.3, curve 1).

4. CONCLUSION

Thus, thanks to the latest achievements in the development of the hydrogen masers at IEM «Kvarz» the following results are obtained:

- the frequency stability of the active hydrogen maser CH1-75 in an ACFC mode $2 \cdot 10^{-15}$ per 1 h, $(0.7\text{-}1.5) \cdot 10^{-15}$ per 1 d is achieved, i.e. the frequency stability is improved by 3-4 times;
- the experimental prototype of passive maser with frequency stability $6 \cdot 10^{-15}$ per 1 d, small weight (38 kgs) and dimensions, suitable for using as a transportable clock, has been developed.

REFERENCE

1. N.A.Demidov and A.A. Uljanov. Design and industrial production of frequency standards in the USSR. Proceedings of the 22-nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Vienna, Virginia, December 4-6, 1990, p.p. 187-208.

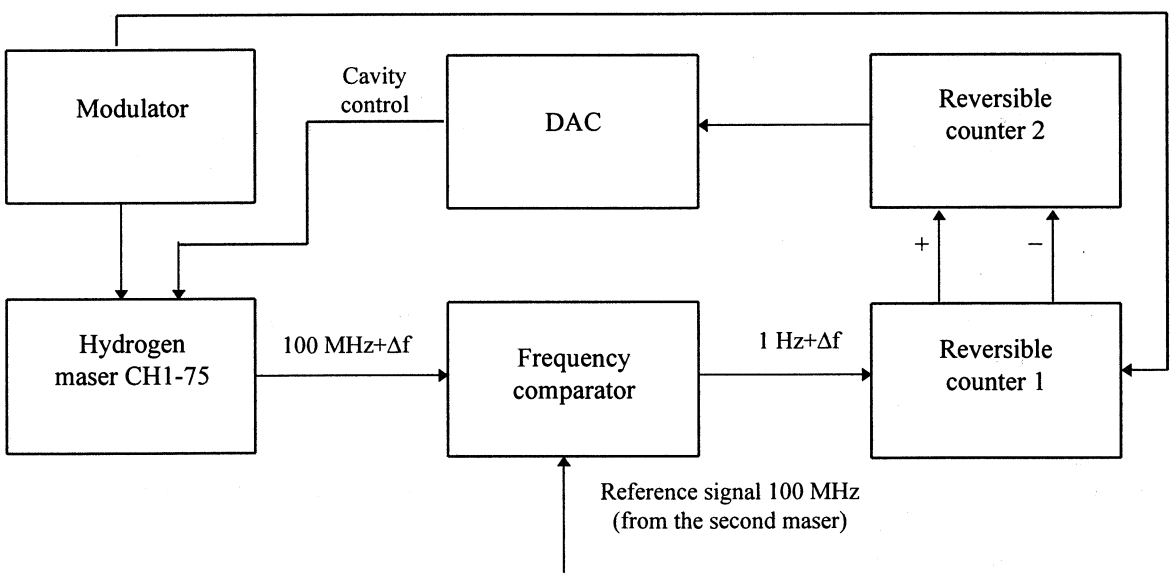


Fig.1. The block-diagram of the ACFC system in the hydrogen maser CH1-75

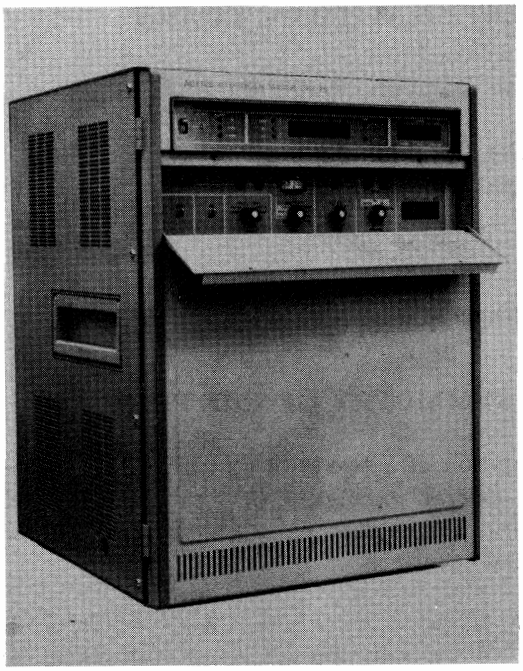


Fig.2. Photo of the CH1-75 hydrogen maser

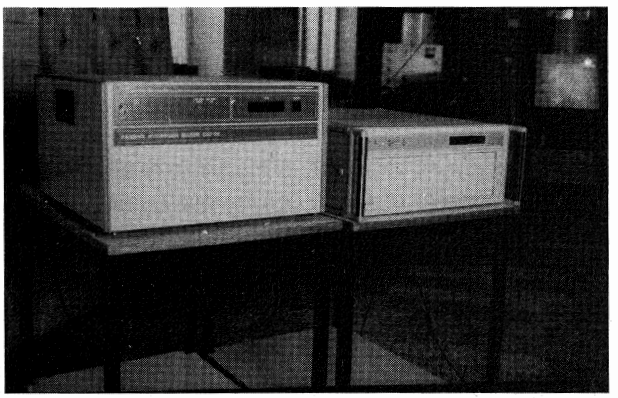


Fig.4. Photo of the passive hydrogen masers CH1-76 (on left) and CH1-86 (on right)

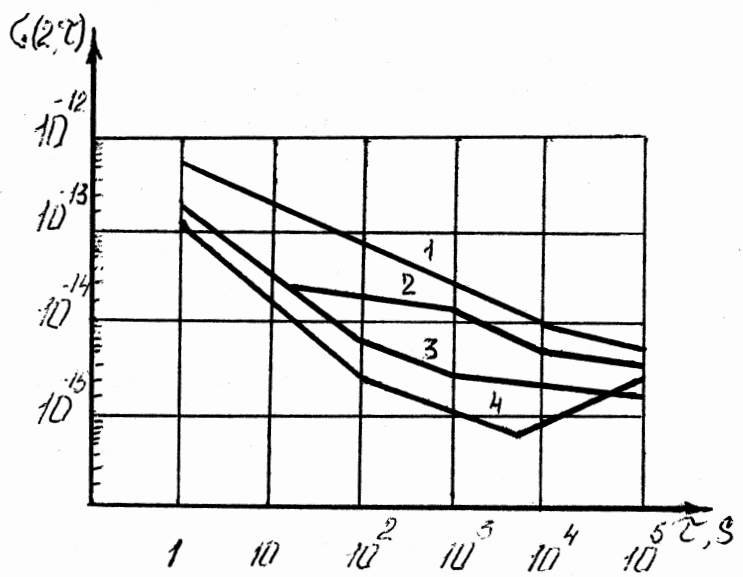


Fig.3. CH1-75 and CH1-86 masers frequency instability:
1 - CH1-86,
2 - CH1-75 (ACFC mode, T=100 s, magnetic field modulation),
3 - CH1-75 (ACFC mode, T=1000 s, magnetic field plus beam modulition),
4 - CH1-75 (free running cavity mode).

RAMSEY FRINGES FOR UNEQUAL INTERACTION TIMES

G Dudle† and P Thomann

Observatoire Cantonal, 58 rue de l'Observatoire, CH-2000 Neuchâtel
†present address: NPL, Teddington, Middlesex TW11 0LW, UK

ABSTRACT

Most frequency standards in which Ramsey's separated oscillatory fields method is used apply two microwave pulses of identical duration to the atoms. Recent proposals to use a continuous beam of laser cooled atoms in frequency standards may, however, result in devices where the interactions are of unequal duration. The purpose of this paper is to establish the symmetry of the Ramsey fringes in such a case. The demonstration is based on properties of the Bloch equations without any assumptions about the temporal dependence of the rf-field. This very general result can therefore also be used for different types of interrogation. In addition, the behavior of the contrast and the width of the central Ramsey fringe as a function of microwave power for both equal and unequal interaction times is also investigated.

Keywords: Ramsey fringes, Cesium standards, Atomic clocks

1 INTRODUCTION

Most cesium frequency standards use the separate oscillating field method first proposed by Ramsey in 1949 [1]. Two microwave pulses separated by a period without rf-field are applied to the interrogated atoms. In a conventional thermal beam tube the two pulses occur in two different regions while the atoms pass through the Ramsey cavity [2, 3]. In atomic fountains the atoms perform a ballistic flight and experience a single rf-field twice, once on their way up and once on their way down [4]. In either case the length of the two microwave pulses is the same. Recent proposals for new designs of frequency standards based on a continuous beam of cold atoms [5, 6] may however result in different interaction times. It is therefore important to establish that the Ramsey fringes remain symmetric with respect to the detuning even if the interaction times are different. In the following section it is shown under what conditions this result is true. The third section contains a study of the behavior of the contrast and the width of the central fringe as a function of the microwave power.

2 SYMMETRY OF THE FRINGES

The evolution of a two-level system in interaction with an external oscillating field can be described by the Bloch equations [3], a system of first order differential equations.

$$\dot{\vec{w}} = \begin{cases} \dot{u} = +\delta v + \Omega_v w \\ \dot{v} = -\delta u - \Omega_u w \\ \dot{w} = -\Omega_v u + \Omega_u v \end{cases} \tag{1}$$

The three elements u, v and w are often regarded as the components of a fictitious spin $\vec{w}$, each of them having a clear physical meaning: u and v are related to the components of the dipole oscillating respectively in phase and with $\pi/2$ phase lead; w is proportional to the population difference between the levels involved ($w = \frac{1}{2}[\rho_{bb} - \rho_{aa}]$ with ρ_{ii} the probability of finding the atom in level $|i>$). The amplitude B of the external oscillating field enters the formula through the Rabi frequency $\hbar\Omega(t) = \mu_B B(t)$ where μ_B stands for Bohr's magneton. Ω_u and Ω_v are kept in Equation (1) to allow for an end to end phase shift between the two microwave pulses ($\Omega^2 = \Omega_u^2 + \Omega_v^2$). Finally, δ designates the detuning between the frequency of the external field and the atomic resonance frequency: $\delta = \omega_{\mathrm{rf}} - \omega_{\mathrm{at}}$.

Equation 1 is a special case of the more general system given in [3] which includes longitudinal and transversal relaxation. For the present purpose the relaxation has been neglected, *i.e.* $T_1 \to \infty, T_2 \to \infty$.

The Ramsey fringes are related to the third component of the fictitious spin $\vec{w}$. To have a symmetrical pattern with regard to δ means that the third component of $\vec{w}$ has the same value after an evolution with $\delta^+ = +\delta_0$ as after an evolution with $\delta^- = -\delta_0$. For the purpose of the demonstration it is helpful to rewrite Equation (1) in terms of $\pm\delta_0$.

$$\dot{\vec{w}}^{\pm} = \begin{cases} \dot{u}^{\pm} = \pm\delta_0 v^{\pm} + \Omega_v(t) w^{\pm} \\ \dot{v}^{\pm} = \mp\delta_0 u^{\pm} - \Omega_u(t) w^{\pm} \\ \dot{w}^{\pm} = -\Omega_v(t) u^{\pm} + \Omega_u(t) v^{\pm} \end{cases} \tag{2}$$

Two systems are contained in Equation 2 which can be combined by computing $\vec{w}_\Delta = \vec{w}^+ - \vec{w}^-$ and $\vec{w}_\Sigma = \vec{w}^+ + \vec{w}^-$. Two independent systems of differential equations can be obtained if the resulting systems are rearranged in the following way:

$$\dot{\vec{r}} \equiv \begin{cases} \dot{u}_\Sigma &= \delta v_\Delta + \Omega_v w_\Sigma \\ \dot{v}_\Delta &= -\delta u_\Sigma - \Omega_u w_\Delta \\ \dot{w}_\Delta &= -\Omega_v u_\Delta + \Omega_u v_\Delta \end{cases} \tag{3}$$

$$\dot{\vec{s}} \equiv \begin{cases} \dot{u}_\Delta &= \delta v_\Sigma + \Omega_v w_\Delta \\ \dot{v}_\Sigma &= -\delta u_\Delta - \Omega_u w_\Sigma \\ \dot{w}_\Sigma &= -\Omega_v u_\Sigma + \Omega_u v_\Sigma \end{cases} \tag{4}$$

The two systems are only independent if $\Omega_v(t) = 0$, *i.e.* if there is no phase shift between the two microwave pulses. It is easy to show that with this condition the norm of the two vectors $\vec{r}$ and $\vec{s}$ does not change in time, *i.e.*

$$\frac{d}{dt}||\vec{r}|| = \frac{d}{dt}||\vec{s}|| = 0. \tag{5}$$

For the rest of this section it is assumed that Ω_v is zero for all times.

The requirement for identical evolution of the third components of $\vec{w}^+$ and $\vec{w}^-$ becomes clear if the initial conditions of the fictitious spin are fixed and the corresponding auxiliary variables $\vec{r}(t=0)$ and $\vec{s}(t=0)$ computed.

$$\left.\begin{aligned} \vec{w}_0^+ &= \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} \\ \vec{w}_0^- &= \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix} \end{aligned}\right\} \rightarrow \begin{cases} \vec{r} = 2 \begin{pmatrix} u_0 \\ 0 \\ 0 \end{pmatrix} \\ \vec{s} = 2 \begin{pmatrix} 0 \\ v_0 \\ w_0 \end{pmatrix} \end{cases} \tag{6}$$

In the event of u_0 being zero at $t = 0$ the norm of $\vec{r}$ is zero. On the other hand, as mentioned above (Equation (5)), this norm does not change with time and will thus remain zero. In particular this means that the third component of $\vec{r}$ ($r_3 = w_\Delta = w^+ - w^-$) which gives the difference of the population differences between the case $+\delta_0$ and $-\delta_0$ is always equal to zero. It is therefore established that — so long as the first component of the fictitious spin is initially zero — the Ramsey fringes are symmetric with respect to δ. This result is surprisingly simple and requires two remarks. First, it must be emphasized that there is no restriction on the initial value of v or of w. The only conditions to be fulfilled are no phase-shift between the two microwave pulses ($\Omega_v = 0$) and $u_0 = 0$. Secondly, it is worthwhile pointing out that the proof relies directly on the differential equation and does not assume any particular interrogation scheme. The symmetry result is therefore correct whatever the particular time dependence of the rf field. It is valid for equal or unequal interaction times (with equal or unequal Rabi frequencies) in any kind of Ramsey cavity (or E-bend or H-bend) as well as in the so-called "multilambda" cavities.

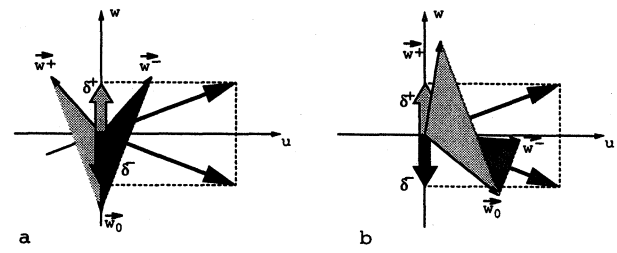


Figure 1: Illustration of the condition $u = 0$.

Figure 1 illustrates the condition for symmetrical Ramsey fringes. Both pictures show the evolution of the fictitious spin $\vec{w}$ for $\delta^+ = +\delta_0$ and $\delta^- = -\delta_0$. The fictitious spin precesses around a vector $\vec{a}$ in the (uw) plane describing thereby a cone the aperture of which is determined by the angle between $\vec{w}_0$ and $\vec{a}$. The vector $\vec{a}$ defining the axis of precession has a u-component related to the rf-amplitude and a w-component proportional to the detuning. In Figure 1 a) the condition $u_0 = 0$ is fulfilled whereas in b) $u_0 \neq 0$. In the first case the fictitious spin describes cones with equal aperture for δ^+ and δ^-. The angular frequency of the precession motion being the same in both cases it is easy to see that at all times the third components of $\vec{w}^+$ and $\vec{w}^-$ are identical. In contrast to this, the cones for positive and negative detuning have different apertures in Figure 1 b). Although the angular frequency of the precession is again the same the third component of the fictitious spin is in general not equal for the two cases.

3 WIDTH AND CONTRAST OF THE CENTRAL RAMSEY FRINGE

Both the width and the contrast of the central Ramsey fringe depend on the amplitude of the interrogating rf-field. The aim of this section is to analyse the variation of these two parameters as a function of the microwave field. Although an overall π-pulse is commonly used to interrogate the atoms, this study gives a useful insight into interesting features of the Ramsey fringes. The following discussion focuses on an E-bend cavity since an analytical solution can be elaborated in the case where the amplitude in each interaction region is constant. Numerical integration of the Bloch equation for an H-bend cavity have however led to analogous results.

The starting point for the analysis is Equation (7), which gives the third component of the fictitious spin after two microwave pulses of length (Rabi frequency) $T_1(\Omega_1)$ and $T_2(\Omega_2)$ respectively, separated by an interval T_L; Equation (7) is obtained

by integrating system (1) with as initial condition $u_0 = v_0 = 0, w_0$ and the two Rabi frequencies Ω_1 and Ω_2 constant but not nececessairly equal. As in the previous section the detuning of the rf-field from the resonance angular frequency is denoted δ:

$$w_f = w_0 \{\kappa \cos(\delta T_L) + \sigma \sin(\delta T_L) + \eta\} \quad (7)$$

The coefficients κ, σ and η stand for the expressions:

$$\begin{aligned}
\kappa &= \frac{\Omega_1\Omega_2}{\tilde{\Omega}_1\tilde{\Omega}_2} \\
&\times \left\{\frac{\delta^2}{\tilde{\Omega}_1\tilde{\Omega}_2}\left(1-\cos\tilde{\Omega}_1T_1\right)\left(1-\cos\tilde{\Omega}_2T_2\right)\right. \\
&\left. -\sin\tilde{\Omega}_1T_1\sin\tilde{\Omega}_2T_2\right\} \quad (8) \\
\sigma &= \frac{\Omega_1\Omega_2}{\tilde{\Omega}_1\tilde{\Omega}_2}\,\delta \\
&\times \left\{\sin\tilde{\Omega}_1T_1\frac{1-\cos\tilde{\Omega}_2T_2}{\tilde{\Omega}_2}\right. \\
&\left. +\sin\tilde{\Omega}_2T_2\frac{1-\cos\tilde{\Omega}_1T_1}{\tilde{\Omega}_1}\right\} \quad (9) \\
\eta &= \left\{\frac{\delta^2}{\tilde{\Omega}_1^2}+\left(\frac{\Omega_1}{\tilde{\Omega}_1}\right)^2\cos\tilde{\Omega}_1T_1\right\} \\
&\times \left\{\frac{\delta^2}{\tilde{\Omega}_2^2}+\left(\frac{\Omega_2}{\tilde{\Omega}_2}\right)^2\cos\tilde{\Omega}_2T_2\right\} \quad (10) \\
&\tilde{\Omega}_i^2 = \delta^2+\Omega_i^2
\end{aligned}$$

In order to define the contrast and the width of the central fringe it is convenient to approximate Equation (7) by the expression:

$$w_f = w_0 \{\zeta \cos(\delta T_L - \psi) + \eta\}. \quad (11)$$

The new variables ζ and ψ are related to κ and σ:

$$\begin{aligned}
\zeta^2 &= \kappa^2+\sigma^2 \quad (12) \\
\tan\psi &= \frac{\sigma}{\kappa}. \quad (13)
\end{aligned}$$

The contrast of the central fringe is defined by the amplitude ζ of the cosine in Equation (11) whereas the width is equal to twice the detuning for which the argument of the cosine is equal to $\pi/2$: $\delta T_L - \psi = \pi/2$ (note that ψ is an odd function of δ (19) and does not imply any shift of the central fringe).

Without restricting the generality of the discussion it is possible to assume identical rf-amplitudes for both pulses. Indeed, Equation (7) contains only the angles of nutation $\phi_i = T_i\Omega_i$ and the interaction times have not yet been fixed.

After the two microwave pulses the overall angle of nutation will be $2\phi_0 = \phi_1 + \phi_2$. The result is most intuitive if the individual angles are expressed in terms of their average ϕ_0 and difference φ:

$$\begin{aligned}
\phi_1 &= \phi_0+\varphi/2 \quad (14) \\
\phi_2 &= \phi_0-\varphi/2. \quad (15)
\end{aligned}$$

Since the discussion applies to the central fringe ($\delta^2 \ll \Omega^2$) further approximations can be made.

$$\left(\frac{\Omega}{\tilde{\Omega}}\right)^2 \simeq 1 \qquad \left(\frac{\delta}{\tilde{\Omega}}\right)^2 \simeq 0 \quad (16)$$

Using these approximations and the previously introduced notations for the nutation angles, the coefficients κ and σ can be rewritten, yielding — according to Equations (12) and (13) — expressions for the contrast and the width.

For equal nutation angles ($T_1\Omega_1 = T_2\Omega_2$), κ and σ take simple forms.

$$\begin{aligned}
\kappa^{eq} &= -\sin^2\phi_0 \quad (17) \\
\sigma^{eq} &= \frac{2\delta}{\Omega}\sin(1-\cos\phi_0). \quad (18)
\end{aligned}$$

For small detunings the contrast as a function of rf-amplitude therefore simply oscillates between -1 and 0 (Equation (18)). While the absolute sign of the contrast is not relevant, it is important to note that in the present case the sign never changes.

As far as the width is concerned, the variable ψ must first be calculated.

$$\tan\psi = -\frac{2\delta}{\Omega}\frac{1-\cos\phi_0}{\sin\phi_0} \quad (19)$$

This term is in general small except for values of ϕ_0 close to an odd multiple of π. By excluding these regions, $\tan\psi$ can be replaced by its argument thus allowing the width of the central fringe to be calculated as a function of microwave amplitude.

Figure 2 summarizes the results of contrast and width as a function of the rf-amplitude. The horizontal scale is normalized to the power producing an overall nutation angle equal to π, *i.e.* $\Omega = 1\Omega_0$ means that $2\phi_0 = \pi$. The vertical axes show on the right the normalised contrast and on the left the full width at half maximum for the given configuration. The geometry is chosen to have interaction times $T_1 = T_2 = 5.8$ ms separated by an interval of $T_L = 66.8$ ms. The grey zones in the figure indicate where the approximation $\tan\psi \simeq \psi$ is invalid.

The variation of the width of the central fringe can be interpreted as a change of the time between the two microwave pulses. In order to make this modulation clearer, the expression for ψ is substituted in Equation (11). The result reads

$$\frac{w_f}{w_0} = \sin^2\phi_0\left\{\cos\left[\delta\left(T_L+\frac{2T_0}{\phi_0}\frac{1-\cos\phi_0}{\sin\phi_0}\right)\right]\right\}+\eta'. \quad (20)$$

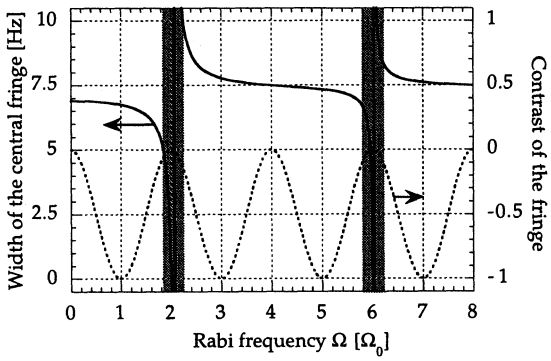


Figure 2: Variation of the width and the contrast of the central Ramsey fringe

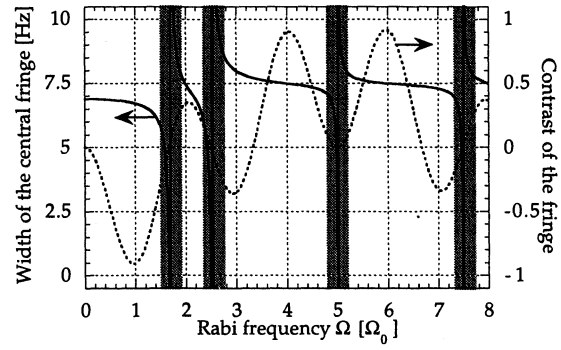


Figure 3: Variation of the width and the contrast of the central Ramsey fringe

For values for which the approximation $\tan\psi \simeq \psi$ holds, one can define T_{eff} such that

$$T_{\text{eff}} = T_L + \frac{2T_0}{\phi_0}\frac{1-\cos\phi_0}{\sin\phi_0} \qquad (21)$$

This time is modified if the microwave amplitude (and hence ϕ_0) is changed and leads to the interpretation of the variation of the width of the central fringe, a phenomenon which has already been recognized earlier [7]. The width is related to the total angle of precession the fictitious spin accumulates between the two microwave pulses. The precession angle itself is connected to the integral over time of the dipole. As soon as the atom enters the first cavity the dipole is slowly built up and the spin immediately starts to precess. By the time the atom leaves the first cavity the fictitious spin has already precessed by a certain angle. For the second cavity the process is inverted: the precession does not stop when the atoms enters but continues until the dipole has completly vanished. It is clear though that the width does not only depend on the time spent between the two microwave pulses but also on the way the dipole is created and canceled in the cavities. For the calculation of precession angle in the microwave cavities one can replace the growth or decrease of the dipole by its average, multiplied by the time spent in each arm. In the case of an E-bend Ramsey cavity with microwave power producing an overall π pulse $T_{\text{eff}} = T_L + 4T_0/\pi$. With an amplitude yielding an overall 3π pulse $T_{\text{eff}} = T_L - 4T_0/3\pi$, *i.e.* a shorter time.

For unequal interaction times the result is more complex. Instead of showing a simple oscillation, the contrast is here the combination of a slow and a fast periodic function,

$$\zeta^{\text{uneq}} = -\sin\phi_1 \sin\phi_2. \qquad (22)$$

Figure 3 summarises the results for contrast and width in this case.

Two new features can be observed as a result of the two oscillation frequencies. First, in opposition to the case before, the sign of the contrast changes with increasing rf-amplitude. Ramsey fringes for different signs are displayed in figure 4. Secondly, the combination of the two oscillation frequencies may transform a fringe that had minimal contrast in the equal interaction time case to a fringe of maximal contrast (compare the contrast in Figure 2 and Figure 3 at $\Omega = 4\Omega_0$). In Figure 4 calculated Ramsey patterns for three different microwave powers are displayed. The interaction times are $T_1 = 7.0$ ms and $T_2 = 4.7$ ms, separated by 66.8 ms. The atoms are assumed to be prepared in the state $F = 3$. If the power is adjusted to produce an overall π-pulse (plain line $\Omega = 1\Omega_0$) the contrast of the fringe is maximum and the atoms are in $F = 4$ after the interactions. Similarly, for $\Omega = 4\Omega_0$ (dashed line) the contrast is also maximal but the atoms are in the $F = 3$ state at $\delta = 0$. The difference of the width between the two cases is easily seen around the detuning $\delta = 40$ Hz where the Ramsey patterns are clearly out of phase. The third case displayed in Figure 4 corresponds to $\Omega = 5\Omega_0$ (dotted line). As predicted, the contrast for the central fringe is zero. It is obvious that for this case it makes no sense to predict the width of the central fringe.

4 CONCLUSION

This paper presents a theoretical study of the Ramsey fringes for unequal interrogation times. The main purpose is to demonstrate under what conditions the Ramsey fringes are symmetric with respect to the detuning δ. It is shown that as long as the first com-

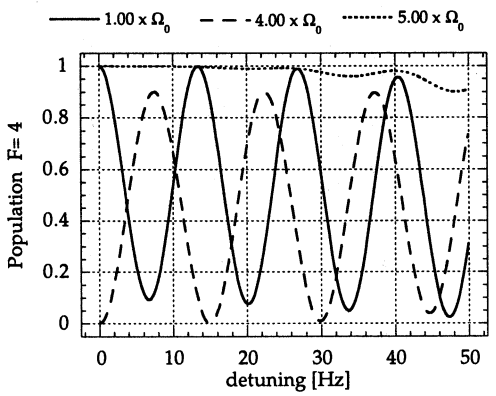


Figure 4: Examples of Ramsey fringes for different rf-powers in the case of unequal interaction times: $T_1 = 7.0$ ms, $T_2 = 4.7$ ms, $T_L = 66.8$ ms.

ponent of the fictitious spin is initially zero ($u_0 = 0$) and no phase-shift exists between the two microwave pulses ($\Omega_v = 0$) the fringes are an even function of δ.

Furthermore a study of the contrast and width of the central Ramsey fringes has been presented. Although most interrogation schemes work with an overall π-pulse this investigation gives some interesting insights of the pattern. Ramsey fringes for unequal interaction times and higher microwave power have been computed.

ACKNOWLEDGMENT

This work has been supported by the Swiss Federal Office of Metrology and the Swiss National Science Foundation. One of the authors (G.D.) is currently the recipient of a grant from the Royal Society.

References

[1] N. F. Ramsey, *Phys. Rev.* **76**, 996 (1949).

[2] A. Bauch *et al.*, *IEEE Trans. Instrum. and Meas.* **36**, 613 (1987).

[3] Jacques Vanier and Claude Audoin, *The Quantum Physics of Atomic Frequency Standards.* Adam Hilger, 1989.

[4] A. Clairon *et al.*, *Europhys. Lett* **16**, 165 (1991).

[5] G. Dudle *et al.*, In *Proceedings of the Fifth Symposium on Frequency Standards and Metrology, 15 - 19 October 1995, Woods Hole, MA, USA.*

[6] N. Dimarcq *et al.*, In *Proceedings of the 9th European Frequency and Time Forum, 8 - 10 March 1995, Besançon,France.*

[7] P. Thomann, In *Proceedings of the 22nd PTTI, Vienna, VA, USA*, 331 (1990).

INFLUENCE OF THE NEIGHBORING TRANSITION ON POLARIZATION GRADIENT COOLING

Kyouya Fukuda, Atsushi Hasegawa, Noboru Kotake, Takao Morikawa and Masatoshi Kajita

Communications Research Laboratory
4-2-1, Nukui-Kitamachi, Koganei, Tokyo 184, Japan
TEL: +81-423-27-6921
FAX: +81-423-27-6694
E-mail: kyo@crl.go.jp

[ABSTRACT]

We discuss the efficiency of polarization gradient cooling of a Cs atom with a laser, red-detuned from the $F_g = 4 \rightarrow F_e' = 5$ transition. The time of flight experiment shows that the polarization gradient cooling works effectively when the detuning is larger than 200 MHz, while it is not effective when the detuning is around 170 MHz. This result is explained considering the influence of the $F_g = 4 \rightarrow F_e' = 4$ transition. We have performed numerical analysis of polarization gradient cooling effect, taking into account two excited states.

1. INTRODUCTION

With the development of the diode laser operating in the near-infrared region, the Cs atom has become one of the most advantageous samples for demonstrating laser cooling, mainly because of its large mass[1]. Furthermore, the development of the laser cooling technique using the Cs atom has been helpful in improving the accuracy and stability of the atomic clock, because the present frequency standard is defined by the hyperfine transition of Cs [2,3].

Employing polarization gradient cooling (PGC), the lowest temperature of a Cs atom becomes much lower than the Doppler limit. Previously, the PGC has been discussed mainly in the case in which the ground state is doubly degenerated. The deceleration by the PGC is proportional to Ω^4/δ^3 , and the equilibrium temperature is in the order of Ω^2/δ, where Ω and δ denote the Rabi frequency and the detuning, respectively.

In this paper, we discuss the cooling efficiency of PGC with an actual Cs $F_g = 4 \rightarrow F_e' = 5$ transition. The characteristic of PGC can not always with the previous simplified model.

2. EXPERIMENT

We measured the temperature (T) of cooled atoms by using a time of flight (TOF) method with various values of laser detuning. Cs atoms were cooled and trapped by a magneto-optical trap (MOT) which was made with three mutually orthogonal pairs of counterpropagating trapping beams having opposite circular polarizations and an inhomogeneous magnetic field. Our experiment was performed in a low Cs vapor pressure cell as in [4]. The trapping beams were derived from a stabilized extended-cavity laser diode and are detuned a few linewidth to the red of the $F_g = 4 \rightarrow F_e' = 5$ transition. These beams were cut off with an acousto-optic modulator. The laser is offset locked to

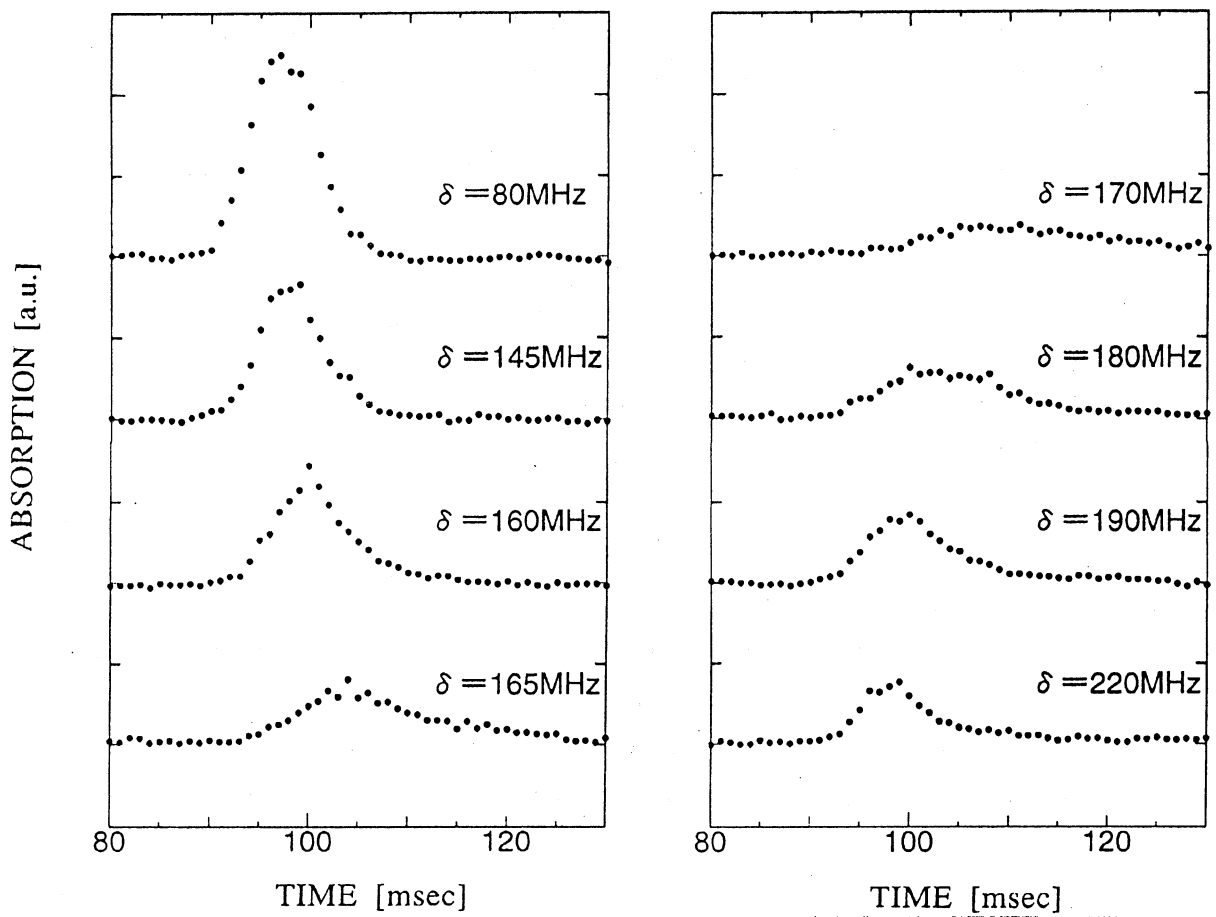


Fig.1: Time of flight (TOF) signal observed with different values of detuning (δ). Here t denotes the time after the cooling laser is turned off.

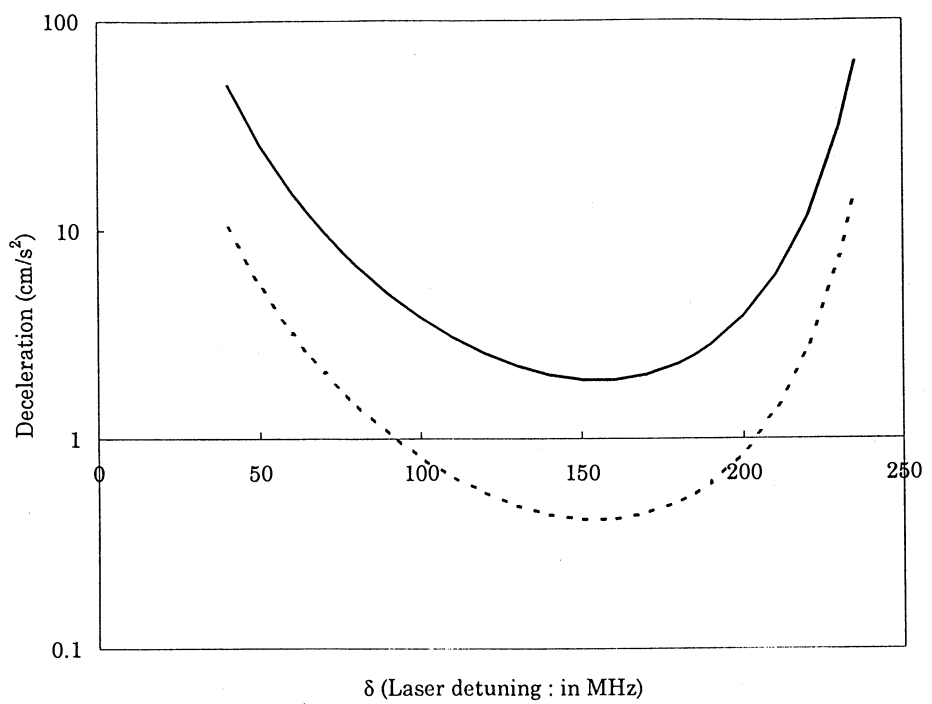


Fig. 2 Atomic deceleration as a function of δ, assuming Ω = 7 MHz, for example velocity is 4 cm/s (solid line) and 15 cm/s (dotted line).

this transition by using a saturated absorption spectroscopy. To prevent optical pumping to the $F_g = 3$ ground state we also excite the $F_g = 3 \rightarrow F_e' = 4$ transition by an another laser diode. To perform a PGC, the locked point of the cooling laser frequency was switched to make detuning much larger. The laser has a power of 13mW with a beam diameter of 12mm (Ω =16MHz). After PGC was performed for 10msec, the cooling laser is turned off. The probe laser beam for TOF method is put 50 mm under the molasses. We measured the population of the atoms passing through the TOF detection region as a function of the intensity of absorption.

3. RESULTS AND DISCUSSION

3.1 Experimental results

Figure 1 shows the experimental result. The TOF signal drastically changes. It is gradually broadened from a detuning is ~-145MHz to ~-170MHz. It seems that the temperature of the molasses increases in this region. Moreover, the peak of the TOF signal shifts to later time. On the other hand, over -170MHz region, the temperature of the molasses seems to decrease until the detuning is ~-220MHz. The peak of the TOF signal again moves to the earlier time. It is noticeable phenomenon. These results cannot be explained by the previous model considering one excited state and the ideal model of the PGC.

3.2 Theory and interpretation

In the case of the Cs atom, the theoretical treatment is much more complicated than the previous treatment, for the following two reasons.

(1) The ground state is formed by nine sublevels.

(2) With PGC, the detuning of the laser is much larger than the natural width. At a detuning of larger than 150 MHz, the laser light functions also as a blue-detuned light of a $F_g = 4 \rightarrow F_e' = 4$ transition.

Taking the sublevels of ground state and two excited states into account, we have performed a numerical analysis of PGC. To our knowledge, this is the first trial taking two transitions into account. We discuss this using a 1-D model with linear⊥linear configuration.

The PGC is caused by the difference of the optical potential at each sublevel. As the first order approximation, the optical potential of ground state atoms concerning two transitions is the sum of potentials calculated with single transitions. Also the transition probabilities to F_e=5 and 4 should be calculated independently. Note that it can be transformed by spontaneous emission not only to F_g=4 but also to F_g=3, when the atom is excited to F_e=4. If it is transformed to F_g=3, it is repumped by another laser that is resonant with the $F_g = 3 \rightarrow F_e' = 4$ transition. Once a transition to F_e=4 occurs, the recoil is caused by more than one photon (in average 12/7 photons), while the recoil after the transition to F_e=5 is caused by one photon. To discuss the efficiency of the PGC, the actual change of kinetic energy was calculated by solving equations given in Ref.[5] numerically.

We have calculated the atomic deceleration taking Ω = 7 MHz and the initial velocity v (for example, at the two cases of the initial velocity is large (v =15 cm/s) and small (v = 4 cm/s)). Figure 2 shows the deceleration as a function of δ. The deceleration is minimum with δ = 160 MHz. Figure 3 shows the equilibrium temperature as a function of δ with Ω=7MHz. Taking the detuning larger than 180 MHz, also the equilibrium temperature becomes lower, because the heating effect by the photon recoil becomes small [5].

From the experimental result, we observed that the deceleration by the PGC becomes minimum at

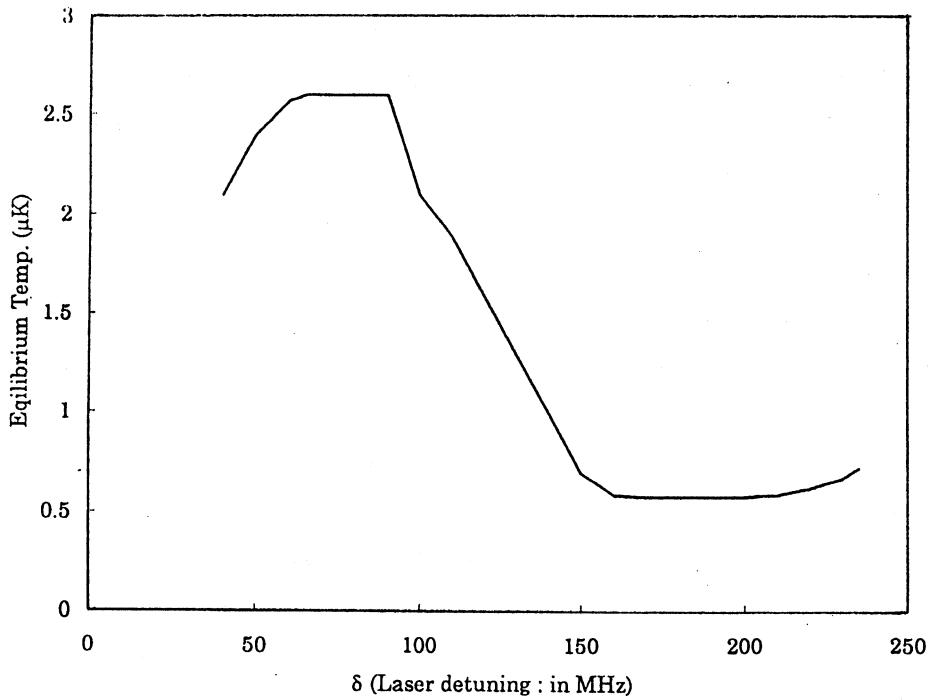


Fig. 3 Equilibrium temperature as a function of δ , assuming Ω =7 MHz. Theoretically, it is preferable to use a detuning larger than 200 MHz, to get both a high deceleration and low equilibrium temperature.

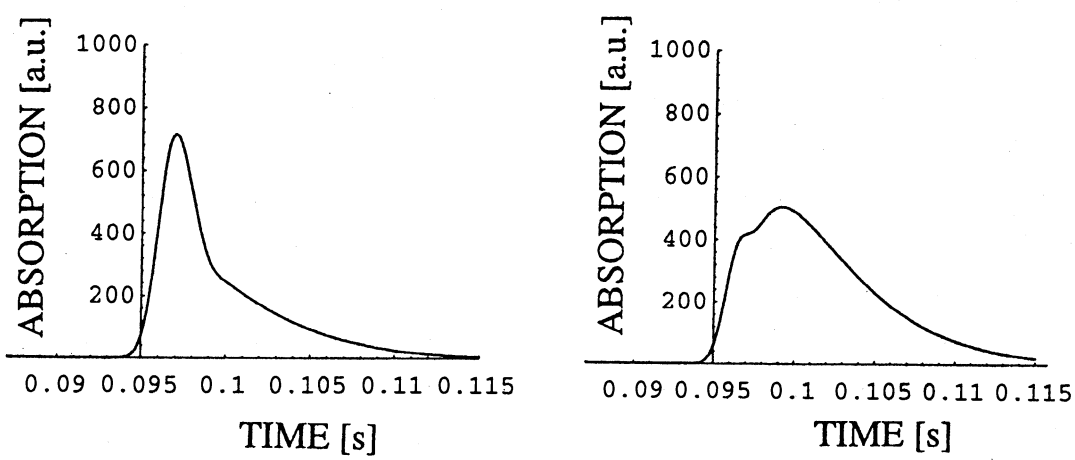


Fig.4 Calculative T.O.F. signal assuming that the initial temperature is 500 μK. Here it is also assumed that the upward velocity is 15 cm/s. The deceleration by the PGC is described by α/v, where v is the atomic velocity and α is a constant. For example, case 1 and 2 show the result with $\alpha = 10^5$ m^2/s^3 and α =200 m^2/s^3, respectively. If the deceleration is large (case 1), the TOF signal is sharp. The deceleration is small (case 2), the peak of the signal shifts and widens.

around δ=170MHz. At this detuning, the temperature does not establish to the equilibrium while the cooling laser is on (10 ms). Therefore, it is considered that the TOF signal becomes sharp when δ<~140MHz or δ>~200MHz, while it is broadened at around δ=170MHz.

Furthermore, figure 1 shows that the peak of the TOF signal shifts when δ is around 170MHz. To explain this phenomenon, we considered that the atoms have a initial velocity of the upward direction ; that is possible if there is power imbalance between up and downward laser beams. At our experimental setup, Cs vapor cell is long to the vertical direction. Therefore, the trap beam is strongly influenced with the absorption. Figure 4 shows the calculative TOF signal, assuming that initial upward velocity is in average 15cm/s and the initial molasses temperature is 500 μ K. This signal includes two components. One of the component is a signal of atoms that establish to the equilibrium while the cooling laser turn on (for 10msec). Another component is a signal which does not establish to the equilibrium for 10msec. In the latter case, atoms are lifted up by the initial velocity and fall down.

If the atomic deceleration changes, the upward velocity at the end of PGC procedure changes too. The deceleration decrease, the upward velocity at 10msec gradually increases. When the atomic deceleration is minimum (the detuning is around 170MHz), the shift of the TOF peak is maximum. Therefore, it is considered that the peak shifts in the TOF signal as a function of the laser detuning. Also in the figure 1, it seems that the signal has two components (for example, at δ =160MHz, one of the component is near 95msec of the time axis, another component is near 100msec).

4. CONCLUSION

We have investigated the PGC characteristics of the Cs atom. We have measured the temperature of cooled atoms with PGC by using a TOF method. The temperature is low (in the order of several microKelvin) when the laser detuning is larger than ~ 200MHz or smaller than ~ 140MHz, while the temperature is high when the detuning is around 170MHz. To explain this phenomenon, we considered the influence of the neighboring transition. It was shown that the PGC is effective when the detuning is larger than ~200 MHz, while it is not effective when the detuning is around 170 MHz. Theoretically it was shown that it is preferable to use a detuning larger than ~200 MHz, to get both a high deceleration and low equilibrium temperature.

References

[1] J. Dalibard, C. Salomon, A. Aspect, E. Arimondo, R. Kaiser, N. Vansteenkiste, and C. Cohen-Tannoudji, in *Proceeding of the 11th Conference on Atomic Physics*, S. Harsche, J. G. Gay, and G. Grynberg, eds. (World Scientific, Singapore, 1989)

[2] J. Vanier and C. Audoin, *The Quantum Physics of Atomic Frequency Standard*, Adam Hilger, Bristol and Philadelphia, pp. 603, (1986).

[3] A. Clairon, C. Salomon, S. Guellati, and W. D. Phillips, Europhys. Lett. **16,** pp.165 (1991).

[4] E. Raab, M. Prentiss, A. Cable, S. Chu, and D. Pritchard, Phys. Rev. Lett. 59, pp2631 (1987)

[5] M.Kajita, N.Kotake, K.Fukuda, A.Hasegawa, and T.Morikawa, Appl. Phys. B to be published

Cs ATOM TRAPPING OUTSIDE OF THE REPUMP BEAM

Atsushi Hasegawa, Kyouya Fukuda, Masatoshi Kajita and
Takao Morikawa

Communications Research Laboratory, Koganei-shi,
Tokyo 184, Japan
e-mail: hasegawa@crl.go.jp

ABSTRACT

We observed that Cs atoms can be trapped even in the region where the molasses is not directly influenced by the electric field of the repump beam. With increasing distance between a narrow repump beam and the trapped atoms, the population of trapped atoms gradually decreases. The limit of the beam separation which can still trap atoms is determined by the molasses size.

Keyword: repump, molasses size, magneto-optical trap

1. INTRODUCTION

The trapping of some neutral atoms by a magneto-optical trap (MOT) has been investigated.[Ref. 1] The MOT of Cs atom is one of the important task. The reduction of the speed of the atom by using a MOT or some other technique is important to improve the accuracy of frequency standards.[Ref. 2] The transition between the two ground hyperfine states (F=3 ⇔ F=4) of Cs is used to define the standard of frequency. It is the so called clock transition. The optical transition between the $6S_{1/2}$ F=4 ground state and the $6P_{3/2}$ F'=5 excited state (cyclic transition) is used for laser cooling. This laser cooling is disturbed by collisions etc. In the weak intensity region of trapping light or by a slight misalignment, it is reported that hyperfine transitions between ground-state atoms is caused.[Ref. 3] If the ground state of the atoms changes to F=3, the optical transition no longer occurs. Therefore, the F=3 ground state is not affected by the cooling process. In order to prevent this phenomenon, we need additional light to return atoms to the F=4 ground state. This is called the repump beam. The characteristics of the repump beam have not been investigated yet. In this paper, we discuss a new feature of the repump beam.

2. EXPERIMENT

The experimental setup is shown in Fig.1. The MOT experiment was performed in a handmade glass cell. Extended-cavity laser diodes were used as the excitation light sources. The linewidth of these laser diodes was less than 1MHz. We performed two experiments with the repump beam. One was narrow beam excitation. In this case, the repump beam was incident on the cell uniaxially and in a different direction from the trap beams. The other was the measurement of the molasses size. In this case, a wide repump beam was used coaxially with one of the trap beams. In this paper, we call

the former case NU (narrow uniaxial) repumping and the latter case WC (wide coaxial) respectively. The NU repump beam was moved into the trap beam region. The magnetic-field gradient was 20G/cm. The profile of the molasses was observed by the CCD camera. The fluorescence from the molasses was collected by a photo-multiplier. In order to increase the signal to noise ratio of the fluorescence intensity, one of the trap beams was chopped and the signal was detected by the lock-in amplifier. The frequency of the trap beam and repump beam were red-detuned from the resonant energy of the F=4 $\leftrightarrow$ F'=5 transition and the F=3 $\leftrightarrow$ F'=4 transition respectively. The temperature of the cooled Cs atoms was $\approx$300 μK, which is measured by the time of flight method.

3. RESULTS AND DISCUSSION

We increased the power density of the repump light by decreasing the beam size. In the coaxial incidence case, the power density of the repump beam is less than 1 mW/cm^2. We have observed that the Cs atoms are still trapped when the power density of the repump beam is increased up to 15 mW/cm^2. This is the maximum density of our system. In this case, the radius of the repump beam is $\approx$ 0.35 mm. Moreover, we moved this NU repump beam away from the center of the molasses. It is noticeable that the Cs atoms seem to be trapped even in the region where the molasses is not directly influenced by the electric field of the repump beam. The picture taken by a CCD camera is shown in Fig.2. The narrow white line is the repump beam and the white dot besides the line is the trapped atoms. The radius of the molasses is $\approx$ 0.4 mm.

In order to quantitatively observe this phenomenon, we measured the intensity of the fluorescence from the molasses as a function of the beam separation. The beam separation means the distance between the repump beam and the molasses. The result is shown in Fig.3. With increasing separation, the population of trapped atoms gradually decreases. Even 3mm from the center, we can observe the fluorescence of molasses. Taking the size of NU repump beam into consideration, it becomes clear that the atoms are not trapped by the residual electric field of the repump light. The Cs atoms are trapped outside of the repump beam.

The trapping region depends on the trap beam intensity.(Fig.4) As the trap light intensity is increased, the region over which atoms can be trapped becomes larger. It also depends on the magnetic-field gradient generated by the anti-Helmhortz coil.(Fig.5) By decreasing the magnetic-field gradient, the trapping region becomes wider. When we change the detuning of the trap light, the trapping region does not change significantly.

We also estimate the size of the molasses in the case of WC repumping by measuring the absorption of the additional probe light. The molasses size increases as the trap laser intensity increases (Fig.6) and the magnetic-field gradient decreases (Fig.7). It does not depend on the detuning of the trap light. These results are reported theoretically by Steane et al. [Ref. 4] The radius of the molasses is given by

$$\langle r_z^2 \rangle \approx \hbar\Gamma(\Omega / \Gamma)^2 (dB/dz)^{-1} (5\Gamma^2 + 4\delta^2) / (\Gamma^2 + 4\delta^2) ,$$

where Ω is the Rabi frequency, Γ the transition linewidth, and δ the laser detuning.

It is found from this equation that the radius of the molasses depends on the magnetic-field gradient and the intensity. Moreover, it does not depend much on the detuning. Our results are in good agreement with this theory.

By comparing the result of the trapping region of the NU repump beam with the molasses size in the case of WC repumping, both results tend to show the same behavior. The trapping region measured by the scanning of a NU repump beam is nearly the same as the molasses size. Therefore, it can be said that the limit of repumping by a NU beam is determined by the molasses size. It can be also said that we directly observe the trapping region in various experimental conditions by using the NU repump beam.

4. SUMMARY

We have observed that Cs atoms can be trapped outside of the repump region. It means that the molasses is not directly affected by the electric field of the repump light. We also measured the molasses size. The trapping region in the case of the NU repumping shows the same tendency with the variation of the molasses size in the case of WC repumping for different trapping conditions. Therefore, the trapping region is determined by the molasses size. The mechanism of this trap is not clarified yet under present conditions. We need more information to do that.

REFERENCES

[1] For example, E.L.Raab, M. Prentiss, A.Cable, S.Chu, and D.E.Pritchard, Phys. Rev. Lett. **59**, 2631(1987).

[2] K. Gibble and S. Chu, Phys. Rew. Lett. **70**, 1771 (1993).

[3] D.Sesko, T. Walker, C. Monroe, A. Gallagher, and C. Wieman, Phys. Rew. Lett. **63**, 961 (1989).

[4] A. M. Steane, M. Chowdhury, and C. J. Foot, J. Opt. Soc. Am. **B9**, 2142 (1992).

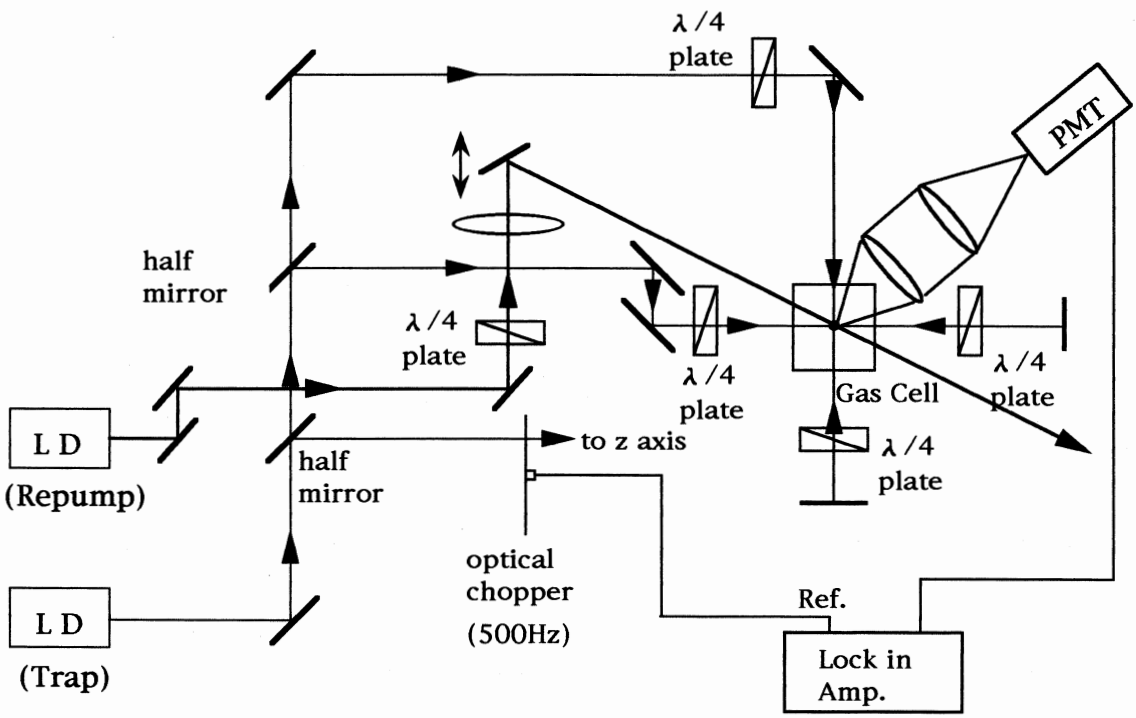


FIG.1. Experimental setup. The repump beam is incident on the cell in a uniaxial configuration.

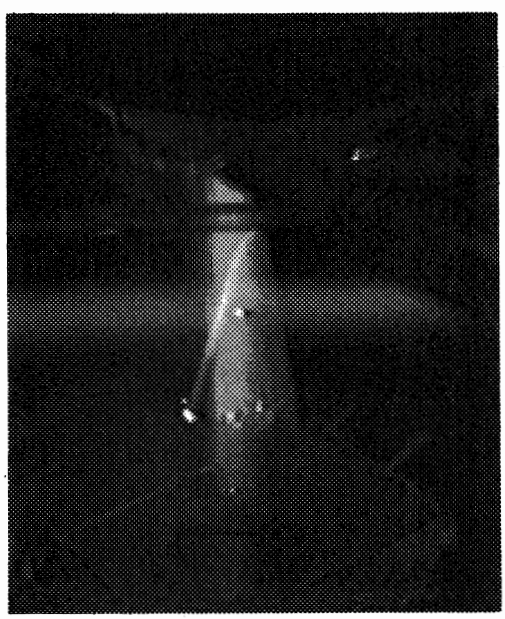


FIG.2 Cs atom trap outside of the repump light. The white line is repump beam and white dot is molasses. Two crossed wide white lines are trap beams. The radius of repump beam is 0.35 mm and molasses 0.4mm. The trap beams are 12ϕ.

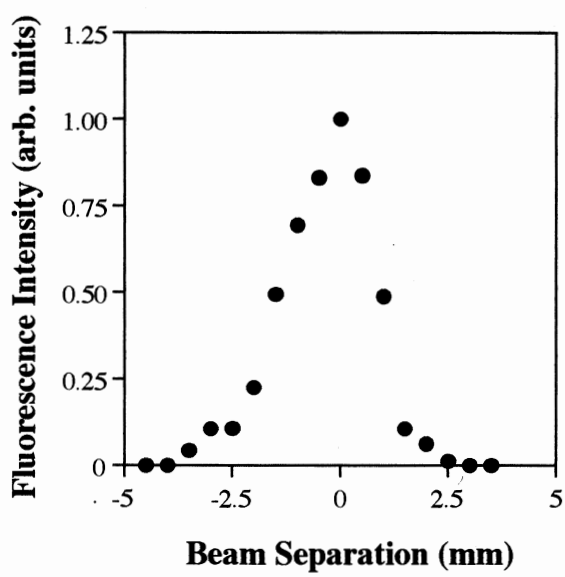


FIG.3 Fluorescence intensity from the molasses as a function of separation of repump beam from the molasses.

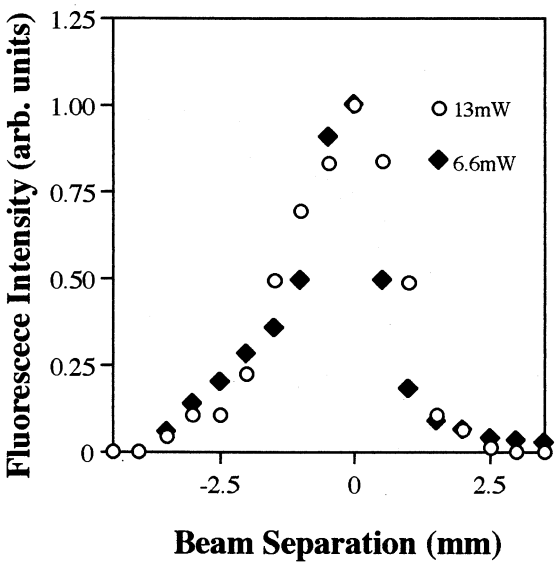


FIG.4 Fluorescence intensity for different trap light intensities.

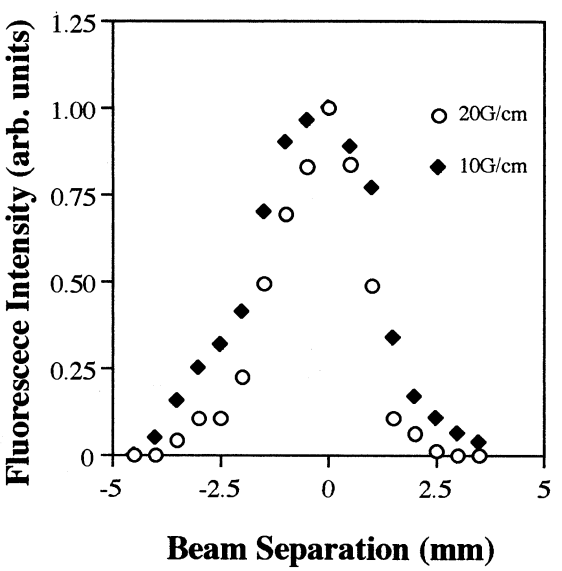


FIG.5 Fluorescence intensity for different magnetic-field gradients.

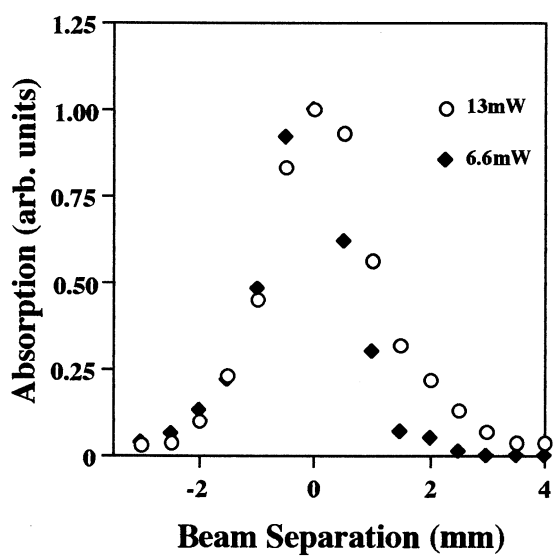


FIG.6 Absorption of the additional probe light to measure the molasses size as a function of separation of probe. Intensity dependence of trap beam light. The zero point of the abscissa represents the highest intensity of the fluorescence from the molasses.

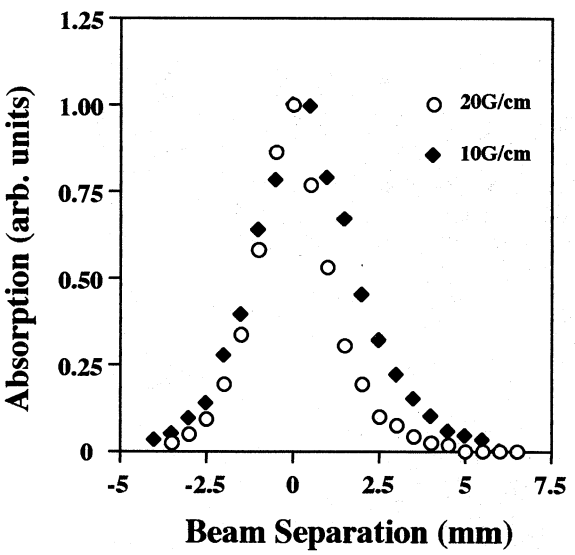


FIG.7 Variation of the absorption for different magnetic-field gradients.

RF-STORED CA^+: PROSPECT FOR A FREQUENCY STANDARD IN THE VISIBLE DOMAIN

Martina KNOOP, Jürgen RINK, Guilong HUANG, Michel VEDEL, and Fernande VEDEL

Université de Provence, Physique des Interactions Ioniques et Moléculaires (UMR 6633 CNRS-UAM1)
Centre de St.Jérôme, Case C21, 13397 Marseille Cedex 20, France

ABSTRACT

Single ions confined in radiofrequency traps offer the best performances for the realization of an atomic frequency standard in the visible domain. We concentrate on the study of the Ca^+-ion, which is one of the most interesting candidates for a metrological application. Preliminary experiments concerning laser-cooling, optimization of experimental conditions and the systematic evaluation of the different neutral gases for line broadening are presented.

Keywords: Ion trap, laser cooling, diode lasers, collisional de-excitation, j-mixing.

INTRODUCTION

Single ion experiments promise to be among the best candidates for the realization of an atomic frequency standard (ref.1). Most of the proposed elements belong to the earth-alkaline group, the ion has then only one outer electron. This gives rise to a simple Λ energy level scheme, where the ion's resonance line (nS-nP) can be used for laser cooling while the (forbidden) clock transition relates a metastable (n-1) D-level to the ground state. Due to a non-negligible (nP-(n-1)D)/(nP-nS) branching ratio (usually a few percent), this configuration requires the recycling of the ions by back pumping of the nP-(n-1)D transition.

The Paul (or radiofrequency) trap has turned out to be the best adapted tool for this kind of high-resolution laser spectroscopy. It offers very long interaction times (hours to days), allowing to neglect transit time broadening effects. There are no interactions with the walls. The influence of collisions with the background buffer gas - even under ultrahigh-vacuum conditions - has to be taken into account, but it can be evaluated (ref.2). Laser cooling which is needed if the clock transition is in the visible domain, allows access to the Lamb-Dicke regime for which the first order Doppler effect vanishes.

THE Ca^+ ION

We started to study a possible use of the Ca^+ ion confined in a Paul trap. The proposed clock transition $4S_{1/2}$-$3D_{5/2}$ with a wavelength of 729 nm has a natural linewidth inferior to one Hertz (ref.3). This gives rise to a theoretical line quality factor $\nu/\Delta\nu$ of about 2×10^{15}. One of the main advantages of the Ca^+-ion is the fact that all the wavelengths needed for the laser cooling and the probing of the transitions can be obtained from laser diodes either directly or by frequency-doubling (Figure 1). A small and compact design for the experimental set-up may then be imagined.

Additionally, calcium has an isotope (^{43}Ca) with an odd nuclear spin (I=7/2) which will allow for applications of extremely high precision to work with atomic transitions which - in first approximation - are independent of magnetic field.

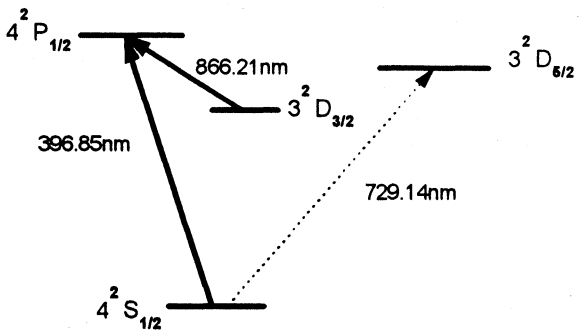


Figure 1: Lowest energy levels of Ca^+.

EXPERIMENTAL SETUP

The experiments described in the following have been carried out in a medium-size Paul trap (r_0=7.2mm). The endcaps of this trap are made from molybdenum mesh with a transmission of 86%, necessary for the observation of the fluorescence of the ions, the ring has been made from stainless steel. The trap is placed in an ultrahigh vacuum vessel at pressures below 7×10^{-10} mbar.

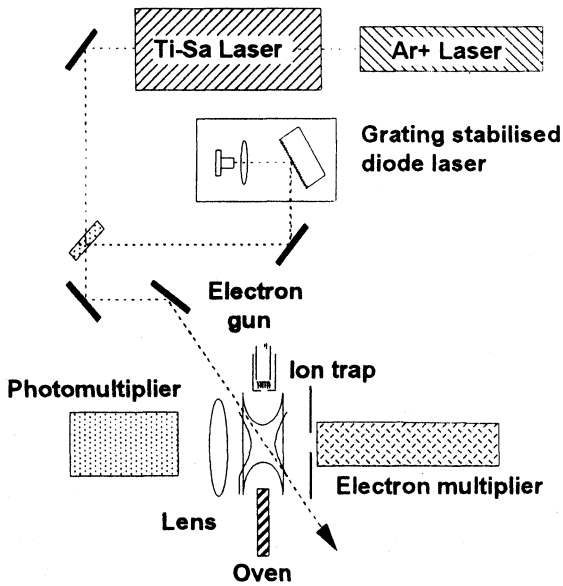


Figure 2 : Experimental set-up

The lasers used for the preliminary laser-cooling experiments are a commercial intra-cavity frequency-doubled titanium-sapphire laser (Coherent 899-21) and a SDL diode laser placed into an external cavity. The TiSa-laser provides up to 15mW of singlemode output power at 397 nm - the wavelength necessary for the laser-cooling of the calcium ions. The diode laser at 866 nm is used to repump the ions which relax into the $3D_{3/2}$-level. It is stabilized on a 300 MHz Fabry-Perot cavity using an electronic feedback method giving a stability better than 200 kHz, the output power used is lower than 2 mW. Both lasers cross the trap diagonally (Figure 2).

OPTIMIZATION OF TRAPPING PARAMETERS

While looking for the best conditions in the stability diagram for the observation of a cooled-ion cloud, we observed that maxima of fluorescence lay on curves almost parallel and close to the experimental limits of the stability. Experiments were carried out on two slightly laser-cooled clouds (T≈100K, ion number close to 20, for a small cloud, and to 200 to 500 for the big one). While moving the working point, it was observed that the ion number was kept (for each experiment we checked that the height of the maximum were not changed, moving back to the point where this maximum was found). However, the fluorescence signal changed as a function of the working point, the variations could reach up to a factor of ten. As mentionned above, the maximum fluorescence points were always found near the boundaries of the stability diagram (β_r=0, for Fig.3).

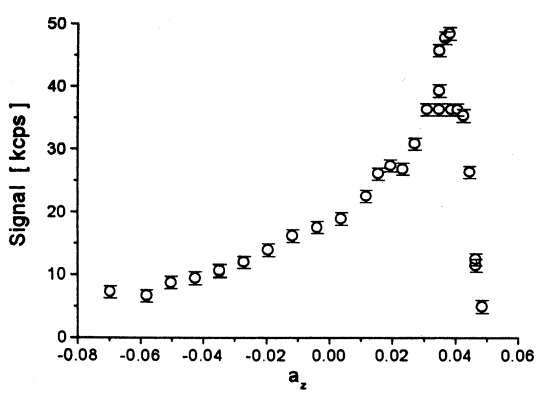


Figure 3 : Fluorescence signal of a small ion cloud as a function of a_z (q_z=0.65).

The real limits of the stability diagram were measured. These limits are defined as the points where the ions are definitively lost. While approaching the limits, the fluorescence signal of the ions drops abruptly (Fig.3 : a_z>0.04), but no ions are lost. The whole ion cloud is lost suddenly for one given a_z-value. This is due to the low temperature of the ion cloud, inferring a kinetic energy very small in comparison with the potential depth, making possible the confinement even in a very shallow potential. The measured limits of the stability region are found inside the theoretical diagram. This behavior has been systematically observed before in various experiments (ref.5) and is foreseen by all the models taking into account anharmonities, either from the trapping field or due to space charge.

Loci of the fluorescence maxima obtained for both clouds are represented in Fig.4 where all the data are put in together. It was checked that for each point the cloud temperature was kept constant (T = 100 ± 20K). For a given cloud, the observed maxima close to one of the boundaries lie almost isobeta lines.

By using a method similar to this one described in ref.5, the ω_x and ω_z secular frequencies of the motion can be measured for each maximum. An additional alternating voltage (so-called « tickle ») of small amplitude is applied to the end caps of the trap (dipolar mode). When the frequency of this tickle is close to one of the motion-frequency spectrum of the axial component, their excursions and their kinetic energy will increase.

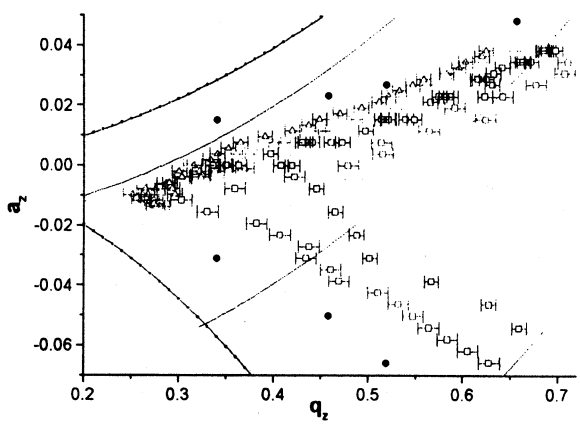


Figure 4 : Lines of maximum fluorescence in the Mathieu stability diagram. The measured limits of the stability diagram for a small ion cloud are represented as (●), the loci of maximum fluorescence differ for a small (Δ) and a big ion cloud (□). Both clouds have approximately the same temperature (T≈100K).

During a resonance, ions absorb energy in a quantity depending on the duration and/or the excitation of the tickle, the corresponding increase of the cloud size leads to a decrease of the density. While scanning the tickle frequency minima of the fluorescence will permit, the location of the motion frequencies (Fig.5). The slight quadrupole contribution allows to excite also the radial frequency (300mV for ω_x, while 10mV are sufficient to evidence ω_z). To monitore the tickle heating and keep the same ion number, the cooling laser was applied red-shifted of 200 to 300 MHZ.

Understanding our fluorescence measurements requires the knowledge of the dynamic properties of the cloud. The power for the cooling laser (1 to 4mW) and for the back pumping diode laser (≈2mW) were high enough to saturate both transition and make sure that the limit temperature is the same, as probed with the Doppler profile measurements done at each involved working point.

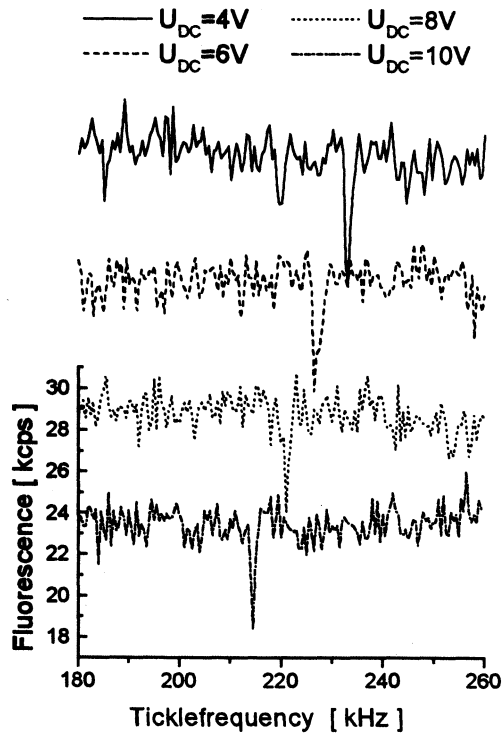


Figure 5 : Fluorescence of an ion cloud as a function of the ticklefrequency applied to the endcaps. The variation of ω_z with the working point is clearly resolved. The measurements were carried out at V_{AC}=402V (q_z=0.65).

Generally speaking, the laser waists are almost or slightly smaller than the ion cloud dimensions (170-200µm for the Ti-Sa laser, 500µm for the diode laser). Because the lasers are oriented oblique regarding the principal axes of the trap, the number of excited ion number decreases when the axial and/or radial cloud dimensions increases, as soon as these dimensions are bigger than the corresponding projection of the waists. This leads to a fluorescence decrease due to the ion density one. Previous studies in our laboratory show that the cloud dimension (σ_x, resp σ_s) increase when approaching the diagram borders (β_x=0, resp β_B=0). This remark allows us to predict the signal decrease when a_z is increasing, as for example in Fig.3. On the other hand, during the deplacements of the point where the fluorescence is measured, the occular frequencies (ω_x, resp.ω_z) vary oppositely considering the geometric parameters. Then, ions will oscillate slowly and will be seen longer by the laser beams. Moreover their velocity will longer compatible with laser frequency (belonging to the ion cloud Doppler profile). The fluorescence will then increase, while moving to the stability limits.

Two effects, the geometric effect and the secular frequency variations are superimposed. Globaly, the fluorescence will increase until the cloud dimension exceeds the laser beam ones, provoking the signal decrease. This geometric factor plays a more important role (effect of 2D), the decrease is faster then than the increase (due to the secular variation frequencies). Since secular frequencies and spatial dimensions follow, for a given component, isobetas, while for the other one, they varied a little, maximas shall be lain on the corresponding isobetas.

When the ion number is bigger, the spatial dimension of the cloud will be also bigger. Actually, for a bigger cloud, cooling the ions will increase the Coulombian interactions more efficient. The cloud dimension will exceed waists for points farer from the boundaries of the diagram. That can be seen in Fig.4, on the side of the β_x=0 line.

Space charge is not big enough (in comparison with the anharmonicities) to render the boundaries really depending on it, as comfirm the measures described above. In the case of the small cloud no maximum were observed near the β_z=0 line. Indeed, the cloud dimensions versus the isobetas - z lines are smaller and they will not reach the threshold of the waist when moving towards the limit.

These measurements show that it is possible to have a good information on the geometric dimensions of the ion cloud, as well as optimal points for the best fluorescence signal exist.

QUENCHING AND J-MIXING

Collisions with the particles of the background buffer gas are one of the major reasons for line broadening under extreme conditions (ref.4). It is therefore necessary to know the rate constants of the different gases present in the ultrahigh vacuum vessel to be able to evaluate systematic effects. Former, we started to look at the influence of the different neutrals (ref.2). We considered two kinds of neutral gases: noble gases like helium, neon and argon, as well as molecular gases (nitrogen, methane and hydrogen).

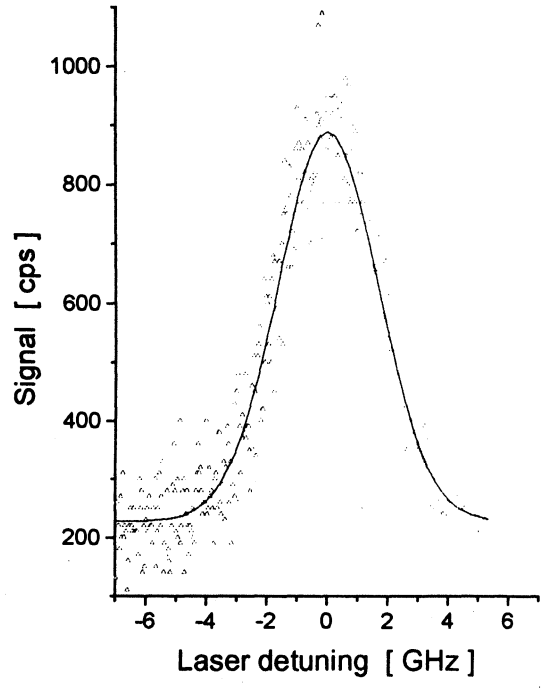


Figure 6 : Excitation of the forbidden $4S_{1/2}$-$3D_{5/2}$ transition at p(CH_4)= $6 \cdot 10^{-8}$ mbar. Laser power at 729 nm is 1060 W.

We directly populate the considered 3D-level, that is we excited one of the forbidden 4S-3D transitions using the fundamental mode of our TiSa-laser (Figure 6). After a given delay the level investigated is probed by excitation of the corresponding 3D-4P transition and observation of the fluorescence in the near ultraviolet (4P-4S). This method assures that only the level investigated is populated while the other atomic states are kept empty. Initial conditions are thus well defined. The obtained decay curves can be fitted with the theoretical equation describing populating and depopulating mechanisms. The applied technique is described in detail in reference 2.

For each of the above mentioned gases we measured the rate constants for quenching (collisional de-excitation of the metastable 3D-doublet to the ground state) and j-mixing (collisional population transfer from $D_{3/2}$ to $D_{5/2}$ level or the inverse). As the pressure in the ultrahigh vacuum vessel is a composition of different gases, each measurement is defined by a set of partial pressures. A multilinear fit can then determine the relative influence of the different gases on the reduced lifetime of the atomic state.

The values found for the collisional de-excitation of the Ca^+ 3D-doublet towards the ground-state are represented in

Table 1. Due to the small energy difference between the $3D_{3/2}$ and the $3D_{5/2}$-level, the rate constant of both states are considered as being identical. Here it is most evident that molecular gases play a very important role for quenching the 3D-doublet. On the other side, atomic gases as helium or neon tend to provoke mostly elastic collisions while having a minor influence on the inelastic de-excitation process.

Gas	m [amu]	Γ [$10^{-12}cm^3s^{-1}$]
H_2	2	(37±14)
He	4	(1.05±0.4)
CH_4	16	$(54\ ^{+91}_{-17})$
Ne	20	(0.9±0.7)
N_2	28	(170±20)
Ar	40	(29.5±17)

Table 1 : Quenching rates for the Ca^+ 3D-doublet for different atomic and molecular gases.

The caracteristics found for the quenching process are confirmed for the j-mixing collisions. In any case the j-mixing rate constants are at least one order of magnitude higher than the quenching rates. This is in part explained by the energy difference of the levels involved (see ref.2), a rough estimation can be made using the Landau-Zener model. The measured j-mixing values are reported in Table 2.

Gas	γ_{35} [$10^{-10}cm^3s^{-1}$]	γ_{53} [$10^{-10}cm^3s^{-1}$]	$\frac{\gamma_{35}}{\gamma_{53}}$
H_2	--	(3±2.2)	--
He	(2.24±0.1)	(1.2±0.7)	(1.9±0.9)
CH_4	(34±14)	(41±12)	(0.85±0.6)
Ne	(5.5±2.2)	(1.3±0.6)	(4.2±3.7)
N_2	(28±3)	(12.6±1.0)	(2.2±0.45)
Ar	(7.7±3.4)	(2.7±1.1)	(2.85±2.4)

Table 2 :j-mixing values for the 3D-doublet of the calcium ion. γ_{35} stands for the $3D_{3/2}$ versus $3D_{5/2}$ mixing rate, and γ_{53} for its inverse.

The principle of detailed balancing gives a theoretical value for the ratio of γ_{35} to γ_{53} of 1.5. A detailed discussion of the values obtained as well as comparison with existing ewperimental an theoretical work is given in Ref.7

OUTLOOK

To realize the probing of the clock transition, we start to stabilize a laser diode at 729 nm. Unfortunately, in this wavelength range, only multimode laser diodes are commercialized, which requires to mount the diode in a grating-stabilized external cavity (ref.8). It is the only possibility to render the diode output singlemode and to obtain the desired wavelength. Further stabilization of the diode is important once the ion stored in the radiofrequency trap has attained the Lamb-Dicke regime (see below). An additional electronic feedback stabilization on a 300MHz Fabry-Perot reference cavity has given stabilities below 1MHz. Optical feedback on the same cavity is a more difficult technique, due to mode competition between the different cavities. Nevertheless the stability performances obtained are of the same order of magnitude as for the electronic feedback. For future applications we plan to use the FM sideband modulation stabilization technique (ref.9) which should allow to attain the laser linewidths necessary for a high-resolution application.

The challenge of a metrological ion trap experiment is the reduction of the observed linewidth, which is mainly due to first order Doppler broadening. One important advantage of the radiofrequency storage is the fact that it opens the way to the Lamb-Dicke regime: If the ion's excursion amplitude in the trap becomes smaller than the emitted wavelength a frequency modulation effect will lead to a discrete frequency

spectrum (ref.10). This will suppress the influence of the first-order Doppler effect. The emitted spectrum is then composed by the center wavelength (ν_0) as well as side-bands separated by integer multiples of the ion's frequency of motion ($\nu_0+\omega$, $\nu_0-\omega$, $\nu_0+2\omega$, $\nu_0-2\omega$...). These frequencies will only be broadened by the second order Doppler effect ($\sim v^2/c^2$). One of the practical conditions of attaining the Lamb-Dicke regime is to have high frequencies of motion ($\geq$ 5 MHz) and a cold (single) ion. For technical reasons this is easier to fulfill by reducing the dimensions of the trap.

We are now building a miniature ion trap system ($2r_0$=1.4mm) to be used for high-resolution spectroscopy. The reduced dimensions do not permit to machine an ideal (hyperbolic) trap. For reasons of simplicity we chose a ring geometry which has been tested successfully elsewhere (ref.11). It has the advantage of being a very open structure allowing to pass the different laser beams and to collect the fluorescence signal of the ion within a very large solid angle.

Ref.1 R.BLATT, P.GILL, R.C.THOMPSON, J.Mod.Opt. **39**, 193-220 (1992)

Ref.2 M.KNOOP, M.VEDEL, F.VEDEL, Phys.Rev.A **52**, 3763 (1995)

Ref.3 G.WERTH, Proc. Summer School "Appl. Laser Spectr.", San Miniato, W.Demtröder and M.Inguscio (eds.), Springer, Berlin 1989

Ref.4 R.ALHEIT, C.HENNIG, R.MORGENSTERN, F.VEDEL, G.WERTH, Appl.Phys.**B61**, 277 (1995) and M.VEDEL, J.ROCHER, M.KNOOP, J.RINK, F.VEDEL, submitted

Ref.5 F.VEDEL, Int. J. Mass Spectrom. Ion Processes and ref. therein, **106**, 33 (1991)

Ref.6 D.J.WINELAND, J.C.BERGQUIST, J.J.BOLLINGER, W.M.ITANO, F.L.MOORE, J.M.GILLIGAN, M.G.RAIZEN, D.J.HEINZEN, C.S.WEIMER, C.H.MANNEY, Proc. Intern. School of Physics "Enrico Fermi", Course CXVIII, E.Arimondo, W.D.Phillips, F.Strumia (eds.), North-Holland, Amsterdam 1992

Ref.7 M.KNOOP, M.VEDEL, F.VEDEL, to be published

Ref.8 C.E.WIEMAN and L.HOLLBERG, Rev.Sci.Instr. **62**, 1 (1991)

Ref.9 R.W.DREVER, J.L.HALL, F.V.KOWALSKI, J.HOUGH, G.M.FORD, A.J.MUNLEY, and H.WARD, Appl.Phys. B**31**, 97 (1983)

Ref.10 R.H.DICKE, Phys.Rev. **89**, 472-473 (1953)

Ref.11 C.A.SCHRAMA, E.PEIK, W.W.SMITH, H.WALTHER, Opt.Comm. **101**, 32-36 (1993)

RUBIDIUM FREQUENCY STANDARD GOES TO DEEP SPACE

Rudolf Kohl *, Michael Bird **

* Dornier Satellitensysteme GmbH, POB 801169, D-81663 München, Germany
** Radioastronomisches Institut der Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany

1. ABSTRACT

A compact and highly reliable Rubidium frequency standard has been developed for the Doppler Wind Experiment (DWE) of the joint NASA/ESA deep space mission Cassini/Huygens to Saturn and its moon Titan. The primary scientific objective of the experiment is to measure the direction and strength of Titan´s zonal winds with an accuracy better than 1 m/s. The Probe´s wind-induced horizontal motion will be derived from the residual Doppler shift of its S-Band radio link to the Cassini Orbiter. The DWE instrumentation consists of a Rb-based Ultra-Stable Oscillator (USO) used to generate the transmitted signal from the Probe, and from another USO used to extract the frequency of the received signal on the Orbiter. The capabilities of these USOs under the rugged environmental conditions on Titan and some results from the DWE-USO acceptance tests and Huygens pre-launch test program will be described.

Keywords:
Rubidium frequency standard, Ultra-Stable Oscillator, Doppler Wind Experiment, deep space mission, rugged environment, test results.

2. INTRODUCTION

DWE is one of six investigations to be performed during the Titan atmospheric descent of the ESA Huygens Probe. DWE is a high-precision tracking investigation to determine the direction and magnitudes of the winds in the Titan atmosphere. The prime science objective, a height profile of the wind velocity, will be derived from the Doppler shift of the Probe Relay Link (PRL) signal from the Huygens Probe to the Cassini Orbiter. After correcting for all known Doppler shifts due to orbit and propagation effects, the wind-induced motion of the Probe will be determined to an accuracy better than 1 m/s, commencing with parachute deployment at an altitude of ca. 160 km down to impact on the Titan surface.
Secondary scientific objectives of the experiment are to monitor the probe descent dynamics like spin and parachute swing, to establish position and orientation of the Huygens probe at and after impact on Titan and to determine strength and spatial scales of turbulences in the atmosphere by measurement of Doppler modulation, Doppler variations and fluctuations.
The most severe constraints on the accuracy of the DWE wind measurement are trajectory errors and instability of the probe oscillator used to generate the PRL signal. It is expected from tolerance analyses that a zonal wind height profile can be recovered with a mean error less than ± 1 m/s. This can be achieved only with a highly stable PRL signal ($df_o/f_o \leq 2e\text{-}10$ or $df_o \leq 0.4$ Hz at S-Band) over the duration of the descent in order to exclude contamination of the measurement by oscillator drift. The frequency stability of the transmitter is guaranteed by using an Rb-based USO (TUSO) to generate the PRL carrier signal. In addition to this transmitter USO, it is necessary to incorporate an additional USO as local reference oscillator into the receiver (RUSO) on the Cassini Orbiter, where the frequency measurement is recorded.
The TUSO/RUSO combination represents the first use of Rubidium based oscillators on a deep space planetary mission.

3. DOPPLER WIND EXPERIMENT - DWE

The Huygens descent mission on Titan is presently scheduled to occur on 27 November 2004, about five months after Cassini arrives at Saturn and seven years after launch. Approximately 22 days prior to its mission, the Probe is separated from the Orbiter and targeted for entry into Titan´s atmosphere. Fig.1 shows the Probe descent profile. The Probe is decelerated at atmospheric entry up to a maximum of 16.1 g at an altitude near 250 km. Parachute deployment at a speed near Mach 1.5 marks the start of the descent phase (time = t_o). A smaller drogue parachute is deployed at t_o + 15 minutes in order to decrease the descent time. DWE measurements commence upon establishment of the radio link shortly after the parachute is deployed, and the decelerator and the nose cap are released. The link will continue for at

least 135 ± 15 minutes until touchdown on Titan. In case of an impact on a solid surface, an impact deceleration of 40 g (approximately 10 g if in a liquid) is expected.
The DWE experimental configuration is shown in Fig. 2. The TUSO is used to generate the radio signal of Transmitter A. An internal TCXO oscillator serves as back-up in the event of a TUSO failure. The TUSO output frequency at 10 MHz is multiplied by 204 to S-Band and transmitted to the dedicated Probe receiver A in the Orbiter. The semiannual cruise checkouts can be conducted via the RF-BITE across the umbilical separation mechanism (USM). At Titan the signal is amplified (HPA) for free-space transmission via the Probe Antenna (PTA) to the Orbiter's high gain antenna (HGA).
Timing and signal generation for the Receiver A are controlled by the RUSO. In order to maintain interchangeability, the RUSO was fabricated as an exact clone of the TUSO. Consistent with the Probe strategy, an integrated TCXO can be substituted for the RUSO if needed. The phase-locked loop control in Receiver A is monitored by a numerically controlled oscillator (NCO), the output of which is recorded to provide the DWE frequency measurement at a sample rate of 8 Hz. The signal level is monitored via the AGC in parallel at the same sample rate. In addition to these DWE "science data", temperatures and internal lock status of TUSO and RUSO are recorded as housekeeping data.
The TUSO will be powered well in advance of the start of transmission from the Probe (≈ 30 minute head start), in order to warm-up and achieve the required frequency stability. The RUSO will be switched on even earlier, together with the rest of the receiver on the Orbiter.
The USOs are designed to withstand the Cassini/Huygens launch accelerations and vibrations, a 7-year interplanetary cruise phase leading to a total radiation dose of more than 10 krad, the above mentioned decelerations during Probe entry as well as significant changes in temperature (ΔT = 15°C) and pressure (ΔP = 1.6 bar) during Probe descent.
The most critical factors for the TUSO, and major drivers in the selection of an Rb-based USO, are the peak deceleration of up to 16.1 g during the entry phase, the pressure variation during descent and the short warm-up time of 30 minutes to reach and maintain a frequency drift stability better than 2e-10. The robustness against mechanical stresses and the quick achievement of frequency stability cannot be guaranteed with quartz crystal oscillators. The high mechanical load during entry might cause a deformation of the internal quartz fastening system in combination with an unpredictable frequency offset and an unknown frequency relaxation time. A similar problem exists with continuously varying mechanical stresses on the quartz housing.
The total frequency drift stability dfo/fo under the rugged environmental conditions on Titan is assumed to be given by the root sum of squares of the frequency drifts from each individual source. The four specific factors affecting the frequency stability are changes in temperature, pressure, acceleration and magnetic field. With the worst case excursions expected during the Huygens descent phase (see Table 1) and the environmental sensitivities given in Table 2, the total frequency drift stability of the DWE TUSO and RUSO can be calculated to be less than 1.9e-10, which is still within specification.
More details on the DWE concept and mission are provided in Ref.1.

		TUSO	RUSO
Temperature:	ΔT	15 °C	5 °C
Pressure:	ΔP	1.6 bar	0 bar
Acceleration:	ΔA	2 g	0.1 g
Magnetic Field:	ΔG	1 mG	0.1 mG

Table 1: Worst Case Environmental Variations

4. ULTRASTABLE OSCILLATOR - USO

The DWE USOs are compact, atomic resonance frequency-controlled oscillators. The highly frequency-stable output signal is obtained from a 10 MHz voltage controlled crystal oscillator (VCXO), whose frequency is referenced and locked to the atomic resonance frequency at approximately 6.834 GHz provided by the ground-state hyperfine transition of ^{87}Rb. Figure 3 represents the detailed block diagram of the DWE USO with the functional blocks on circuit board level.
Core of the USO is the Physics Package with the functional elements
a) Rb-lamp with lamp coil and heater
b) resonator containing Rb-cell, frequency multiplier, microwave couple loop and photo detector
c) magnetic-field coil
d) common heater for resonator and quartz crystal
e) and the crystal itself
all mounted in a common magnetic shield made from Mu-metal.
The Rb-lamp, which provides light by an radiofrequency-excited plasma discharge is mounted in a temperature-controlled (≈113°C) housing. The coil located around the lamp is part of the exciter circuit. The remaining elements of the lamp oscillator, which operates at a frequency of about 83 MHz, are placed on the lamp board. The Rb-cell filters and resonates with the light from the lamp. The corresponding dip in the light intensity is

detected by the photo cell which provides the loop error signal to the servo board. The resonator and the stress compensated (SC-cut) crystal oscillator are operated at 78°C. Lamp and cell are actively heated, whereas the crystal is slaved to the resonator. However, independent thermal sensors are provided for all three components.
A frequency compensating magnetic control loop has been implemented to improve and speed-up the frequency stability during warm-up (see Fig. 6). Significant improvement was achieved by adapting the magnetic field strength via a feedback loop using the thermal sensor in the crystal oscillator housing.
Starting at the upper left and moving clockwise in Fig.3, the other functional blocks alongside the Physics Package are:

a) Lamp Board: controls heating of the Rb lamp and provides RF energy for the lamp excitation
b) Oscillator Board: provides the 10 MHz sinusoidal output signal via VCXO through a buffer amplifier
c) Servo Board: synchronously demodulates the photo detector current and generates the error signal for the VCXO control. Additionally, lock status information is provided.
d) Synthesizer Board: generates all frequencies required for signal processing from the VCXO output, and performs phase modulation of the input signal to the step recovery diode of the frequency multiplier.
e) DC/DC Converter Board: converts the primary supply into secondary voltages and provides floating supply lines and insulation between primary and secondary voltages. The converter switching frequency is synchronized to the USO 10 MHz output signal.

The small volume, low mass and low power consuming USOs have been realized by the application of advanced technological concepts. The Physics Package, shown in Fig. 4, is encapsulated by a low thermal conductive foam within a two-layer inner housing of the USO. The inner housing forms the second magnetic shield required for the Physics Package. Proper stiffness of the inner box and simultaneously low mass is achieved by using only a thin Mu-metal layer which is mechanically supported by a second layer made from aluminium. The electronics, grouped into functional blocks, are arranged on five printed circuit boards. The physical arrangement with the printed circuit boards surrounding the inner housing of the USO can be seen in Fig.5. Radiation sensitive parts like bipolar transistors and MOS-field-effect transistors are encapsulated by tantalum caps that provide shielding in addition to the structural elements of the USO.
More details on the USO functional design and operating principles are given in Ref.2.

The actual physical and electrical characteristics of the DWE-USO are presented in Table 2.

Long term frequency stability	< 2e-11/month
Freq. drift 30 min after switch on	< ± 2e-10
Freq. Setting Accuracy	< 1e-8
Short term stability	< 2e-11 for τ = 1s
	< 6e-12 for τ = 10s
Phase Noise [dBc/Hz]	< -80 for df=1Hz
	< -110 for df=10Hz
	< -130 for df≥100Hz
Frequency Variations with:	
- temperature from -10 to +40 °C	< ± 3e-12 / K
- pressure	< ± 1e-13 / mbar
- acceleration	< ± 2e-12 / g
- magnetic field	< ± 3e-12 / G
- radiation	< ± 1e-14 / rad
Output Signal Level [dBm]	0 ± 2 into 50 Ohm
Harmonics [dBc]	< -36
Spurious [dBc]	< -60 up to 50 MHz
Return Loss [dB]	< -30
Supply Voltage [V]	22 to 43
DC current [mA]	< 675
Warm-up Power [W]	< 18.4 for t<30min
Steady state Power [W]	
at -30°C in air	11.5
at 25°C in air	8.7
at -30°C in vacuum	8.8
at 25°C in vacuum	7.1
Warm-up time at 25°C [min]	
at 25°C in air	< 16
at 25°C in vacuum	< 12
Mass [g]	1898 ± 2
incl. 150g for radiation shielding	
Dimensions [mm]	170 x 117 x 118
	(2.3 ltr.)

Table 2: DWE-USO Characteristics

5. DWE- USO TEST PROGRAM

A total of six Rb USOs have been developed and built by Dornier Satellitensysteme (DSS) in Ottobrunn (near Munich), Germany. The Rb glassware, lamp and cell, has been supplied by Efratom Elektronik, also a German company, after a burn-in and aging performance test.
A single structural and thermal model (STM) was delivered in April 1994. Two engineering models (EM), one TUSO and one RUSO, were delivered in November 1994 to the Huygens / Cassini program for system testing. Three interchangeable units of flight models (FM), TUSO and RUSO FM and one qualification flight spare (QFS) were delivered in March 1996.

Some examples of measurement results achieved during the acceptance and qualification test programs at unit and system level are displayed in Figures 6 - 8.
The behaviour of the USO output frequency as a function of time from switch-on is shown in Fig. 6 for two engineering test models, one with and one without a temperature-compensating magnetic control loop (refer to section 4). After reaching its asymptotic value, the output frequency of both models easily stays within the 2e-10 tolerance band. However, the model without magnetic compensation was found to require too much time to warm-up.
Figure 7 shows a short-term stability result from a typical test run at unit level. Running measurements of the Allan variance are made with integration times of 1 and 10 sec., respectively. Each point in Fig. 7 represents 100 such measurements.
An example from the system end to end tests is shown in Fig.8. The radio link frequency derived from the recorded NCO control word is shown as a function of time from start of the test. The first part of the test was performed using the internal TCXOs in the transmitter/receiver chain A. The TUSO/RUSO combination was switched on from standby to operation at the time t = 104 min until t = 168 min. The characteristics of the received frequency change significantly in this time interval from a series of discontinous jumps ($df_o/f_o \approx$ 1e-7) and unpredictable drifts to a stable value near zero frequency offset.

6. CONCLUSIONS

The Huygens Doppler Wind Experiment is designed to determine the velocity of Titan´s zonal winds. The wind is measured over a height range from 0 - 160 km from ist Doppler signature on the Probe´s radio relay signal to the Cassini Orbiter. Rubidium USOs have be selected as reference oscillators for the Huygens transmitter and for the Cassini receiver due to their excellent frequency stability (df_o/f_o < 2e-10) and insensitivity to their environment. The DWE USOs are able to cope with rapid variations of the ambient conditions, like temperature, pressure, acceleration and mechanical shocks during the 3-hour Huygens Probe mission on Titan. Minimum aging under heavy cosmic ray dosages (10 krad) is expected during the cruise phase of 7 years from Earth to Saturn and Titan. A warm-up time shorter than half an hour can be achieved with warm-up - and steady state dc power limited to 19 W and 12 W, respectively. A sophisticated thermal, mechanical and electrical design resulted in low mass (1.9 kg) and volume (2.3 ltr).

Unlike precision quartz oscillators, which suffer frequency stability degradation in such hostile environments, the DWE USOs have been demonstrated to withstand rigorous qualification and acceptance testing.

ACKNOWLEDGMENTS

This paper presents results of the DWE-USO project. Design and development of the USO and manufacturing of the Huygens Probe units were funded by the Deutsche Agentur für Raumfahrt-angelegenheiten (DARA). Fabrication of the Cassini Orbiter RUSO was funded by ESA.
Many thanks in particular to Klaus Wagner from DSS and to Dirk Plettemeier and Rainer Wohlmuth from the University of Bochum for their tireless efforts and the sleepless nights.

REFERENCES

Ref.1: Bird, M. K. et al., Rubidium Ultra-Stable Oscillators at Titan: The Huygens Doppler Wind Experiment, Proceedings of the Workshop: Scientific Applications of Clocks in Space, 7-8 Nov.1996, Pasadena, CA/USA
Ref.2: Huygens DWE Ultra-Stable Oscillator Experiment User Manual, DWE Project

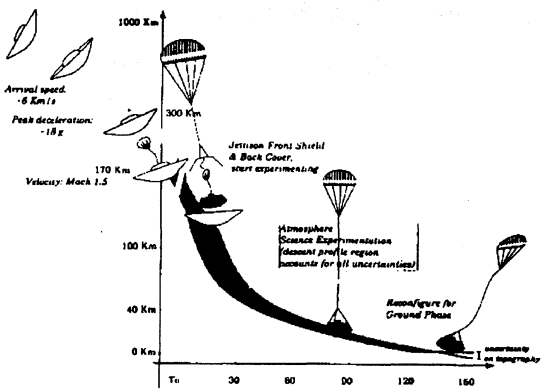


Figure 1: Huygens Probe Descent Profile

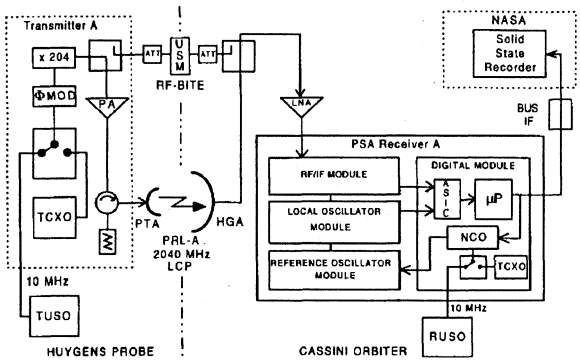


Figure 2: DWE Experimental Configuration

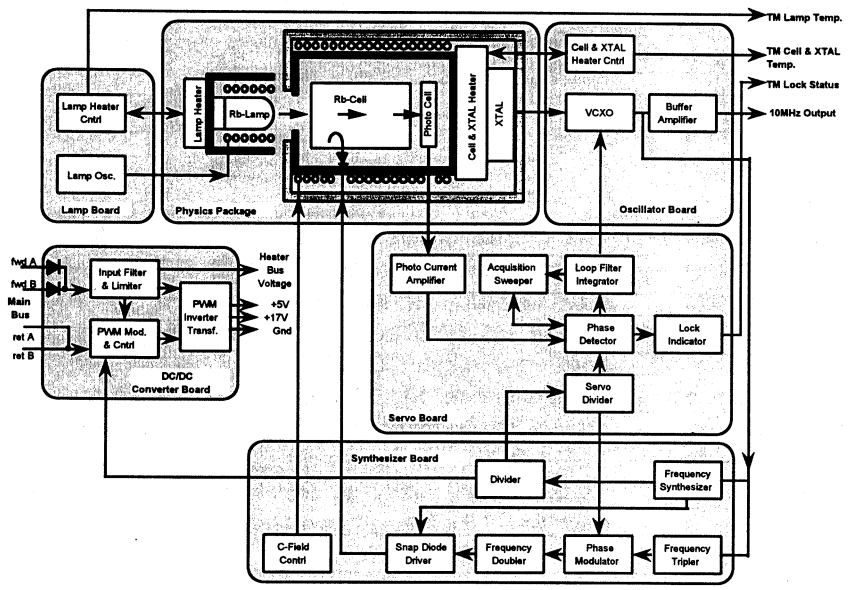


Figure 3: DWE USO Blockdiagram

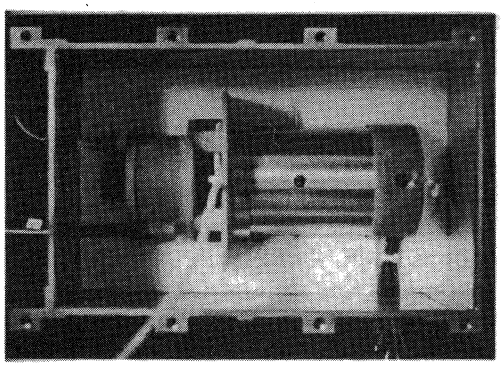


Figure 4: Physics Package and Inner Housing

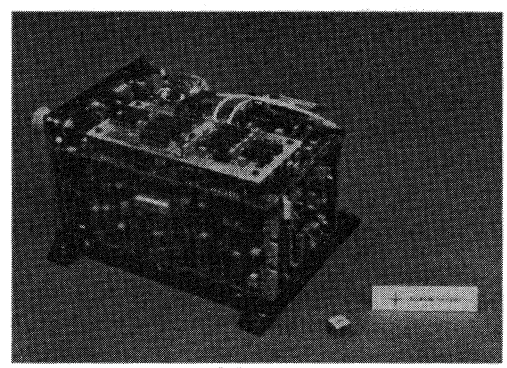


Figure 5: DWE USO Cover and Side Walls Off

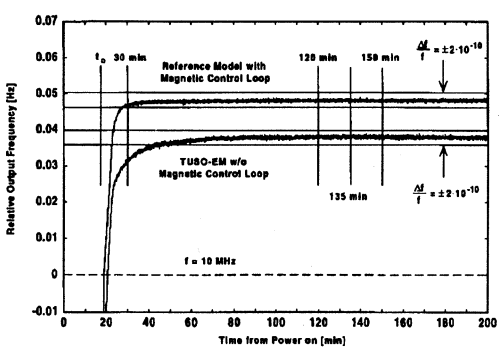


Figure 6: Frequency Drift Stability over Mission

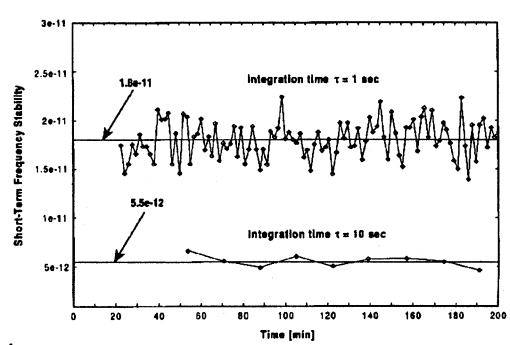


Figure 7: Short Term Frequency Stability

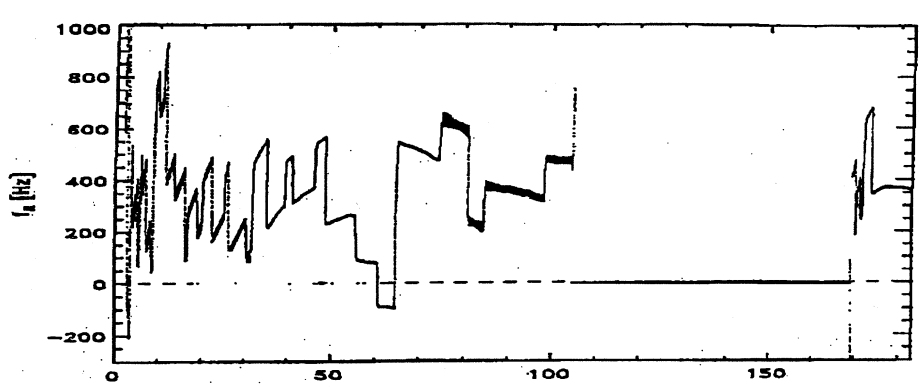


Figure 8: Huygens/Cassini End-toEnd Radio Link Test with Rb USOs and TCXOs

A COMPARATIVE STUDY OF THE VELOCITY DISTRIBUTION DETERMINATION METHODS IN A CESIUM BEAM FREQUENCY STANDARD

A.Makdissi and E. de Clercq
Laboratoire Primaire du Temps et des Fréquences
Bureau National de Métrologie - Observatoire de Paris
61, Avenue de l'Observatoire
75014 Paris - France
E-mail : Alaa.Makdissi@obspm.fr

ABSTRACT

We focus on the use of different approaches to determine the Time Of Flight (TOF) distribution of the atoms in frequency standards in order to reduce the uncertainty on the determination of some frequency shifts that influence the accuracy of a standard. We present a new method to compute the TOF from a transient response of the beam to a frequency step in the interrogation microwave (MW) signal. We propose a comparison criterion to evaluate the accuracy of the estimated TOF distribution and apply it to some existing methods.

Keywords : Time of Flight, Transient response, Atomic Frequency Standard.

1. INTRODUCTION

Using one method to determine the TOF distribution may lead to a biased estimator of the actual distribution. This bias may be due to a systematic error in the method. The constancy of one method to many data sets does not yields necessarily the accuracy of the method, it revels the repeatability of the method rather than its accuracy.
In this paper we propose a new method to deduce the TOF distribution from the transient response of the beam to a frequency step in the MW signal. In the next section we show how to compute this transient response and we compare it with a measured one. Finally, a comparison between some existing method is established.

2. THE TRANSIENT RESPONSE

We suppose that the interrogation angular frequency at time $t < 0$ is ω_1. At time $t = 0$, we apply a frequency step from ω_1 to ω_2. For any time t, we can write :

$$\omega(t) = \omega_1 + \dot{\phi}(t) \qquad (1)$$

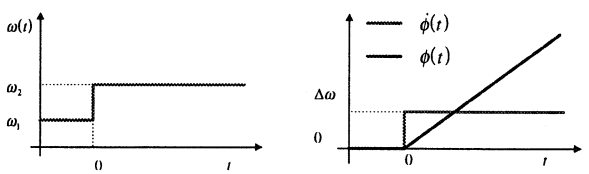
where :

$$\dot{\phi}(t) = \begin{cases} 0 & \text{for } t < 0 \\ \omega_2 - \omega_1 = \Delta\omega & \text{for } t > 0 \end{cases} \qquad (2)$$

$\phi(t)$ is the phase of the interrogation signal and $\dot{\phi}(t)$ is its derivative with respect to time. Integrating (2), the phase in the case of a frequency step is written as :

$$\phi(t) = \begin{cases} 0 & \text{for } t < 0 \\ \Delta\omega t & \text{for } t > 0 \end{cases} \qquad (3)$$

Figure 1.a shows the frequency step and figure 1.b shows the corresponding shape of $\phi(t)$ and $\dot{\phi}(t)$.



(a) (b)

***Fig.1** : (a) frequency step, (b) the corresponding phase.*

In standards which use a Ramsey cavity with drift lenght L and interrogation lenght l , $l << L$, (Figure 2), the signal $y(t)$ measured by a detector located at distance D from the MW cavity is written [1]:

$$y(t) = y_0 + \int_0^\infty f(\tau)\sin^2(b\tau)\cos\left[\omega_1 a_1 \tau + \Phi(t,\tau)\right] \qquad (4)$$

where :

- $a_1 = \frac{L}{\ell}$, $a_2 = \frac{D}{\ell}$, $\alpha_1 = \frac{\ell}{D+L}$ and $\alpha_2 = \frac{\ell}{D}$
- τ is the time of flight in one arm of the cavity
- $f(\tau)$ is the TOF distribution.
- b is the Rabi angular frequency.
- $\Phi(t,\tau)$ is the phase difference seen by an atom of a TOF τ and arriving to the detector at time t.

If we neglect the electrical phase shift between the two arms of the cavity, then $\Phi(t,\tau)$ could be written as :

$$\Phi(t,\tau) = \phi(t - T_D) - \phi(t - T_D - T) \qquad (5)$$

where $T_D = a_2\tau$ is the time spent by the atom from the second arm of the cavity to the detector, $T = a_1\tau$ is the time spent between the two arms (Figure 2).

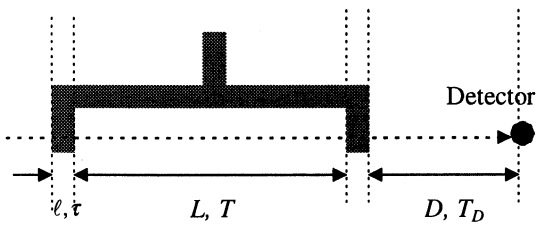


Fig. 2: Ramsey configuration.

Using equation (3), $\Phi(t,\tau)$ could be written for a fixed TOF τ:

$$\begin{cases} t < T_D \Rightarrow \Phi(t,\tau) = 0 \\ T_D < t < T_D + T \Rightarrow \Phi(t,\tau) = \Delta\omega(t - T_D) \\ t > T + T_D \Rightarrow \Phi(t,\tau) = \Delta\omega T \end{cases} \quad (6)$$

In order to calculate the measured signal $y(t)$, we need to write $\Phi(t,\tau)$ versus τ for a fixed time t. Equation (6) leads to:

$$\begin{cases} \tau > \alpha_2 t \Rightarrow \Phi(t,\tau) = 0 \\ \alpha_2 t > \tau > \alpha_1 t \Rightarrow \Phi(t,\tau) = \Delta\omega(t - a_2\tau) \\ \alpha_1 t > \tau > 0 \Rightarrow \Phi(t,\tau) = \Delta\omega a_1 \tau \end{cases} \quad (7)$$

Figure 3.a shows $\Phi(t,\tau)$ versus the time t for a fixed TOF τ and figure 3.b shows the same function versus the TOF τ for a fixed instant t. This last form of $\Phi(t,\tau)$ (Fig 3.b) will be used to compute $y(t)$.

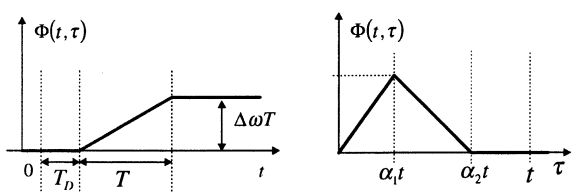


Fig. 3: $\Phi(t,\tau)$ (a) versus the time t, (b) versus the TOF τ.

The integration in (4) can be decomposed in the form :

$$y(t) = \int_0^{\infty} \ldots d\tau = \int_0^{\alpha_1 t} \ldots d\tau + \int_{\alpha_1 t}^{\alpha_2 t} \ldots d\tau + \int_{\alpha_2 t}^{\infty} \ldots d\tau \quad (8)$$

substituting (7) in (8), we obtain :

$$y(t) = y_1(t) + y_2(t) - y_3(t) - y_4(t) + C \quad (9)$$

where :

$$y_1(t) = \int_0^{\alpha_1 t} f(\tau)\sin^2(b\tau)\cos[\omega_2 a_1 \tau]d\tau$$

$$y_2(t) = \int_0^{\alpha_2 t} f(\tau)\sin^2(b\tau)\cos[\omega_1 a_1 \tau - \Delta\omega a_2 \tau + \Delta\omega t]d\tau$$

$$y_3(t) = \int_0^{\alpha_1 t} f(\tau)\sin^2(b\tau)\cos[\omega_1 a_1 \tau - \Delta\omega a_2 \tau + \Delta\omega t]d\tau$$

$$y_4(t) = \int_0^{\alpha_2 t} f(\tau)\sin^2(b\tau)\cos[\omega_1 a_1 \tau]d\tau$$

Figure 4 shows the four terms of Equation (9) in the case of frequency step from -50 Hz to 50 Hz (The Ramsey linewidth of our standard is 100 kHz) for $b = b_{opt}$ (the Rabi frequency that maximise the signal at resonance). They were computed with the functional parameters of our standard and a theoretical weighted by τ maxwellian TOF distribution.

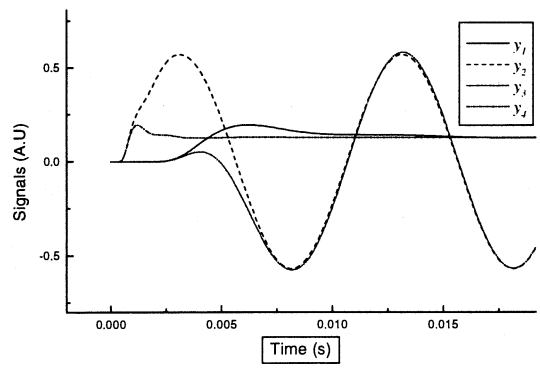


Fig. 4 : The four terms forming the transient response.

In our standard at LPTF, we measured this transient response for many different frequency steps. The measurement were made using an acquisition card which samples the signal with a period of 60 μs. For each configuration (frequency step) we averaged over 100 measurement to have a good Signal to Noise ratio. The experimental results fitted well with the theoretical ones. Figure 5 shows a measured transient response compared to the computed one in the same conditions as before. (In our set-up we have L = 1000 mm, l = 10 mm and D = 250 mm).

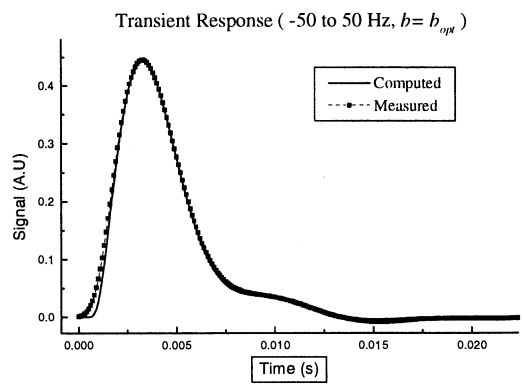


Fig. 5 : Transient response of the LPTF standard.

3. THE TOF COMPUTATION

Equation (9) is the sum of four terms. Each term is a Volterra integral so $y(t)$ could be seen as a Volterra equation with the general form :

$$y(t) = \int_0^t K(\tau, \omega_1, \omega_2, b) f(\tau) d\tau \qquad (10)$$

where $K(...)$ is the known kernel of the equation (see equation (9)), $y(t)$ is the known (measured) left-hand side and $f(\tau)$ is the unknown solution.

The discretization of (10) in t and τ leads to linear system of the form:

$$Y = AX \qquad (11)$$

Where :

- Y is a vector of N measured data (discretization of y(t))
- X is a vector of M (M < N) points of the unknown solution (discretization of $f(\tau)$).
- A is a matrix (NxM) representing the discretization of the kernel $K(...)$.

The linear system (11) is generally ill-conditionned. In order to determine its solution, here we applied the regularization technique already applied to compute the TOF from one Ramsey pattern [2]. Figure 6 shows the TOF distribution deduced of the data of figure 5. We used N = 300 data points spaced by Δt = 100 μs, and computed the TOF distribution in M= 250 points spaced by $\Delta\tau = \alpha_1 \Delta t = \Delta t / 125$ = 0.8 μs. In the same figure we showed the theoretical weighted maxwellian distribution for comparison.

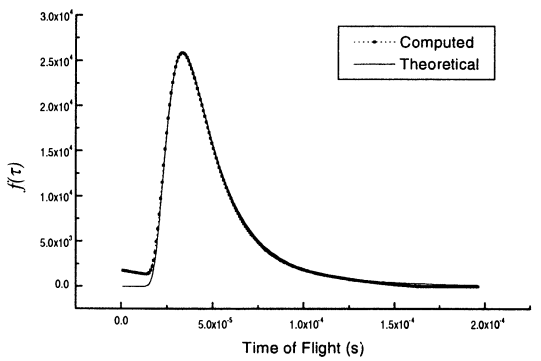


Fig. 6 : TOF distribution deduced from the transient response of figure 5.

4. COMPARISON

Many methods were proposed to calculate the TOF distribution. When applied to calculate the Doppler effect shift using many data sets, each of these methods showed a repeatability of some 10^{-15} when the data sets were changed. But the difference between two methods was sometimes more than 10^{-14}. This led us to think that there was a systematic error or bias introduced by each method. This bias may be due to an approximate model in some methods or to the effect of noise in all the data processed by these methods. In order to adopt a method for our standard we need a criterion to compare these methods.

We have shown before [3] that the frequency shift due to the second order Doppler effect can be computed directly from one measured Ramsey pattern without knowing the TOF distribution $f(\tau)$ or the Rabi frequency b with a very good approximation. So we choose this method (we call it the *Direct method*) to compute the Doppler effect as a reference, and we will compare the computed Doppler shift by the TOF estimated by some proposed methods with the value obtained by the Direct method.

The selected methods are:

1) **REGULARIZATION :** Data in frequency domain, No approximation, One Ramsey Curve, [2].
2) **SHIRLEY :** Data in frequency domain, Approximation, Three Ramsey Curves, [4].
3) **BOULANGER :** Data in MW level domain, No approximation, One $I(b)$ + One Ramsey curve, [5].
4) **HIGH MW :** Data in frequency domain, Approximation, A part of one Ramsey curve at very high MW level, [3].
5) **TRANSIENT :** Data in time domain, Approximation, one transient response to a frequency step, this paper.

Figure 7 shows the value of the relative Doppler shift computed by the precedent methods. The straight line (Direct method) shows the value computed directly from an experimental Ramsey pattern without the use of a TOF distribution.

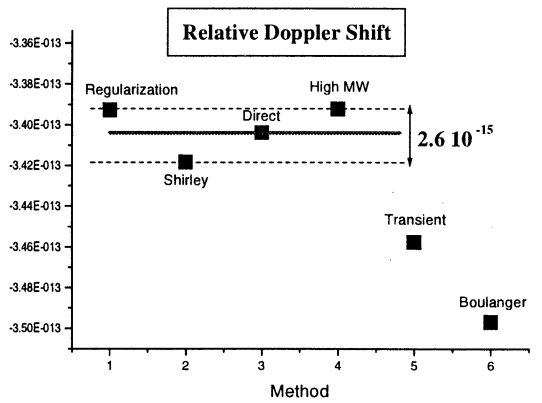


Fig .7 : The Relative 2nd order Doppler shift computed by the TOF distribution estimated by some existing methods.

The deviation of the Boulanger method from the others may be due to a miss of accuracy in the measurement of the MW level or to a the influence of the neighbouring transitions level.

The deviation of the transient method may be due to a time jitter when repeating the same frequency step in order to make an averaged measurement or to the detector or syntheziser response which were neglected.

From figure 7, we can see that the second order Doppler effect is estimated with an accuracy of 2.6×10^{-15} at worst using four different methods.

5. CONCLUSION

We have shown how to compute the transient response of the beam to frequency step in the interrogation signal. The experimental transient response fitted well with computed one. We used this transient response to deduce the TOF distribution in our standard. The computation of the TOF distribution is based on the solution of integral Volterra equation by the regularization method. Finally, we compared some existing methods of evaluation of the TOF distribution. The comparison is based on the computation of the Doppler effect. We showed that using four different methods we can estimate the Doppler shift with an uncertainty less than 2.6×10^{-15} and consequently reducing the uncertainty on the evaluation of some frequency shifts.

REFERENCES

[1] J. Vanier and C. Audoin, The Quantum Physics of Atomic Frequency Standards, Bristol: Adam Hilger, Ch. 5, 1989.

[2] A. Makdissi and E. de Clercq, to be published in IEEE Trans. On UFFC, May 1997.

[3] A. Makdissi and E. de Clercq, to be published in IEEE Trans. On Instr. & Meas., April 1997.

[4] J. H. Shirley, in Proc. 34rd. Sympos. On Frequency Control, p. 162, 1989.

[5] J. S. Boulanger, Metrologia, Vol. 23, pp. 37 - 44, 1986.

CESIUM MAGNETO-OPTIC TRAP IN ROMANIA

Cipriana Mandache , M.Bengulescu , T.Acsente

National Institute for Laser, Plasma and Radiation Physics, Nuclear Fusion and Plasma Physics Laboratory, R-76900 , Bucharest, Magurele, P.O.Box MG-36, Romania

ABSTRACT

The goal of our work in INFLPR - Romania is to develop a cesium magneto-optical trap and finally to realize an atomic clock using cold Cs atoms. We investigated how the size of a ring-shaped cloud of cold, magneto-optically trapped cesium atoms depends on the trapping parameters. We have studied numerically the capture process in a magneto-optical trap. When the optical Bloch equations are combined with the dynamical equations that describe the motion of the center of mass coordinate of the atom in the semiclassical approximation, a set of nonlinear ordinary differential equations in 406 real variables is obtained. The results show surprising features, as an atom being launched with a velocity towards the center of the trap deviates from the trajectory which passes through the center. This could lead to a decrease of the capture rate for slow atoms.

Keywords :

cesium clock, cold atoms, magneto-optical trap

1. INTRODUCTION

The experimental device developed in INFLPR for cooling and stopping the cesium atoms is quite simple : a main fused silica cylinder (50 mm diameter) with four orthogonal arms (18 mm diameter), and, at each end, optical windows (fig.1).

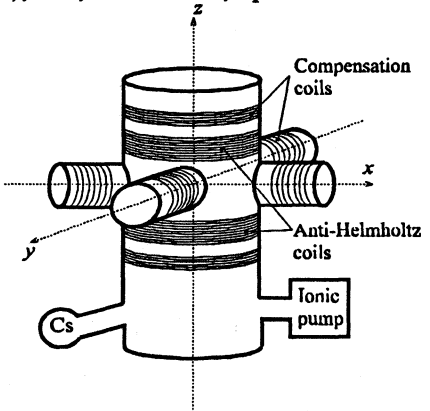


Fig.1

The cesium reservoir is attached to the glass envelope in a "cold finger" configuration . All the trap cell (including the"cold finger" of cesium) is attached to an ionic pump. Two anti-Helmholtz coils, disposed on the main cylinder are able to produce a magnetic gradient field between 5 and 15 G/cm, while three pairs of Helmholtz coils are used to compensate the earth magnetic field. To obtain a magnetic field gradient of 10 G/cm, it is enough to inject a current of 375 mA in the two anti-Helmholtz coils. The trapping beams come from an SDL 5402 laser diode in an extended cavity configuration (cavity length 4 cm) (Ref.1). Various components and special devices necessary to the test of these components have been realized in our laboratory, as follows : - the laser diode driver with a noise 0.6 $nA/\sqrt{Hz}$ (between 100 Hz and 100 kHz) ; -a modulation laser device ; - a current source for the Helmholtz coils; - a laser diode characterization device .

The acquisition system coupled to the computer is able to analyze the experimental data.

2. SIZE OF CLOUDS OF COLD CESIUM ATOMS

If we have N atoms in a stable ring-shaped spatial distribution of radius r and the trapping laser beams misaligned, the atom is subjected to a motion given by the equation (Ref. 2):

$$m\frac{d^2r}{dt} = -Kr - \gamma\frac{dr}{dt} + \frac{\alpha N}{x^2}r + \zeta(r)kxr \qquad (2.1)$$

where the spring constant K, is:

$$K = \frac{16\hbar k\Gamma\Delta\Omega_0^2\left(\frac{d\omega}{dx}\right)}{\left(4\Delta^2 + \Gamma^2 + 12\Omega_0^2\right)}$$

and Ω_0 is the Rabi frequency; $d\omega/dx$ is the change of the atomic frequency due to the magnetic field; $\gamma = kK / (d\omega/dx)$ is the damping constant due to the Doppler shift.

The third term arises from the strong range coupling between the atoms . The last term is the macroscopic vortex force F_v, which arises because of the imbalance between the two spontaneous forces acting in opposite directions in the case of misaligned Gaussian laser beams. With this assumption we obtained the radius ring for of stable ring trajectory :

$$r = \left[\frac{\alpha N'}{K - m\omega^2(r)}\right]^{1/3} = \left[\frac{\alpha N'}{K^2 - \left(\frac{m\zeta_{s0}^2}{\gamma^2}\right)\exp^2\left(-\frac{r^2+2s^2}{w^2}\right)}\right]^{1/3}$$

where (Ref.2) :

$$\zeta(r) = \zeta_0 \exp\left(\frac{-2\left(r^2+2s^2\right)}{w^2}\right)$$

$$\zeta_0 = \frac{8\hbar k\Gamma\Delta\Omega_0^2\left(\frac{s}{w}\right)^2}{\left(4\Delta^2+\Gamma^2+12\Omega_0^2\right)}$$

We investigated how the size of a ring-shaped clouds of cold, magneto-optically trapped cesium atoms depends on the trapping parameters : magnetic field gradient (dB/dz), laser intensity, detuning (Δ) and misalignment. We investigated the size of the radius ring-shaped clouds for the following values : dB/dz between 5 and 20 G/cm, Δ between -1 and -165 MHz and the misalignment between 0,1 and 2 mm.

When the number of trapped atoms is small and also for the small detunings the radius of atomic ring depends strongly on the misalignment. For N $> 10^8$ the radius depends strongly on the number of trapped atoms (fig 2 and fig.3).

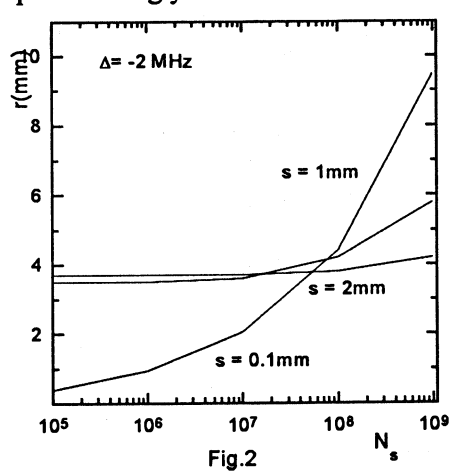


Fig.2

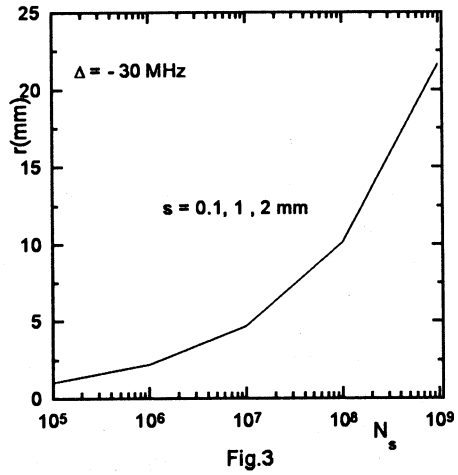


Fig.3

For the detuning $\Delta > -20$ MHz, the radius does not depend on the misalignment. For an misalignment of 1 mm, the maximum of ring radius is obtained for a detuning Δ= -1,65 MHz. We see also that the ring radius depends on the number of trapped atoms (fig.4)

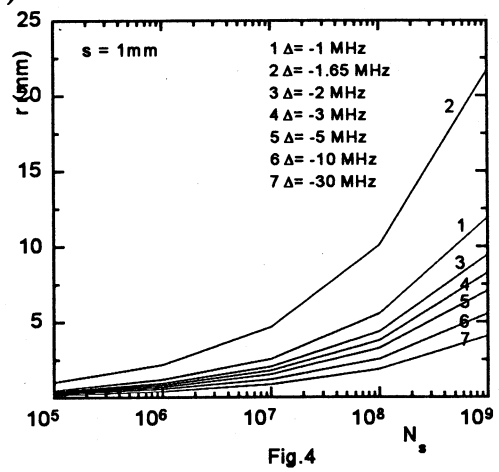


Fig.4

We have found also that the ring radius decreases with the magnetic field gradient. This is expected because an increase
in dB/dx produces a trap with a stronger spring constant.. For a magnetic field gradient of 10G/cm the number of trapped atoms is maximum.

3. 3-D MOTION OF CESIUM ATOMS

We have performed numerical computations taking into consideration the actual sublevel structure of the F=4→ F=5 transition in Cesium, the gradient of magnetic field and the exact structure of the laser electric field (the phase difference of the three pairs

hyperfine splitting. The internal state of the atom is described by the density matrix $\sigma(t)$ in the rotating frame. This density matrix has as dimension the sum of the number of sub-states of the lower level and the excited level. Here we chose as the basis states the sub-states M_F = -4 to 4 of the F = 4 ground state and the M_F = -5 to 5 sub-states of the F = 5 excited state with respect to the Oz axis. Other states do not enter our problem because the excited state (F=5) decays only to F = 4 and even if collisions or side wing pumping by the trap lasers succeeded in bringing the atom in the F = 3 ground state, the repumping laser would quickly return the atom to the F = 4 → F = 5 cycle. If we separate this matrix into four blocks corresponding to the two levels, we obtain the submatrices σ_{ee} and σ_{gg} for the excited-state and lower-state manifolds, and the submatrices σ_{eg} and σ_{ge} containing the optical coherences. Likewise, we separate the electric-dipole operator of the atom as :

$$\mu = \mu_{eg} + \mu_{ge} \qquad (3.3)$$

into its raising and lowering part.
In the rotating-wave approximation, the atom-field coupling is governed by the Rabi operator :

$$\Re = \mu_{eg} \bullet E_+ / \hbar \qquad (3.4)$$

which generalizes the Rabi frequency. Obviously, this operator $\Re$ has nonvanishing matrix elements only between a lower and an upper substate.
The evolution equations for the submatrices of σ may be put in the form (Ref.7,8) :

$$\frac{d\sigma_{ee}}{dt} = -\Gamma\sigma_{ee} - iL\sigma_{ee} + i\Re\sigma_{ge} - i\sigma_{ge}\Re^+$$

$$\frac{d\sigma_{gg}}{dt} = \Gamma\sum_\beta Q_\beta^+\sigma_{ee}Q_\beta - iL\sigma_{gg} + i\Re^+\sigma_{eg} - i\sigma_{ge}\Re$$

$$\frac{d\sigma_{eg}}{dt} = -\left[\frac{\Gamma}{2} - i\Delta + iL\right]\sigma_{eg} + i\Re\sigma_{gg} - i\sigma_{ee}\Re$$

$$\frac{d\sigma_{ge}}{dt} = -\left[\frac{\Gamma}{2} + i\Delta + iL\right]\sigma_{ge} + i\Re^+\sigma_{ee} - i\sigma_{gg}\Re \qquad (3.5)$$

The spontaneous decay rate is indicated by Γ , and $\Delta = \omega - \omega_0$ (where ω_0 is the average atomic frequency) is the detuning of the light frequency.

The Liouville operator L accounts for the level splitting within the two manifolds *e* and *g* and is basically the commutator with the Zeeman (or hyperfine) Hamiltonian. In the case of weak coupling of the magnetic field, which, in the case of Cs ($g_{F=5}$ = 0.550 , $g_{F=4}$ = 0.251), goes up to about 10 Gauss due to the hyperfine structure intervals of about 100 MHz of the excited state, we can write

$$-iL = -\frac{i}{\hbar}[\mu_M \bullet B,..] \cong \frac{i\mu_B g_F}{\hbar}[F \bullet B,..] \qquad (3.6)$$

within the ground- and excited state manifolds. Finally, we have introduced the dimensionless dipole operators Q_β to account for the anisotropic repopulation of the lower level by spontaneous decay. These operators are defined by their matrix elements as Clebsch-Gordan coefficients (Ref.7,9) :

$$\langle F_e M_e | Q_\beta | F_g M_g \rangle = \langle F_e M_e | F_g M_g ; 1\beta \rangle \qquad (3.7)$$

for β = -1, 0, 1,where F_e and F_g equal 5 and 4. Contrasting with (Ref.7,10), we do not adiabatically eliminate the optical coherences (since kv > Γ), nor perform the low intensity approximation (since the Rabi frequency is of the order of Γ). The electric fields $E_\pm$ inherent in the operators $\Re$ and $\Re^+$ should be taken at the time-dependent position **R**(t) of the atom. Since, in the spirit of the semiclassical picture, we assume that the atom moves with the given velocity **v**, this position must be taken as :

$$\mathbf{R}(t) = \mathbf{R}_0 + \mathbf{v}\,t \qquad (3.8)$$

For an atom in a radiation field, the Heissenberg operator for the force **f** is determined by the commutator of the momentum operator with the Hamiltonian. This gives :

$$\mathbf{f} = \nabla(\mu E) = \hbar(\nabla\Re) \qquad (3.9)$$

For the component j of the force F this gives :

$$F_j = 2\,\mathrm{Re}\,\mu_+ \cdot \nabla_j E_- \qquad (3.10)$$

where μ_+ is the positive frequency part of the mean dipole :

$$\langle \mu(t) \rangle = \mu_+ e^{-i\omega t} + \mu_- e^{i\omega t} \qquad (3.11)$$

given by :

of σ_+ - σ_- beams being equal to 0, in the sense that at x = y = z = 0 the z - pair produces a linearly polarized electric field along $\hat{e}_y$ with a phase of $-\pi$ / 2, the x-pair a field along $\hat{e}_z$ with the same phase and the y - pair a field along $\hat{e}_x$ again with the same phase). In the upper range of velocities that are important for the capture process (1-10 m/s) the sub-Doppler cooling mechanisms like polarization gradient cooling can be ignored because the atoms traverse the gap between two successive identical field polarizations in a time much shorter than the ground-state sublevel pumping time. Anyway, in our model they are taken into account as we have used the Optical Bloch Equations (OBE) which contain the effects of the light-shift and population pumping (Ref.3). Also, we have tested the usual hypothesis which is made in the usual one-dimensional model of the capture process (Ref.4,5), which is that the different M_F sublevels are equally populated, and we can state that even though that the relatively strong magnetic field at the outskirts of the trap mixes the various sublevels, in the mean and at each instant there are substantial population differences between these levels. In the end, even considering all the possible simplifications (like the hermitian nature of the density matrix), a system of ordinary first-order nonlinear differential equations in 406 (real) variables has to be solved, with a characteristic time of evolution of about 10 ns (due to the value of the Larmor frequency and the strength of the laser-atom coupling on one side, and the dependence of the laser field with position on the other side). Essentially, we have used a semiclassical description of the interaction between the translational degrees of freedom of the atom and the laser field excited internal state variables. This is fully justified for the capture process, as both the velocity is greater than the recoil velocity for cesium (3.49 mm/s) and the recoil energy is much smaller than ($\hbar$ times) the spontaneous decay rate Γ. Moreover, in this initial attempt, we have not incorporated the velocity diffusion process due to mainly two reasons : firstly, in contradiction to one-dimensional models, to account for diffusion either the motion of the atom has to be known beforehand or one deals with a velocity distribution which changes little over a recoil velocity (this would apply to the final stages of trapping) and secondly one does not need to obtain a Fokker-Planck type equation in order to obtain the final velocity distribution (and temperature) but rather is interested if the atom is captured or not. We can (and shall) add the diffusion effect by superposing a small random velocity kick to the velocity as obtained from our deterministic equations (keeping track of the various angular distributions involved). We are aware of (Ref.6), where a Monte Carlo wave-function approach was applied in the case of the 3D optical molasses with a discretization step of the atomic momentum of one recoil momentum and a range of the order of a few tens of atomic recoil momenta. Evidently, a full quantum-mechanical treatment would be desirable, but the amount of computation required, even with the Monte Carlo treatment (which instead of the solution of N^2 OBE's requires the (stochastic) evolution of N variables, averaged a number of times to obtain good enough statistics, where N is the number of states (in our case 20 times the number of points on the momentum grid, which, for an initial velocity of 10 cm/s, is about 188,000). Moreover, as we suppose that the evolution of the dynamical variables is strongly dependendent of the initial condition, the number of averages needed to obtain convergence can be large.

3.1. Equations of motion

We consider the situation when a Cesium atom is entering the region of interaction with the laser beams (of frequency ω) of a MOT with a velocity v. The electric field of the lasers (we work in the dipole approximation of the interaction Hamiltonian which is excellent for allowed transitions as the cesium $6S_{1/2} \mid F=4 \rightarrow 6P_{3/2} \mid F=5$ transition at 852 nm) can be decomposed as :

$$E(R,t) = E_+(R)e^{-i\omega t} + E_-(R)e^{i\omega t} \qquad (3.1)$$

In the case of the MOT :

$$E_+(R) = E_0 \exp\left(-\frac{y^2+z^2}{2w^2}\right)\left(\frac{-1}{\sqrt{2}}\left(\hat{e}_y + i\hat{e}_z\right)e^{ikx} + \frac{1}{\sqrt{2}}\left(\hat{e}_y - i\hat{e}_z\right)e^{-ikx}\right) + (x \rightarrow y \rightarrow z) \qquad (3.2)$$

The field drives the transition between a lower level indicated by the index g (for ground state) and an excited level indicated by e. Both levels may consist of a manifold of degenerate or nearly degenerate sub-states, corresponding to Zeeman degeneracy or

$$\mu_+ = \mathrm{Tr}\ \sigma_{eg}\ \mu_{ge} \qquad (3.12)$$

In order to solve together the Bloch equations (3.5) and the kinematical equation (3.8), all that is now needed, besides Newton's second law, is to write down the expression of the cuadrupolar magnetic field, which is :

$$\mathbf{B(R)} = dB/dz\ |\ (\ z\,\hat{e}_z\ -\ (x\,\hat{e}_z\ +\ y\,\hat{e}_y\ )/2) \qquad (3.13)$$

3.2 Results

Due to the fact that the density matrix must always be hermitic, we can choose as variables in our system of ODE's the three coordinates x, y, z and the three components of the velocity of the atom together with the superior diagonal components of the density matrix.

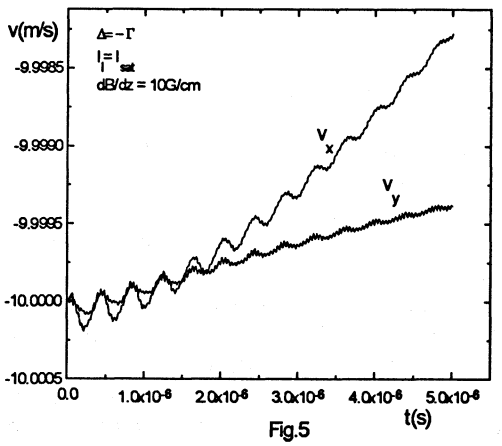


Fig.5

As the MATLAB routine for solving ODE's we are using deals only with real numbers, and the coherences really are complex, we are left with 406 variables. To speed up the solution, we have vectorized the equations and also used some tricks in order to insure that the populations always remain positive and the trace of the density matrix stays equal to one. The solutions (Figs. 5 and 6) show, on the small scale, oscillations of the velocity due to the dipole force (with a period of about 0.4 μm) and another, more pronounced oscillation with a temporal period of 0.4 μs, which we, up to now, failed to interpret. It may be due to oscillations in the density matrix as it relaxes from the initial condition. Anyway, it is far from the Larmor frequency. On the large scale, things are even more dramatic, in that the uniform deceleration we expected to observe is replaced by an interesting dynamics, with different decelerations on the three axes and periods of acceleration after a long period of deceleration.

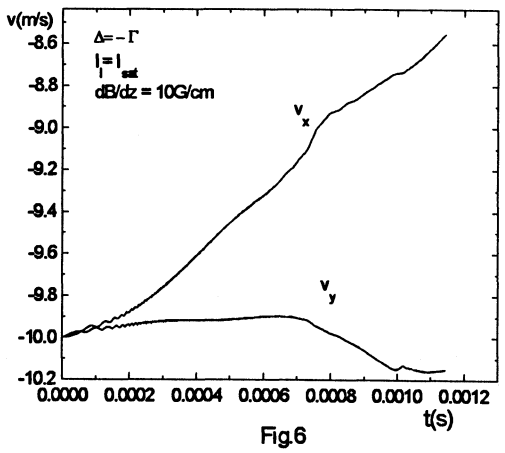


Fig.6

This, for smaller velocities, could lead to the fact that even at velocities that traditionally are supposed to lead to capture, the atom may escape from the trap. To check the consistency, we repeated the calculations with different tolerances (and thus time steps) and found excellent agreement.

REFERENCES

1. C.Mandache, P.Thomann , "Miniaturised, extended cavity for laser diode", European Space Agency Reports, CEC CIPA 3510PL (1993)

2 I.Guedes, M.T.de Araujo, D.M.B.P. Milori, G.I.Surdutovich, V.S.Bagnato and S.C.Zilio," Forces acting on magneto-optically trapped atoms", J.Opt.Soc.Am B11, 1935 (1994)

3. C.Cohen-Tannoudji, Frontiers in Laser Spectroscopy edited by R.Balian, S.Haroche and H.Liberman (North-Holland,Amsterdam, p.1(1977)

4. K.Lindquist ,M.Stephens and C.Wieman, Phys.Rev.A ,Vol 46, Nr. 7, p.462 (1991)

5. K.E.Gibble, S.Kasapi and S.Chu, Optics Lett., Vol.17,Nr.7, p.526 (1992)

6. Y.Castin and K.Mølmer, Phys.Rev.Lett. A Vol 74, Nr.19, p.3772 (1995)

7. G.Nienhuis and P.van der Straten, Phys.Rev.A, Vol.4, Nr.1, p .462 (1991)

8. G.Nienhuis, Phys.Rep. 138, 151, (1986)

9. J.Javanainen, Phys.Rev.A, Vol.44,Nr.1, p.5857 (1991)

10.A.Hemmerich, M.Weidemüller, T.Esslinger and T.W.Hänsch, Europhys. Lett., 21 (4), pp. 445-450 (1993)

OPTIMIZED OPERATION OF THE PASSIVE HYDROGEN MASER

Claudiu MIRICĂ, Liviu GIURGIU, Octav GHEORGHIU

Atomic Clock Team, IFTAR Bucureşti-Măgurele
PO Box MG 7, Bucureşti-Măgurele, Romania

ABSTRACT

A theoretical method for the passive Hydrogen maser optimization of the operating conditions: interrogation signal level and atomic beam intensity, and of the designing parameters: the input and output coupling factors, is introduced. The method do not depend on the modulation or the locking technique (Ref. 1). Some new notations and quantities of interest are introduced, the results calculated numerically are plotted, and the relations required to determine other quantities of practical interest are provided. A normal size maser and a reduced size one are calculated as example.

Keywords: passive Hydrogen maser
optimization

1. A BRIEF THEORY OF THE PASSIVE HYDROGEN MASER

1.1 The Hydrogen Maser is a primary time-frequency standard with an outstanding frequency stability. As physical principle, this maser uses the atomic Hydrogen resonance at the hyperfine splitting frequency (ν_0=1.42 GHz) of the Hydrogen atom fundamental level. This transition is used due to the high quality factor of this atomic line (10^9). The atomic resonance is detected by stimulated emission of radiation in a microwave resonant cavity (of Q_0 quality factor and V_c volume). Within the cavity, the H atoms are confined in a constant phase region (of V_b volume, η' filling factor and T_b storage time) (Ref. 2: 6.2.120). In the storage region a magnetically selected atomic beam (of total intensity I_t, having an r proportion of atoms in the useful level $|F=1, m_F=0>$, representing a net intensity $I = rI_t$, the rest being in the level $|F=1, m_F=1>$) is collimated. The stored atoms get relaxed through non-radiant processes with the T_1 longitudinal and T_2 transversal relaxation times. In the passively operated maser, the resonant cavity is coupled with an input circuit with a coupling factor β_1, and an output circuit) with a coupling factor β_2. The input circuit contains a V_g interrogation voltage generator with a Z_0 matched internal resistance. The output circuit has as load resistance the matched Z_0 input resistance of the amplifier, on which is collected the V_{out} voltage. The total coupling factor becomes $\beta=\beta_1+\beta_2$, that reduces the cavity's quality factor to $Q_c=Q_0/(1+\beta)$ (Ref. 2: 6.3.26).

1.2 Considering that the interrogation voltage has the pulsation $\omega=2\pi\nu$, with a $\delta\omega=\omega_0-\omega$ mistuning correction, and that the cavity resonance pulsation of the is ω_c, with a $\delta\omega_c=\omega_0-\omega_c$ mistuning correction, than the relation between V_g and V_{out} is:

$$V_{out}(\delta\omega)=\frac{\sqrt{\beta_1\cdot\beta_2}}{1+\beta+\beta_m(\delta\omega)+j\cdot\frac{\delta\omega_c-\delta\omega}{\omega}\cdot Q_0}\cdot V_g \quad (1)$$

where β_m is the coupling factor with the atomic medium, and it is a negative quantity due to the signal amplification by stimulated emission.

1.3 For the quantitative description of β_m one should define the quantities related to the effects of spin exchange relaxation and transition saturation.

Since the threshold intensity that has been introduced for the active H maser has no more physical meaning for the passive maser (Ref. 2: 6.3.33 & 37), another reference and normalized atomic beam intensity are introduced to simplify the definitions:

$$I_{t1}=\frac{2}{T_t}\cdot\frac{1}{\bar{\sigma}_{ex}\cdot\bar{v}_r}\cdot\frac{V_b}{T_b}=\frac{I_{th}}{q}\quad;\quad\tilde{I}=\frac{I_t}{I_{t1}}=q\cdot\frac{I_t}{I_{th}} \quad (2)$$

I_{t1} has the physical meaning of the total atomic beam intensity at which the spin exchange longitudinal relaxation time is $T_t=\sqrt{(T_1)_0\cdot(T_2)_0}$ (the geometrical mean of the longitudinal and transversal relaxation times without spin exchange effect); $\bar{\sigma}_{ex}$ is the spin exchange mean cross section; $\bar{v}_r$ is the mean relative velocity between H atoms at the cavity temperature Θ_c=313K, ($\bar{\sigma}_{ex}\cong 23.3\cdot10^{-20}\cdot m^2$, $\bar{v}_r$ = 3626 m/s).

In respect to the above mentioned definition, the relaxation times with spin exchange effect can be described in the easiest way:

$$T_i^{-1}=(T_i)_0^{-1}+i\cdot\tilde{I}\cdot T_t^{-1},\ \text{for } i=1...2 \quad (3)$$

(Ref. 2: 6.3.29 & 30)

Another quantity that should be introduced is the specific quality factor of the Hydrogen maser,

introduced for the unloaded cavity, with a "0" subscript:

$$q_0 = \frac{\overline{\sigma}_{ex} \cdot \overline{\nu}_r \cdot \hbar}{2 \cdot \mu_0 \cdot \mu_B^{\,2}} \cdot \frac{T_b}{T_t} \cdot \frac{1}{\eta' \cdot Q_0 \cdot r} = \frac{q}{1+\beta} \qquad (4)$$

A further simplifying notation is introduced (Ref. 2: 6.3.39):

$$u = \sqrt{\frac{(T_1)_0}{(T_2)_0}} = \sqrt{\frac{3}{4}} \ldots 1 \; ; \; c = 2 \cdot u + u^{-1} = 2.89 \ldots 3 \qquad (5)$$

The coupling factor with the atomic active medium becomes:

$$\beta_m = \beta_{m0} \cdot \frac{1 + j \cdot \delta\omega \cdot T_2}{1 + \delta\omega^2 \cdot T_2^{\,2} + S_0} = \frac{\beta_{m0}}{1 + \frac{S_0}{1 + \delta\omega^2 \cdot T_2^{\,2}}} \cdot \frac{1 + j \cdot \delta\omega \cdot T_2}{1 + \delta\omega^2 \cdot T_2^{\,2}} \qquad (6)$$

The unsaturated coupling factor at resonance is defined by:

$$\beta_{m0} = -\frac{\tilde{I}}{q_0 \cdot \left(1 + c \cdot \tilde{I} + 2 \cdot \tilde{I}^2\right)} \; ; \; \alpha = -\frac{\beta_{m0}}{1+\beta} \qquad (7)$$

emphasizing the relation with the previously introduced oscillation parameter α (Ref. 2: 6.3.38) that describes the maser's capacity to enter in oscillation for $\alpha>1$.
The saturation factor at resonance is:

$$S_0 = \frac{V_{out}^{\,2}}{2 \cdot \beta_2 \cdot Z_0 \cdot P_1 \cdot q_0 \cdot \left(1 + c \cdot \tilde{I} + 2 \cdot \tilde{I}^2\right)} \qquad (8)$$

$$P_1 = \frac{\hbar \cdot \omega_0}{2} \cdot I_{t1} \qquad (9)$$

The power P_1 has the physical and direct meaning of the potential power brought into the cavity by the magnetically selected atomic beam with the intensity I_{t1}, and it's use simplifies all the relations.
1.3.8 The voltage gain produced by the active medium is:

$$G = \left(1 + \frac{\beta_m}{1+\beta}\right)^{-1} \; ; \; G_0 = \left(1 - \frac{\alpha}{1+S_0}\right)^{-1} \qquad (10)$$

with G_0 being noted the gain at resonance (Ref. 2: 6.5.16).

2. THE OPTIMIZATION THEORY

2.1 In any of the modulation and locking techniques that one might use, the physical quantity whose variation is detected in the passive Hydrogen maser is the output voltage V_{out}. Considering this reason, the meaningful quantity for the stability evaluation is the dispersion value at resonance:

$$d = \left|\frac{\partial V_{out}}{\partial \omega}\right|_{\delta\omega=0} = V_{out}|_{\delta\omega=0} \cdot T_2 \cdot \frac{\alpha}{1 + S_0 - \alpha \cdot \frac{1-S_0}{1+S_0}} \qquad (10)$$

2.2 Added to the real mistuning voltage, the coupled cavity produces also a noise voltage that has the following value at resonance:

$$V_n = \sqrt{2 \cdot \beta_2 \cdot Z_0} \cdot \sqrt{\frac{4 \cdot k_B \cdot \Theta_c \cdot f}{1 + \beta + \beta_m|_{\delta\omega=0}}} \qquad (11)$$

where f is the Fourier frequency band of the filtered noise. It should be mentioned that this is only the thermal white frequency noise of the loaded cavity and it does not contain any other noise source (e.g. intermodulation noise induced by the modulation technique (Ref. 3), electronic components flicker noise etc.) that have to be evaluated separately. The white frequency noise approximation is valid for medium time intervals. As for the white frequency noise produced by the phase locking loop, it is to be considered through the amplifier's noise figure, that can be measured without any special problem. This approach will be used bellow.
2.3 The stability of the VCXO locking, in the assumption that the control loop operates without error, is:

$$\sigma_y(f) = \frac{|\Delta\omega|}{\omega_0} = \frac{V_n(f)}{\omega_0 \cdot d} \qquad (12)$$

due to the presence of the noise voltage V_n that induces a false error signal equal to $\omega_0 d$ in the control loop, shifting the frequency by $\Delta\omega$. In the time domain, for white noise, the Fourier frequency f transforms to $(2\tau)^{-1}$, τ being the integration time in a stability measurement.
The stability's expression becomes:

$$\sigma_y(\tau) = \frac{1}{\omega_0 \cdot T_t} \cdot \sqrt{\frac{2 \cdot k_B \cdot \Theta_c}{P_1}} \cdot \frac{u + 2 \cdot \tilde{I}}{\sqrt{q_0 \cdot \left(1 + c \cdot \tilde{I} + 2 \cdot \tilde{I}^2\right)}} \cdot X(\alpha, S_0) \cdot \tau^{-1/2} \qquad (13)$$

where the quantity $X(\alpha,S_0)$ is the adimensional function:

$$X(\alpha, S_0) = \frac{1 + S_0 - \alpha \cdot \frac{1-S_0}{1+S_0}}{\alpha \cdot \sqrt{S_0 \cdot \left(1 - \frac{\alpha}{1+S_0}\right)}} \qquad (14)$$

This result is similar with the relation 6.6.66 of Ref. 2, for the power spectral density of the fractional amplitude fluctuations, since it accounts for the envelope detection and amplitude integration of the mistuning voltage.
This quantity is the stability indicator of the interrogation signal level, through the saturation effect, when all the other parameters are kept

constant. The first level of optimization is to minimize the X value by finding the optimum saturation for a given oscillation parameter α.

2.4 The study of X, as a function of S_0 having parameter α, shows the minimum of this quantity. This minimum is given in two disjunct domains of α values, by two roots of the equation $\partial X/\partial S_0 = 0$, providing the optimum relation between S_0 and α:

$$S_0(\alpha) = \begin{cases} \sqrt{1-\alpha}\ ; & \text{if } \alpha \le \frac{8}{9} \\ \frac{3}{2}\cdot\alpha - 1 + \frac{1}{2}\cdot\sqrt{9\cdot\alpha^2 - 8\cdot\alpha}\ ; & \text{if } \alpha > \frac{8}{9} \end{cases} \quad (15)$$

By setting the saturation factor according to (15) the number of variables is reduced by one, as in:

$$X(\alpha) = X[\alpha, S_0(\alpha)] \quad (16)$$

leaving the stability to depend on the atomic beam intensity and the maser quality factor q.

For an already existent maser, with ionic pumps that do not support an atomic beam intensity as high as the optimum level, and with fixed couplings, the optimization stops at this first level.

The optimum saturation factor is to be determined through (15). The output resonance voltage results from (8), and the carrier voltage amplitude V_g of the interrogation signal generator is to be calculated by (1).

2.5 Since the oscillation parameter α is a function of the beam intensity and of the quality factor:

$$\alpha(q, \tilde{I}) = -\frac{\tilde{I}}{q\cdot(1 + c\cdot\tilde{I} + 2\cdot\tilde{I}^2)} \quad (17)$$

and a parameter of the X stability indicator, the following stability indicator is defined:

$$Y(q, u, \tilde{I}) = \frac{u + 2\cdot\tilde{I}}{\sqrt{q\cdot(1 + c\cdot\tilde{I} + 2\cdot\tilde{I}^2)}}\cdot X[\alpha(q, \tilde{I})] \quad (18)$$

to account for the atomic beam intensity influence on the maser stability.

The second level of optimization is realized by varying the beam intensity to obtain the minimum value of Y, keeping q and u as parameters independent on the operating conditions. This minimum is numerically calculated in the equation:

$$\frac{\partial Y(q, u, \tilde{I})}{\partial \tilde{I}} = 0 \quad (19)$$

The calculated dependence of the normalized beam intensity $\tilde{I}$ and the indicator Y, for c=3 (i.e. $(T_1)_0=(T_2)_0$), are plotted in Fig. 1 and, respectively, Fig. 2. For lower values of c, the calculated values of the normalized beam intensity and of the indicator Y are reduced by up to 15%.

The optimum normalized intensity takes values nearby 0.5, with a threshold at q=0.188, when α=8/9, where the (15) solution has its first derivative discontinuity. As a comment, it would worth to mention that this result is smaller then Ref. 2: 6.7.51 that predicts a optimum of 0.845.

The same threshold appears for the stability indicator Y, but, for q>0.188, the dependence becomes nearly linear, as predicted by Ref. 2: 6.7.52.

With the second level of optimization one should determine the optimum normalized beam intensity from Fig. 1 and proceed with the first level of optimization, as described above.

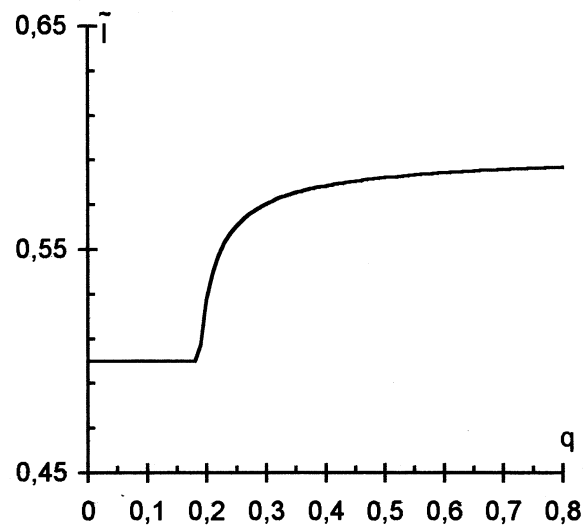


Fig. 1 Calculated optimum normalized intensity for masers with quality factors q<0.8 and c=3

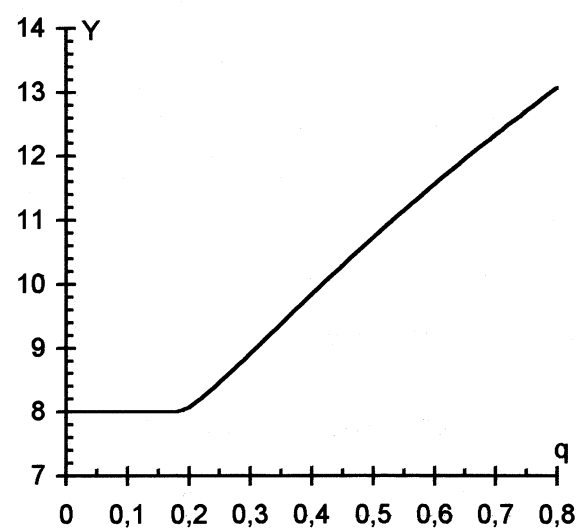


Fig. 2 Calculated optimum stability indicator Y for masers with quality factors q<0.8 and c=3

For an already existing maser whose coupling factors cannot be modified, the optimization stops

at this level, restricted to the operating conditions: interrogation signal level and atomic beam intensity.

The intrinsic stability of the maser can be determined at this optimization level because the external noise is not introduced yet.

2.6 In order to determine the optimum coupling factors, the external noise should be considered. The maser stability, including the external noise, is defined by the following relation:

$$\sigma'_y = \sqrt{F} \cdot \sigma_y \qquad (20)$$

using the total input and output noise figure F of the locking loop.

Assuming that β_1 is negligibly small compared to β_2 (Ref. 2: page 1124, par. 2), and considering the equivalent noise temperature (reflected in the output circuit of the cavity) of the amplification chain Θ_a, the noise figure, at resonance, is:

$$F \cong 1 + \frac{1+\beta+\beta_m}{\beta} \cdot \frac{\Theta_a}{\Theta_c} = 1 + \frac{1+\beta}{\beta} \cdot \frac{1}{G_0} \cdot \frac{\Theta_a}{\Theta_c} \qquad (21)$$

It is easy to observe that low coupling factors increase the noise figure while high coupling factors increase q factor and the Y value. Also, the variation of q implies changes of the optimum value of the gain G_0. There is an optimum coupling between these two extreme situations, where a third stability indicator named Z

$$Z(q_0, \frac{\Theta_a}{\Theta_c}, \beta) = \sqrt{F} \cdot Y\left[q_0 \cdot (1+\beta)\right] \qquad (22)$$

takes the minimum value possible. This optimum total coupling factor is plotted in Fig. 3:

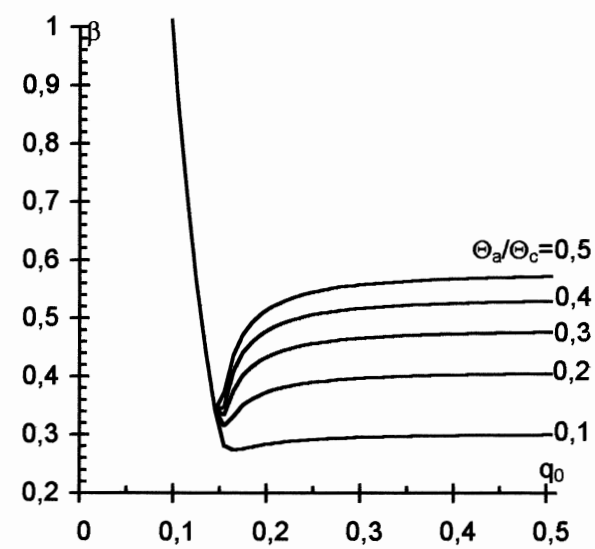


Fig. 3 Calculated optimum total coupling factor β for masers with quality factors q_0<0.5, c=3, and ratio Θ_a/Θ_c=0.1,0.2..0.5

For q_0<0.15, the optimum β is:

$$\beta = \frac{0.188}{q_0} - 1 \qquad (23)$$

in order to obtain the threshold value q=0.188 mentioned in sect. 2.5, and this optimum value does not depend on Θ_a.

For q_0>0.15, the optimum β is numerically calculated from the equation:

$$\frac{\partial Z\left(q_0, \frac{\Theta_a}{\Theta_c}, \beta\right)}{\partial \beta} = 0 \qquad (24)$$

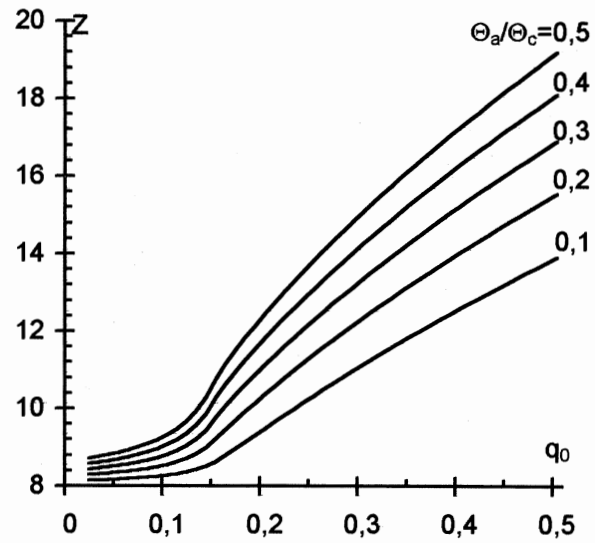


Fig. 4 Calculated optimum stability indicator Z for masers with quality factors q_0<0.5, c=3, and ratio Θ_a/Θ_c=0.1,0.2..0.5

Despite that the optimum β has a first derivative discontinuity, the stability indicator Z does not present this discontinuity and tends to Y for Θ_a tending to zero.

This third level of optimization is usable to masers where the coupling factors can be modified to optimally match them with the control loop noise.

With the coupling factor determined through this last optimisation level of the designing parameter β, the previous levels are appliable.

Also, the stability of the maser could be predicted.

An experimental iterative method to set the optimal values of the operating conditions involves stability measurements and is very difficult.

3. EXAMPLES

As an exemplification of the above optimization theory, the algorythm was applied to a normal and a reduced size types of maser cavities, token from Tab. 1 of Ref. 4.

The first two examples are for a normal size cavity maser. For all the examples $(T_1)_0=(T_2)_0$ and $\Theta_a=$ 75 K were supposed.
In the first example, the atomic beam intensity was left unchanged and only the saturation factor was optimized.
In the second example the atomic beam intensity was optimized, with a notable improvement of the predicted stability.
In the third example, was considered a reduced size cavity maser. A full optimization of the coupling factor, beam intensity and saturation was calculated.

Tabel:

Qty	Ex. 1	Ex. 2	Ex. 3
Q_0	60 000	60 000	6 000
V_b	$2.35\ 10^{-3}$m	$2.35\ 10^{-3}$m	$1.15\ 10^{-3}$m
T_t	0.4 s	0.4 s	0.2 s
q_0	0.0435	0.0435	0.5
q	0.087	0.087	0.72
I_{t1}	$2.67\ 10^{13}s^{-1}$	$2.67\ 10^{13}s^{-1}$	$5.23\ 10^{13}s^{-1}$
$\tilde{I}$	0.122	0.5	0.59
P_1	$1.26\ 10^{-11}$W	$1.26\ 10^{-11}$W	$2.46\ 10^{-11}$W
α	1.004	1.916	0.118
S_0	0.911	0.954	0.921
Y	10.1	8.0	17.0
σ_y	$7.4\ 10^{-14}\tau^{-1/2}$	$5.9\ 10^{-14}\tau^{-1/2}$	$1.8\ 10^{-13}\tau^{-1/2}$
Θ_a/Θ_c	75/313	75/313	75/313
β	1	1	0.44
σ'_y	$8.2\ 10^{-14}\tau^{-1/2}$	$5.9\ 10^{-14}\tau^{-1/2}$	$2.0\ 10^{-13}\tau^{-1/2}$
levels appl.:	1st level	2nd level	3rd level

4. CONCLUSIONS

The stability of the output error signal was used as performance criterium for the optimization problem. For the definition of the stability was used the power spectral density of the fractional thermal amplitude fluctuations, because it is the quantity that affects the envelope detection and amplitude integration of the mistuning voltage.
A three level optimization algorythm was developed:

1. the saturation factor optimization: for the masers were the atomic beam intensity and the coupling factors are given; this level provides the optimum interrogation generator voltage and the corresponding output voltage;
2. the atomic beam intensity optimization: for the masers were the coupling factors are given; this level is to be followed by the first one;
3. the total coupling factor optimization: for the masers were the couplings are to be designed, or could be changed; this level takes account on the equivalent noise temperature of the noise of the entire amplification, detection and feed-back loop reflected in the output circuit of the cavity; this level is to be followed by the second one.

The calcutated optimal stabilities for maser cavities taken from literature (Ref. 4) have proved to be better then the results of the quoted reference.
All the other practically interesting quantities, such as the voltages, could be easily calculated with the provided relations, in any particular case of specific proportion between β_1 and β_2.

5. REFERENCES

1. Busca G. *Proc. 33th Ann. Symp. on Freq. Contr.* p. 563 (1979)
2. Audoin C., Vanier J. *The Quantum Physics of Atomic Frequency Standards* Adam Hilger, Bristol (1988)
3. Audoin C., Candelier V., Dimarcq N. *Conf. on Precision Electromagn. Measurements* (1990)
4. Audoin C., Viennet J., Lesage P. *J. Physique **42** Suppl. C8 p. 159* (1981)

CURRENT PROGRESS WITH THE NPL CAESIUM FOUNTAIN FREQUENCY STANDARD

M. Zucco, G. Dudle, P.B. Whibberley, S.N. Lea and D. Henderson
National Physical Laboratory, Teddington, Middlesex TW11 0LW, UK

ABSTRACT

We report progress on NPL's prototype caesium fountain frequency standard. The vacuum chamber, microwave cavity, and laser systems are complete and we anticipate obtaining our first Ramsey fringes this spring. In this paper we outline the design of our device and discuss our Raman velocity selection scheme.

1 INTRODUCTION

A microwave frequency standard based on an atomic fountain of laser-cooled caesium has now been demonstrated to be amongst the best available primary frequency standards and has been used to provide steering corrections to TAI with estimated accuracy exceeding that of thermal beam standards [1, 2]. We are engaged in the construction of a device of this type at NPL. A demonstration device was constructed in collaboration with NPL by C.J. Foot and co-workers in the Clarendon Laboratory, University of Oxford, as reported at the last Neuchâtel EFTF [3]. As our intention is to produce a well-characterised working standard, we have adopted a design similar to that implemented at LPTF. An analysis of the expected performance of this design has been presented by several authors [4, 5].

2 FOUNTAIN DESIGN

The NPL fountain is housed in the basement of Bushy House, the 17th century mansion at the focus of the NPL site. This location is reasonably thermally stable and is well isolated from sources of vibration, enabling the fountain and associated lasers to sit on an optical table which bears directly onto the floor. One of NPL's two Sigma-Tau hydrogen masers will be installed in an adjacent room, forming a convenient reference for our local oscillator.

The design of our fountain is generally similar to that at LPTF (figure 1). After trapping, cooling and launching, a cloud of cold caesium atoms passes through a microwave cavity 28 cm above the trapping region, continues its ballistic flight above the cavity for around 0.5 s in a weak magnetic field (C-field) region, falls back through the microwave cavity to complete the Ramsey interaction and finally undergoes state-sensitive detection in a region 21 cm below the trapping region. The cooling beam geometry is such that one pair of beams (z-beams) is aligned along the vertical axis of the fountain. The trapping chamber is a ten way cross of low-permeability stainless steel (316LN). The ports at 45° to the x, y cooling beams will be used for Raman velocity selection. Anti-Helmholtz coils for MOT operation and Helmholtz coils for magnetic sub-level separation are wound on formers about the vertical axis of the chamber. The longitudinal C-field is formed by a solenoid wound on a titanium tube which also forms the vacuum vessel for this part of the fountain. The C-field region is enclosed within three layers of mu-metal magnetic shielding and will be temperature-stabilized to control black-body shifts. Above the C-field chamber is a four-way cross providing access for pumping whilst the detection region is a six-way cross attached below the trapping chamber. The vacuum chamber is supported on an aluminium structure which also carries

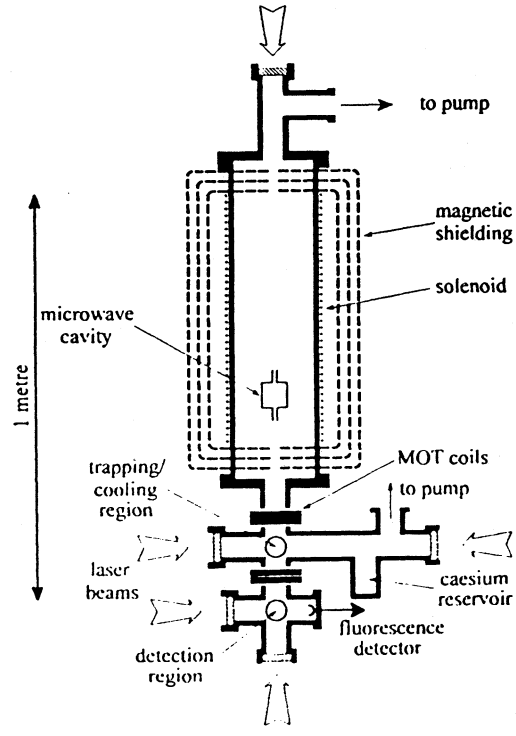


Figure 1: Schematic section of the NPL fountain.

the trapping and cooling optics, detection optics, and magnetic field compensation coils and is completely enclosed in a further layer of mu-metal shielding. Local magnetic field fluctuations ($\simeq$ 100 nT) are caused by ground return currents between local railway lines and may require active servoing of the compensation coil currents.

The microwave cavity is a cylindrical TE_{011} cavity [6] carefully tuned and temperature-stabilised to be in resonance with the caesium frequency. The B-fields of the microwave cavity and C-field solenoid should be accurately parallel to avoid Majorana transitions. The cavity axis has been mechanically aligned to better than 1 mrad of the axis of the C-field chamber.

Master and repumper extended-cavity diode lasers (SDL 5412, SDL 5422) are locked to the $F = 4 - F' = 5$ and $F = 3 - F' = 3$ components respectively of the D2 line in a saturated absorption scheme in thermal caesium vapour cells. The master laser provides the detection beams, whilst the cooling beams are produced by two diode lasers (SDL 5422), injection-locked to the master laser with appropriate detunings provided by AOMs. The repumper beam is shifted to the $F = 3 - F' = 4$ transition by a double-pass through an AOM. As a result of the locking scheme, all the beams are frequency modulated at 80 kHz with peak deviation of 400 kHz. A PC running LabView generates the timing sequence for laser frequency switching and intensity control via the AOMs, magnetic field switching, mechanical shutters on the cooling beams, and detection signal processing as well as tuning the local oscillator frequency between fountain cycles.

Fluorescence from the detection beams is collected by large numerical aperture aspheric lenses on both sides and focussed onto large-area photodiodes. Analogue Modules low-noise amplifiers are used to ensure good signal-to-noise.

3 TRANSVERSE VELOCITY SELECTION

The density of atoms in the fountain is much greater than in an atomic beam and hence collisional processes are of greater importance. In particular, the spin-exchange shift due to collisions between cold ground-state atoms [7, 8] is likely to be one of the limiting factors in the realisation of the potential precision of the atomic fountain, since it is particularly large for ultra-cold atoms. The shift has been measured [8, 9] and it has been shown that for a realistic atom density using optical molasses with atoms selected in $M_F = 0$, the frequency shift is at the 10^{-15} level. The shift can be further reduced to the 10^{-16} level, without degrading the fountain's performance, by transverse velocity selection. The cloud

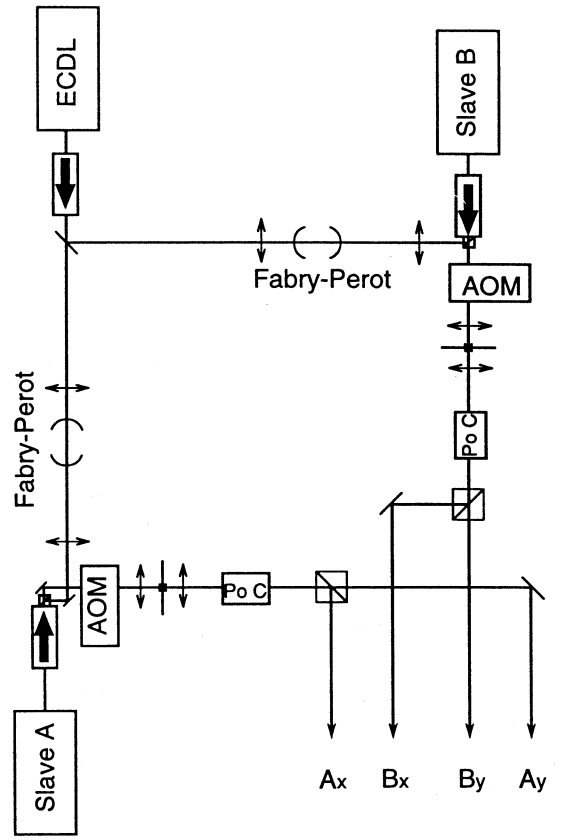


Figure 2: Optical set-up for Raman velocity selection.

of launched atoms thermally expands as it performs its ballistic flight. Atoms in the wings of the transverse velocity distribution, which undergo the first microwave interaction, will not pass a second time through the cavity, thereby contributing to the collisional shift but not to the detected signal. For typical fountain parameters, more than 90% of the atoms are lost. By rejecting these atoms from the fountain before the first microwave interaction, the collisional shift can be reduced without reducing the detected signal. Building on work started at LPTF by one of us [5] we will perform this velocity selection by means of stimulated Raman transitions [10, 11, 12].

The set-up being used on the NPL fountain is shown in figure 2. To drive the Raman transition we require counter-propagating laser beams whose frequency difference is accurately controled and tunable at around the hyperfine transition frequency. This is achieved by a method previously implemented at LPTF by K. Szymaniec and A. Clairon. A laser diode mounted in the extended cavity configuration is directly modulated at 1.532 GHz. The length of the extended cavity is chosen such that the modulation sidebands are resonant, leading to efficient power transfer [13, 14] (figure 3). The third order sidebands are selected by Fabry-Perot filters and

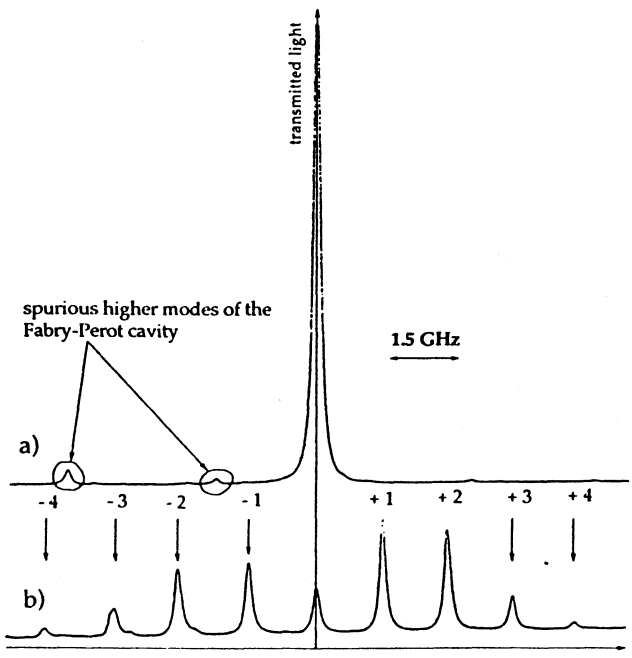


Figure 3: Frequency spectrum of the Raman master laser: a) without and b) with rf modulation ($P_{\rm rf}$ = +6 dBm). The laser parameters are $I_{\rm LD}$ = 80.3 mA, T = 18.4°.

injection-lock slave lasers which provide the Raman beams. The frequency offset between the slave lasers is thus 6 × 1.532 GHz.

It is currently planned to implement 2D Raman cooling [12, 15]. Each slave laser beam passes through an AOM which provides fast control of the optical power. A Pockel cell together with a polarising beam cube selects the direction along which the Raman pulses are applied. In a first attempt we will cool only half the velocity space in each direction ($v_x > 0$ and $v_y > 0$). This reduces both the experimental complexity and sequencing difficulties.

4 CONCLUSION

Construction of the NPL prototype caesium fountain frequency standard is nearly complete and Ramsey fringes will shortly be obtained. Once operational, a full programme of evaluation of systematic shifts will be undertaken, enabling a primary standard to be constructed. The fountain has been designed to include transverse velocity selection to reduce the cold collisional shift; a laser system for driving Raman transitions has been developed.

ACKNOWLEDGMENT

We acknowledge many useful discussions with A. Clairon and co-workers at LPTF and thank them especially for drawings for the microwave cavity. MZ acknowledges support from the EU Human Capital and Mobility programme and GD acknowledges support from the Swiss National Science Foundation and the Royal Society. This work is supported by the UK National Measurement System Foundation Programme. The National Physical Laboratory is operated on behalf of the DTI by NPL Management Ltd, a wholly-owned subsidiary of Serco Group plc.

References

[1] S. Ghezali *et al.*, *Proc. CPEM '96*, to appear in *IEEE Trans. Instrum. Meas.*

[2] *Circular T* **98** (18th March 1996); *ibid* **101** (13th June 1996), BIPM.

[3] A. Michaud *et al.*, *Proceedings 7th EFTF, Neuchâtel*, 1993, 528.

[4] K. Gibble and S. Chu, *Metrologia* **29**, 201 (1992).

[5] S. N. Lea *et al.*, *Physica Scripta* **T51**, 78 (1994).

[6] G. Vecchi and A. De Marchi, *IEEE Trans. Instrum. Meas.* **42**, 434 (1993).

[7] E. Tiesinga *et al.*, *Phys. Rev. A* **45**, R2671 (1992).

[8] K. Gibble and S. Chu, *Phys. Rev. Lett.* **70**, 1771 (1993).

[9] S. Ghezali *et al.*, *Europhys. Lett.* **36**, 25 (1996).

[10] M. Kasevich *et al.*, *Phys. Rev. Lett.* **66**, 2297 (1991).

[11] K. Moler *et al.*, *Phys. Rev. A* **45**, 342 (1992).

[12] S. Ghezali *et al.*, these proceedings.

[13] C. J. Myatt *et al.*, *Opt. Lett.* **18**, 649 (1993).

[14] A. P. Willis and D. M. Kane, *Optics Communications* **111**, 577 (1994).

[15] M. Kasevich and S. Chu, *Phys. Rev. Lett.* **69**, 1741 (1992).

A potassium vapor magnetometer optically pumped by diode laser.

E. Alzetta[a], N. Beverini[a,c], C. Carmisciano[c], A. Donati[b], O. Faggioni[c], E. Maccioni[a], E. Simeoni[b], F. Strumia[a]

[a]*Dipartimento di Fisica dell'Università di Pisa , Pisa (Italy)*
[b]*Kayser Italia SpA, Livorno (Italy)*
[c]*Consorzio Universitario della Spezia, La Spezia (Italy)*

Geophysical investigation, both in regional scale in order to model crustal composition, and in local scale, like the prospection in archeological sites, needs of magnetometer instruments with high sensitivity and accuracy.

Optically pumped magnetometer is the current choice when the highest sensitivity is needed. Optical pumping by circular polarized resonance radiation allows the detection of the Zeeman resonance frequency in alkali ground state atoms, giving the scalar value of the local magnetic field.

Commercial instruments make use of Cesium atom, and of a radiofrequency resonance lamp as pumping source. The use of Potassium, in particular the isotope 41, can however be advantageous. It has a nuclear spin of 3/2 (in spite of 7/2), a value of gyromagnetic ratio two times larger (7 Hz/nT) and a very small hyperfine structure (242 MHz). This means a smaller number of Zeeman transitions, that are more easily resolved: at the typical geomagnetic field of 50000 nT, Cs presents 14 transitions at about 6 Hz of distance one from the other, while Potassium 41 presents only 6 transitions separated by about 960 Hz. The choice of Cesium was dictated by the difficulty of realizing reliable K lamps.

We present here the experimental work performed in order to build an optically pumped potassium magnetometer, where the pumping source is a diode laser emitting at 770 nm, corresponding to the D1 line of the potassium spectrum. From the analysis of the signal to noise ratio it is possible to predict a sensitivity better than 1 $pT/\sqrt{Hz}$.

Diode laser noise conversion in optically thick atomic sample

M. Bahoura, Ph. Laurent, G. Santarelli, and A. Clairon
BNM/Laboratoire Primaire du Temps et des Fréquences,
Observatoire de Paris, 61, Av. de l'Observatoire, 75014 Paris, France
E-mail: Messaoud.Bahoura@obspm.fr

Diode lasers from the visible to the near infrared have opened the way to important developments in frequency standards and atomic spectroscopy. The frequency performances of such optical standards critically depend on the quality of the optical reference signal which may be an absorption signal or a laser-induced fluorescence signal. In many applications, it becomes of interest to understand the characteristics of the noise of the transmitted light through the absorbing medium. Recent experiments [1] have found that the intensity power noise spectrum in absorption of an atomic vapor contains all corresponding hyperfine-splitting resonances. A related experiment has been performed [2] using a broad-bandwidth laser beam which had passed through a Rubidium vapor cell. All these studies are related to weakly saturating laser field.

In our experiment a broadband diode laser beam is directed through an optically-thick Cesium cell and detected with a large band and fast photodetector. The detector output is sent to a spectrum analyzer that is set to a single frequency. Oscillating noise spectra depending on laser frequency tuning from resonance and Cesium cell temperature were obtained. Similar spectra were observed using Rubidium vapor cell [3]. For high Fourier frequencies, the noisy laser spectrum can be thought of as laser field modulated by white frequency noise with continuum sidebands rather a single sinusoidal frequency in FM spectroscopy. So we have to consider the contribution of different components at different Fourier frequencies for the transmitted signal. Our experimental noise spectra are in good agreement with a simple theory [3] based on the ideas of FM spectroscopy. This model assumes that the laser field is consisted of a strong saturating carrier and weak white frequency noise sidebands. The theory takes into account the laser linewidth, Doppler broadening effect, Cesium natural linewidth, hyperfine structure and Cesium cell temperature. The power noise spectrum of the transmitted beam for high spectrum analyzer frequency was found to be extend as far as several Doppler widths.

In conclusion, a significant amount of excess noise is acquired by a diode laser beam after passing through an optically-thick Cesium vapor cell. This noise exhibits an oscillating behavior and depends besides technical noise of the laser, on the dynamics of the dense medium.

References

[1] T. Yabuzaki, T. Mitsui, and U. Tanaka, Phys. Rev. Lett., **67**, 2453 (1991).
[2] D. H. McIntyre, C. E. Fairchild, J. Cooper, and R. Walser, Opt. Lett., **18**, 1816 (1993)
[3] L. Hollberg et al. to be published in Opt. Lett.

SECOND-HARMONIC-GENERATION OF Nd:YAG LASER FOR ABSOLUTE FREQUENCY STABILIZATION AND SPECTROSCOPY OF $^{127}I_2$ HYPERFINE STRUCTURE NEAR 532 nm

E. Bava, V. Busetti, G. Galzerano and F. Massari

Dipartimento di Elettronica e Informazione del Politecnico di Milano,
P.zza Leonardo da Vinci 32, 20133 Milano, Italy
E-mail: bava@elet.polimi.it Phone: +39(2)239936-09/01 Fax: +39(2)23993635

Abstract - This work deals with the realization of Second Harmonic Generation (SHG) of a Nd:YAG continuos wave diode pumped laser operating at 1064 nm and 100 mW. The laser frequency is doubled by using a MgO:$LiNbO_3$ heated to its phase matching temperature of 53.9 °C, placed inside an external built up resonant cavity. With 55 mW of fundamental power 0.8 mW of green light has been obtained.

1. INTRODUCTION

Owing to their intrinsic characteristics of short term frequency stability (improved by stabilization to resonant cavity), high reliability and potentially high power, monolithic diode-pumped Non Planar Ring Oscillator (NPRO) Nd:YAG lasers are attractive sources for applications in optical communications, spectroscopy (Refs. 1,2) and metrology (Refs. 3,4). However, for these applications, long term stability has to be improved. For Nd:YAG laser at 1064 nm an efficient absolute stabilization needs second harmonic generation to lock the doubled frequency to a suitable rovibrational transition of $^{127}I_2$ at 532 nm. For power level lower than 1 W, efficient SHG requires some method of increasing the intensity in doubling crystal. The use of an external resonant cavity to enhance the fields present in the non linear material represents a good solution. It was first demonstrated by Askhin et al (Ref. 5) and subsequently other experiments have produced optimal results (Refs. 6,7).
The stabilized source could then be used for frequency measurements in both fundamental and second harmonic regions.

2. THEORY AND EXPERIMENT

The experiment consists in keeping the frequency of a NPRO Nd:YAG diode-pumped laser at 1064 nm locked to the external doubler cavity shown in Figure 1.
The cavity has four plane mirrors (M_1 has T_1=5%), two lenses (f=50 mm) and a 12 mm long MgO:$LiNbO_3$ crystal, with a nominal 7% of MgO and with surfaces antireflection coated for both wavelengths. This type of cavity, because of losses higher than a monolithic resonator (Ref. 7), does not allow high efficiency, but it can be used in different configurations with the same crystal and, for high circulating power, thermal-lensing effect can be compensated.

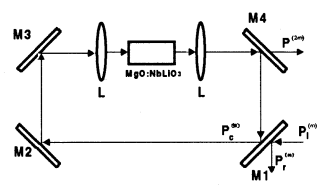


Figure 1. Scheme of the symmetrical ring cavity used.

The SHG conversion efficiency of an external cavity doubler may be determined from the theory of Boyd and Ashkin (Ref. 5).
Let $\eta=P_2/P_i$ be the total SHG efficiency, where P_i and P_2 are the fundamental incident power and harmonic output power. It can be expressed as:

$$\sqrt{\eta}=\frac{T_1\sqrt{\gamma_{shg}P_i\kappa}}{\left[1-\sqrt{1-T_1}\left(1-\gamma_1/2-\sqrt{\gamma_{shg}\eta\kappa P_i}/2\right)\right]^2}$$

where γ_{shg}, γ_1, T_1 and κ are respectively the single-pass SHG efficiency function of the non linear material and of the focusing parameters (scattering, absorption and diffraction), the round-trip power loss (not inherent the process of duplication), the input mirror trasmissivity and the mode-matching coefficient that takes into account a possible mode mismatch.
The total efficiency is optimized at the value η^o when T_1 is:

$$T_1^o=\gamma_1/2\left(1+\sqrt{1+4\gamma_{shg}P_i\kappa/\gamma_1^2}\right)$$

$$\eta^o=\frac{4\gamma_{shg}P_i\kappa/\gamma_1^2}{\left[1+\sqrt{1+4\gamma_{shg}P_i\kappa/\gamma_1^2}\right]^2}$$

It is straightforward to show that $T_1=T_1^0$ corresponds to an impedance-matched cavity condition.

The cavity losses was evaluated to be 4.2%, hence the trasmissivity of the input mirror was chosen as T_1=5%.
$MgO:LiNbO_3$ (7%) was selected as nonlinear material for this experiment because of its large nonlinear coefficient for non critical phase-matching doubling of 1064 nm radiation (at temperature of 53.9 °C) and low photorefractive effect. The temperature of the crystal was maintained by a small oven and temperature controller. In such conditions we have measured γ_{shg}=0.938 W^{-1}, γ_1=4.5 % and κ=0.7.
The locking between the laser frequency and the cavity resonance is obtained through two different techniques either by frequency modulation technique (Ref. 8) using an electrooptic modulator and the piezoelectric transducer bounded on the Nd:YAG crystal (figure 2), or with the polarization spectroscopy exploiting the cavity anisotropy (figure 3) (Ref. 9).

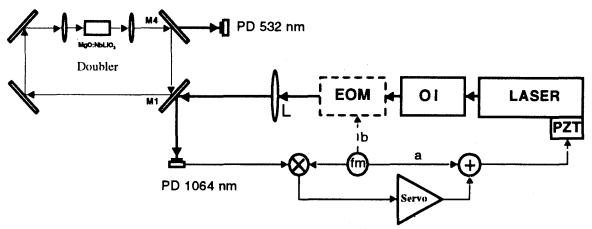


Figure 2. Optical and electronics set-up used to lock the fundamental frequency to the resonant cavity: frequency modulation by mean of piezoelectric -**a**- or by an external electrooptic modulator -**b**-.

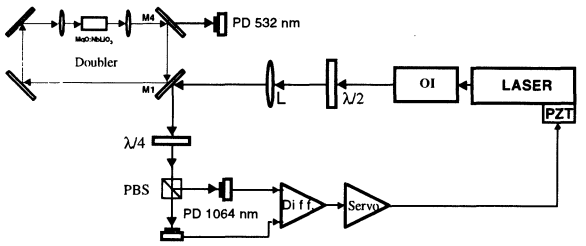


Figure 3 Optical and electronics set-up used to lock the fundamental frequency to the resonant cavity by mean of polarization spectroscopy technique. PBS polarization beam splitter.

From the measured value of γ_{shg}, γ_1, κ and with $T_1=T_1^0$ the theoretical result for total efficiency and output power are shown in figures 4a and 4b as a function of incident power. Also plotted in the figures are our experimental results for η and P_2 versus P_i.
In these configurations we have obtained a maximum output power of 0.8 mW with 55 mW of incident fundamental power.

Such efficiency was limited by the cavity losses and the modal mismatch. We are analysing different cavities geometry (either ring or linear) and combined locking techniques to reduce the round-trip losses and to improve the reproducibility.

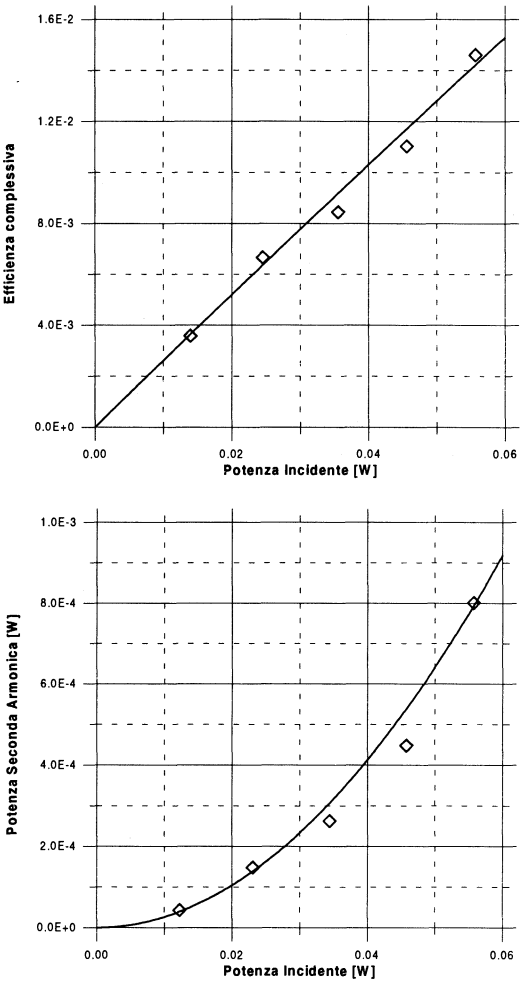


Figure 4. a) Total SHG efficiency and **b)** SHG power versus the fundamental input power P_i. The measured cavity loss are 4.5%, single SHG efficiency in unit of power is $\gamma_{shg}=0.938\cdot10^{-3}$ W^{-1} and κ=0.7.

3. SOME APPLICATIONS

This type of source will be absolute frequency stabilized respect to a rovribational transition of $^{127}I_2$ with the methods of the transfer modulation (Ref. 4) and frequency modulation (Refs. 1,3) Doppler free spectroscopy shown in figure 5.
This stabilized source in addition to a commercial doubled Nd:YAG will be used in our laboratory to extend frequency measurement capabilities around 532 nm.
The sum frequency of the standard He-Ne CH_4 (at 88 THz, 3.39 μm) and He-Ne $^{127}I_2$ (at 474 THz, 633 nm) lasers is only 1.5 THz from the frequency of the doubled Nd:YAG laser.

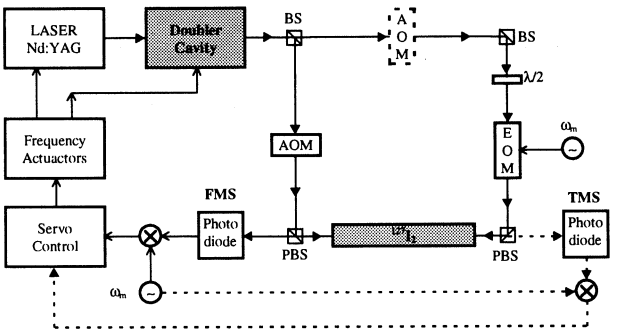


Figure 5. Experimental configuration for frequency stabilization of Nd:YAG laser at 532 nm to a rovibrational transition of $^{127}I_2$ by mean of frequency modulation frequency and transfer modulation techniques (dashed line). AOM acoustic optic modulator, EOM electrooptic modulator, PBS polarization beam splitter and BS beam splitter.

Using a frequency comb generator (Ref. 10), built with an electrooptic crystal placed inside a resonant cavity where the modulation frequency (some GHz) is a integer of the free spectral range of the resonator, we could make absolute

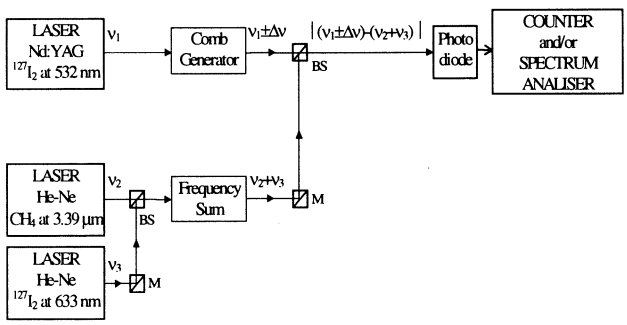


Figure 6. Absolute frequency measurement of the Nd:YAG laser through He-Ne frequency standard. BS beam splitter and M mirror.

frequency measurements of several line of $^{127}I_2$ in the spectral region of 532 nm following the scheme shown in figure 6. In this way it could be possible to measure the frequency of the transition R(56) 32-0 of $^{127}I_2$ (a probable candidate for the realization of the meter) with an uncertainty of 50 kHz. This measurement will be performed in cooperation with the Italian Institute of Metrology IMGC and IEN where the He-Ne standards are available.
Moreover it will be possible a comparison between our laser and the green sources available at metrology institute IMGC.

REFERENCES

[1] A. Arie, and R. L. Byer, *J. Opt. Soc. Am. B*, vol. 11, November 1993.

[2] P. A. Junger, S. Swartz, M. Eickhoff, J. Ye, J. L. Hall and S. Waltman, *IEEE Trans. Instrum. and Meas.*, vol. 44, n. 2, April 1995.

[3] A. Arie, S. Schiller, E. K. Gustafson and R. L. Byer, *Optics Lett.*, vol. 17, n. 17, September 1, 1992.

[4] M. Eickhoff and J. L. Hall, *IEEE Trans. Instrum. and Meas.*, vol. 44, n. 2, April 1995.

[5] A. Ashkin, G. D. Boyd and J. M. Dziedzic, *IEEE J. Quantum Electron.*, vol. QE-2, pp. 109-123, June 1966.

[6] E. S. Polzik and H. J. Kimble, Optics Letters, vol. 16, n. 18, Sept. 1991.

[7] W. J. Kozlovsky, C. D. Nabors and R. L. Byer, IEEE J. Quantum Elect., vol. 24, n. 6, June 1988.

[8] E. Bava and F. Massari, *Rev. Scient. Instr.* (April, 96)

[9] T.W. Hansch and B. Coullaud, Optics Comm., vol.35, n. 3, Dec. 1980.

[10] M. Kourogi, T. Enami and M. Ohtsu, IEEE Phot. Techn. Lett, vol. 6, n. 2, Feb. 1994.

PRESENT FREQUENCY STANDARDS FOR LASERS

D. J. E. Knight

DK Research,
110 Strawberry Vale, Twickenham, Middx., TW1 4SH, UK
E-mail: djek@cix.co.uk

1. ABSTRACT

Since a previous compilation of accurate frequency measurements at laser frequencies (Ref.1) there have been many new measurements, and a new "Mise en pratique" for the definition of the metre has been published (Ref.2). A short list of accurate frequency measurements, including some on cooled atoms and ions, is presented, together with the recommended frequencies from the "mise en pratique".

Keywords: lasers, frequency standards, accurate, list

2. INTRODUCTION

In recent years there have been a number of accurate frequency measurements on reference systems for stabilised lasers, with uncertainties as small as 3 parts in 10^{13} (single-sigma), which amount to a substantial set of new frequency standards. Further, frequency measurements on laser-cooled atoms and ions are beginning to be made at accuracies which exploit the advantages of near-stationary systems for frequency standards, avoiding the km/s velocities of atoms or ions in gases or plasmas at room- or higher-temperatures. Shortly the accuracy of such systems may challenge the accuracy of frequency measurement against conventional primary caesium-beam standards, presently about 1 part in 10^{14}. Nonetheless, the application of laser cooling and e.g. fountain geometry to the caesium microwave standard already produces an accuracy of 3 parts in 10^{15} (Ref.3), keeping it ahead of the optical standards. Highly accurate measurements have also been made on the more calculable atomic and sub-atomic systems, and these are concentrated at the higher frequencies.

This paper presents a list, Table 1, of the more- accurate optical frequency standards, or those which are potentially more accurate, and is mainly concerned with frequency measurement at frequencies above about 100 THz, 3 μm wavelength. A schematic diagram of the main routes of frequency measurement from the caesium standard to these frequencies is shown in Figure 1. Cooled-atom or -ion frequency standards at lower frequencies are also included, together with high-accuracy transfer standards such as OsO_4 near 30 THz, that remain outside the "mise en pratique" for the metre. Some information is drawn from a previous list (Ref.1), and transitions of interest for frequency standards are covered (Ref.4). Stabilised lasers listed in the 1992 revision of the "mise en pratique" of the definition of the metre (Ref.2) are listed at the foot of Table 1, with their recommended frequencies and their uncertainties when operated as prescribed. All uncertainties in this table are single-sigma, but those of the recommended standards are more conservative than those in the measurements table, which each reflect one particular measurement. It is notable that a number of measurements and standards have uncertainties smaller than 1 part in 10^{11}, and these are shown in bold type. Apart from the entries in this Table, many additional laser frequency standards exist, such as have been listed up to 130 THz, 2.3 μm, for spectroscopic calibration (Ref.5) and in the 190-240 THz, 1.6-1.25 μm, region for optical communications etc. (Refs.6,7).

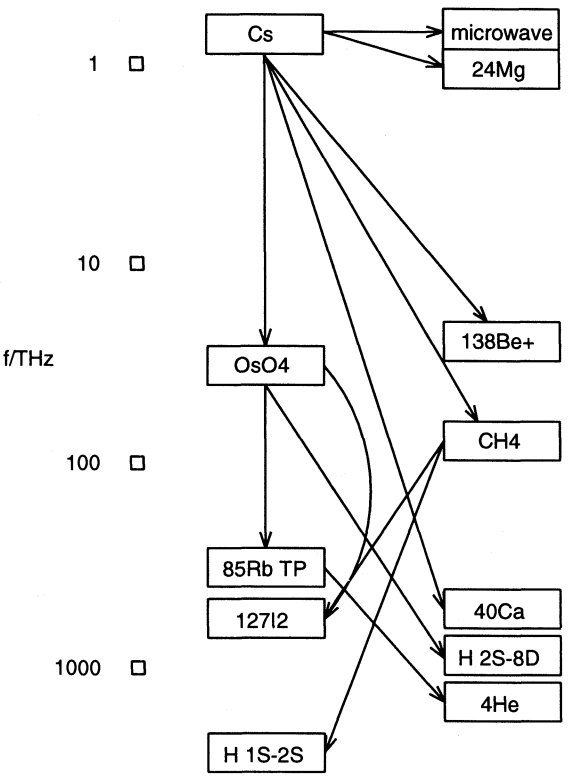


Figure 1: Frequency standard links

3. THE LIST FORMAT

The list in Table 1 is intended for transmission as a text file, or for use as a computer data file, with space to accommodate the frequency to 16 or more significant figures, together with its uncertainty. Engineering exponential notation is used. There is space for up to 3 references, also a brief note on the frequency measurement route used. There is not room however for the wavelengths in addition, except that for multiple gas reference lines, such as in I_2 or Te_2, the lines are given a

TABLE 1: FREQUENCY STANDARDS FOR LASERS

system	app*	reference(s)	label	frequency/Hz	uncert.	notes (line;)std used; etc
H	**B**	**Ud97**		**2466.06141318734**	**+12 3.4 -13**	**1S-2S TP;Cs via CH4**
muonium~	B	Sch95		2455.529002	+12 2.3 -08	1S-2S TP; 488Te2 d4
Ps#	~B	Ch84		1233.607185	+12 1.2 -08	1S-2S TP;D beta viaTe2
D	**B**	**deB97**		**770.8592528515**	**+12 5.7 -12**	**2S1/2-8D5/2TP;Cs/778Rb**
4He	G	Pav94		770.732839058	+12 2.4 -10	2\3S1-3\3P0;778Rb
H	**B**	**Nez93Nez95deB97**		**770.6495615850**	**+12 6.4 -12**	**2S1/2-8D5/2TP;Cs/778Rb**
172Yb+	T	Gi95Tay97		729.4688689	+12 5.5 -10	633I2 via 1m FP
172Yb+	T	Gi95Ro97		642.1167853	+12 1.1 -09	octopole;633I2/467Te2
130Te2	G	Ro97	467Te2	642.1165136	+12 9.3 -10	633I2
130Te2	G	Ba85	486'Te2	616.6788575	+12 4.0 -10	b1;633I2; also b2
130Te2	G	Ba91	488Te2	613.8811491	+12 7.6 -10	d4;633I2;+b2,b1,e3,i3
127I2	G	Jun95	532I2	563.260223480	+12 1.24-10	2f1064,a1;780Rb,633I2
127I2	G	Ed96	637I2	470.76645546	+12 2.2 -10	a4;633I2
40Ca	**T**	**Ri96**		**455.98624049407**	**+12 8. -13**	**3P1-1S0; Cs**
Sr+	T	Mad96Mar96		444.77904414	+12 8.8 -10	633I2
88Sr+	T	Ba95Ba97		444.77904398	+12 1.3 -10	633I2
85Rb	**G**	**To97**	**778Rb**	**385.28514237828**	**+12 5.2 -12**	**5S-5D F=3-5 TP; OsO4**
87Rb	G	Ba93	780Rb	384.22798187	+12 1.5 -10	D2,d/f; 633I2
87Rb	G	Ye96	780Rb	384.2279818773	+12 1.4 -11	D2,d/f; 778'Rb
85Rb	G	Ba93	795Rb	377.10627149	+12 1.5 -10	D1,c'; 633I2
Cs	B	Av86	852Cs	351.7257682	+12 3.5 -08	D2;633I2;1-m,hyp.res.
Cs	G	We96Er64	895Cs	335.116095	+12 9. -08	D1;
C2HD	G	Ye296		281.635363961	+12 1.2 -10	532I2; 1.3Pa
20Ne	L	Po83		260.10324926	+12 3.1 -10	CH4 F
20Ne	L	Sas92		196.7803338	+12 1.6 -08	633I2,1-m; 387Pa
12C2H2	G	Nak96		197.75046656	+12 7.6 -10	v1+v3 R(18)high;778Rb
13C2H2	G	Nak96		196.92974592	+12 7.6 -10	v1+v3 R(23)high;778Rb
12C2H2	G	Nak96		194.91619955	+12 7.7 -10	v1+v3 P(21)low;778Rb
13C2H2	G	Nak95Nak96		193.22000472	+12 5.7 -10	v1+v3 P(30)low;778Rb
HCN	G	Nak95		192.6224469	+12 5. -10	2v1 P(27); 778Rb
CO	G	Kni95		192.113068	+12 2.1 -08	0-3R(21);780Rb;40torr
CO	G	Po283		129.019381761	+12 5.4 -10	R12; CH4 F
CO	G	Po283		128.360848969	+12 5.5 -10	R05; CH4 F
CO	G	Po283		127.598179703	+12 6.3 -10	P01; CH4 F
CO	G	Po283		126.990495584	+12 6.3 -10	P06; CH4 F
CH4 E	G	Do81		88.3731490330	+12 1.9 -11	Cs via CH4 F; 0.26Pa
CH4 E	G	Za83		88.3731490312	+12 1.4 -11	Cs via Rb; 0.27Pa
172Yb+	T	Be92		87.360087	+12 5. -08	515I2, 633I2
OsO4	**G**	**Ac93**		**29.37081407841**	**+12 2.7 -12**	**CO2 R26; OsO4 R10**
OsO4	**G**	**CL85CL88Ac93**		**29.05405744666**	**+12 2. -12**	**CO2 R10; Cs**
192OsO4	**G**	**Do87**		**28.464676938787**	**+12 3.3 -12**	**CO2 P14; Cs**
138Ba+	T	Mad96		24.01204831717	+12 1.8 -11	Cs
24Mg	B	Go87		601.277157860	+09 3. -11	Ramsey; Cs
199Hg+	T	McG78		40.5073479978	+09 2.5 -11	uncooled; Cs via Rb
199Hg+	**T**	**Poi96Ber97**		**40.507347996496**	**+09 1. -13**	**Cs**
113Cd+	T	Tan96		15.199562858	+09 1.3 -10	rf trap; Cs
171Yb+	**T**	**Bau96**		**12.642812118471**	**+09 7. -13**	**Cs**
171Yb+	**T**	**Se95Fi96**		**12.642812118468**	**+09 1.3 -13**	**Cs**
9Be+	**T**	**Bo85Bo91**		**303.016377265070**	**+06 1.9 -13**	**Cs**
Some CIPM Recommendations						
127I2	G	Qu93	515I2	582.49060337	+12 2.5 -10	(a3),Ar+, CIPM1992
127I2	G	Qu93	543I2	551.57948296	+12 2.5 -10	(a9),HeNe,CIPM1992
127I2	G	Qu93	576I2	520.2068084	+12 4. -10	(a1),dye, CIPM1992
127I2	G	Qu93	612I2	489.8803549	+12 3. -10	(o),HeNe, CIPM1992
127I2	G	Qu93	633I2	473.612214705	+12 2.5 -11	(i),HeNe,CIPM1992;Cs
127I2	G	Qu93	640I2	468.2183324	+12 4.5 -10	(g),HeNe, CIPM1992
40Ca	B	Qu93		455.9862405	+12 4.5 -10	thermalbeam,CIPM1992
CH4 F	G,B	Qu93		88.3761816005	+12 2.3 -11	hyp.unres.CIPM1992;Cs
CH4 F	**G,B**	**Qu93**		**88.37618160018**	**+12 3.0 -12**	**hyp.res. CIPM1992;Cs**

Notes
(i) All uncertainties are 1-sigma.
(ii) Labels of I2 and Te2 lines have 3-digit wavelength in nm: e.g. 633I2.
Primes on e.g. 486'Te2 indicate differing lines.
*: app: apparatus:- B: in beam; G: in gas; L: in laser Lamb dip; T: in trap.
#: positronium (Ps)
1m FP: 1-metre Fabry Perot; hyp.res.: hyperfine resolved; 1-m: lambda meter; TP: two-photon.

wavelength-based label placed in the 'reference' column, and which is used to identify it in the 'notes' column. Apart from the 'system' column at the left, for which transition information can appear as the first item in the 'notes' column at the right, there is a single-letter column 'app.', for 'apparatus' which indicates how the system was interrogated: B: beam; G: gas; L: gas laser Lamb dip; or T: trap. ('~B' is entered for thermal evaporation from a nearby surface into vacuum.) This broadly correlates with increasing potential accuracy in the order L, G, B and T. All standards listed, except CO at 196 THz, use a method such as saturated absorption or two-photon excitation (TP) to eliminate the Doppler effect to first order. The Table references have codes in the form of two or three letters for the start of the first author's name plus two digits for the year, and are listed in Section 5.

4. DISCUSSION OF THE MEASUREMENTS

4.1 The entries

The uncertainty range of the measurements in Table 1 lies between about 1 part in10^8 and 3 parts in 10^{13}. It is notable that one of the high-accuracy measurements is for the highest frequency listed - for the H 1S-2S two-photon transition. Measurements made primarily to test the theories of physics, such as to compare with quantum electrodynamic calculations for such atoms as Ps, muonium, H, D and He contribute to the recent high-accuracy data. The trapped-ion or cooled-atom measurements do not yet fully exploit the possibilities of the method, e.g. for the octopole Yb^+ transition at 642 THz, which has an exceedingly narrow natural width, having a measured lifetime of 3700 days (see Ro97). In the latter case technical trade-offs exist, e.g. between signal/noise ratio and minimising the state perturbation, and there is the need to synthesise frequencies in very specific parts of the spectrum. (This list is constructed partly to aid future synthesis for frequency measurement.) Progress with cooled-atom or ion measurements is of interest, so that more than one measurement can be listed for the same transition, if by different groups or with respect to different standards.

A new standard in the list is the simple and accessible two-photon transition in Rb at 385 THz, 778 nm. Now measured to 5 parts in 10^{12} via OsO_4 and exhaustively tested, the standard comprises a gas cell interrogated by a readily-available diode laser. One well-researched and potentially very accurate ion-trap standard is omitted for want of a frequency measurement: the Hg^+ clock transition near 1065 THz, 281.5 nm (Ref.8). Recent investigations into 30 THz standards using CO_2 gas lasers indicate room for improvement (Refs.9,10) towards the accuracies achieved by portable hyperfine-resolved methane-stabilised lasers.

4.2 Technological advances

In the past decade, work on frequency standards has taken advantage of the long life and low cost of solid-state lasers, such as the diode laser, the Ti-sapphire laser or for low noise the Nd^{3+}-YAG laser at 1064 nm. Also for the new high-accuracy experiments, not forgetting high-stability systems used for gravitational wave detection experiments, lasers are being spectrally narrowed by locking to ultra stable, ultra-high-finesse cavities suspended in vacuum, to reach the spectral widths and frequency stabilities needed to interrogate ultra-narrow resonances (Ref.8).

4.3 Dissemination/transportation of optical frequency standards

For the intermediate accuracies of the Table, gas-cell standards, such as using I_2 or Te_2, have been measured in frequency-standards laboratories, and then transported to the experimental laboratory to use as a local reference (see e.g. Sch95). However, for the highest accuracy, complete stabilised-laser systems have had to be transported e.g. from Braunschweig to Munich and back (see Ud97). However, where distances of a few km have permitted installation of dedicated optical fibre, similar stabilised lasers in the two laboratories have been continuously compared over the fibre link. It is notable that at 385 THz, 778 nm, the fibre introduces frequency errors of only a few Hz, although contributing some added noise from e.g. vibration (see deB97).

4.4 Towards optical clocks

The new systems are also moving towards solid-state heterodyne mixers and frequency-combining elements, giving long-term reliability, which should offer the possibility of continuous operation, leading to experiments on the use of such systems as clocks. This is already the case for some of the standards listed in the microwave region, for example, the mercury ion standard at 40.5 GHz, although not yet in a laser-cooled form.

5. TABLE 1 REFERENCES

Ac92: (OsO_4) O. Acef, J. J. Zondy, M. Abed, D. G. Rovera, A. H. Gérard, A. Clairon, Ph. Laurent, Y. Millerioux and P. Juncar, "A CO_2 to visible optical frequency synthesis chain: accurate measurement of the 473 THz HeNe/I_2 laser", Optics Commun. **97**, 1993, 29-34.

Av86: (Cs D_2) G. Avila, P. Gain, E. de Clercq and P. Cerez, "New absolute wavenumber measurement of the D_2 line of caesium", Metrologia **22**, 1986, 111-114.

Ba85: ($^{130}Te_2$) J. R. M. Barr, J. M. Girkin, A. I. Ferguson, G. P. Barwood, P. Gill, W. R. C. Rowley and R. C. Thompson, "Interferometric frequency measurements of $^{130}Te_2$ transitions at 486 nm", Optics Commun. **54** (4), 15 June 1985, 217-221.

Ba91: ($^{130}Te_2$) G. P. Barwood, W. R. C. Rowley, P. Gill, J. L. Flowers and B. W. Petley, "Interferometric measurements of $^{130}Te_2$ reference frequencies for 1S-2S transitions in hydrogen-like systems", Phys. Rev. A, **43**, 1991, 4783-4790.

Ba93: (^{85}Rb, ^{87}Rb) G. P. Barwood, P. Gill and W. R. C. Rowley, "Optically narrowed Rb-stabilized laser GaAlAs diode laser frequency standards with 1.5 x 10^{-10} absolute accuracy", Proc. SPIE - Int. Soc. Opt. Eng. **1837**, 1993, 262-270.

Ba95: (Sr^+) G. P. Barwood, C. S. Edwards, P. Gill, G. Huang, H. A. Klein and W. R. C. Rowley, "precision interferometric frequency measurement of the 674 nm $^2S_{1/2}$-$^2D_{5/2}$ in a single cold Sr^+ ion", IEEE Trans. Instrum. Meas. **44** (2), April 1995, 117-119.

Ba97: ($^{88}Sr^+$, $^{86}Sr^+$) G. P. Barwood, P. Gill, H. A. Klein and W. R. C. Rowley, "Clearly resolved secular sidebands on the $^2S_{1/2}$-$^2D_{5/2}$ 674 nm

clock transition in a single trapped Sr^+ ion", IEEE Trans. Instrum. Meas., in the press, February 1997.

Bau96: ($^{171}Yb^+$) A. Bauch, D. Schnier and Chr. Tamm, "Microwave spectroscopy of $^{171}Yb^+$ stored in a Paul trap", Proc. 5th Symp. on Frequency Standards and Metrology, Ed. J. C. Bergquist (Singapore: World Scientific) 1996, 387-388.

Be92: ($^{172}Yb^+$) A. S. Bell, P. Gill, H. A. Klein, A. P. Levick and W. R. C. Rowley, "Precision measurement of the $^2F_{7/2}$-$^2D_{5/2}$ 3.43 μm interval in trapped $^{172}Yb^+$", J. Mod. Opt. **39** (2), 1992, 381-387.

Ber97: ($^{199}Hg^+$) J. C. Bergquist and D. Berkeland, private communication, 24 February 1997.

Bo85: ($^9Be^+$) J. J. Bollinger, J. D. Prestage, W. M. Itano and D. J. Wineland, "Laser-cooled atomic frequency standard", Phys. Rev. Lett. **54** (10), 11 March 1985, 1000-1003.

Bo91: ($^9Be^+$) J. J. Bollinger, D. J. Heinzen, W. M. Itano, S. L. Gilbert and D. J. Wineland, "A 303-MHz frequency standard based on trapped Be^+ ions", IEEE Trans. Instrum. Meas. **40** (2), April 1991, 126-128.

Ch84: (positronium (Ps), e^+e^-) S. Chu, A. P. Mills Jr. and J. L. Hall, "Measurements of the positronium 1^3S_1-2^3S_1 interval by Doppler-free two-photon spectroscopy", Phys. Rev. Lett. **52** (19), 7 May 1984, 1689-1692.

CL85: (OsO_4) A. Clairon, B. Dahmani, A. Filimon and J. Rutman, "Precise frequency measurements of CO_2/OsO_4 and HeNe/CH_4-stabilized lasers", IEEE Trans. Instrum. Meas. **IM-34** (2), June 1985, 265-268.

CL88: (OsO_4) A. Clairon, B. Dahmani, O. Acef, M. Granveaud, Yu. S. Domnin, S. B. Pouchkine, V. M. Tatarenkov and R. Felder, "Recent experiments leading to the characterization of the performance of portable (He-Ne)/CH_4 lasers: Part II: results of the 1986 LPTF absolute frequency measurements", Metrologia **25**, 1988, 9-16.

deB97: (H and D) B. de Beauvoir, F. Nez, L. Julien, B. Cagnac, F. Biraben, D. Touahri, L. Hilico, O. Acef, A, Clairon and J. J. Zondy, "Absolute frequency measurement of the 2S-8S/D transitions in hydrogen and deuterium: new determination of the Rydberg constant", Phys. Rev. Lett. **78** (3), 20 January 1997, 440-443.

Do81: (CH_4 E line) Yu. S. Domnin, N. B. Koshelyaevskii, Yu. M. Malyshev, Yu. G. Rastorguev, V. M. Tatarenkov, A. N. Titov: [VNIIFTRI report (in Russian) received 3 February 1982 from Yu. S. Domnin, pp. 52-54] and for the stabilised laser: Yu. M. Malyshev, S. N. Ovchinnikov, Yu. G. Rastorguev, V. M. Tatarenkov, A. N. Titov: Kvant. Elektron. **7**, 1980, 655-658, transl.: Sov. J. Quantum Electron. **10**, 1980, 376-377.

Do87: (OsO_4) Yu. S. Domnin, N. B. Koshelyaevskii, A. N. Malimon, V. M. Tatarenkov and P. S. Shumyatskii, "Infrared frequency standards based on osmium tetroxide", Sov. J. Quant. Electron. **17** (6), June 1987, 801-803; also Kvantovaya Elektron. (Moscow) **14**, June 1987, 1260-1263.

Ed96: ($^{127}I_2$) C. S. Edwards, G. P. Barwood, P. Gill, F. Rodríguez-Llorente and W. R. C. Rowley, "Frequency stabilised diode lasers in the visible region using Doppler-free iodine spectra", Optics Commun. **132**, 1996, 94-100.

Er64: (Cs D_1) Erikson et al. Archiv for Fysic **28**, 1964, 233ff.

Fi96: ($^{171}Yb^+$) P. T. H. Fisk, M. J. Sellars, M. A. Lawn and C. Coles, "A microwave frequency standard based on trapped, buffer gas-cooled $^{171}Yb^+$ ions - some accuracy issues", Proc. European Frequency Time Forum 5-7 March 1966, (London:IEE) Conf. Publ. 418, 1996, 212-217.

Gi95:($^{172}Yb^+$) P. Gill, H. A. Klein, A. P. Levick, M. Roberts, W. R. C. Rowley and P. Taylor, "Measurement of the $^2S_{1/2}$-$^2D_{5/2}$ 411-nm interval in laser-cooled trapped $^{172}Yb^+$ ions", Phys. Rev. A **52** (2), August 1995, R909-R912.

Go87: (^{24}Mg) A. Godone, E. Bava and C. Novera, "High-resolution frequency measurement of the $^3P_1$-$^3P_0$ ^{24}Mg transition", Metrologia **24**, 1987, 133-138.

Jun95: ($^{127}I_2$) P. A. Jungner, S. Swartz, M. Eickhoff, J. Ye, J. L. Hall and S. Waltman," Absolute frequency of the molecular iodine transition R(56)32-0 near 532 nm", IEEE Trans. Instrum. Meas. **44** (2), April 1995, 151-154 and M. L. Eickhoff and J. L. Hall, "Optical frequency standard at 532 nm", idem, 155-158.

Kni95: (CO) D. J. E. Knight, K. I. Pharaoh and D. A. Humphreys, "Absolute frequency measurement of the 3-0 R(21) transition of CO at 1.5605 μm for optical communication standards", Proc. EFTF'95, Besançon, France, March 8-10 1995, 516-519.

Mad96: (Ba^+, Sr^+) A. A. Madej, K. J. Siemsen, B. G. Whitford, J. E. Bernard and L. Marmet, "Precision absolute frequency measurements with single atoms of Ba^+ and Sr^+", Proc. 5th Symp. on Frequency Standards and Metrology, Ed. J. Bergquist (Singapore: World Scientific) 1996, 165-170, and A. A. Madej and K. J. Siemsen, "Absolute heterodyne frequency measurement of the $^{88}Sr^+$ 455-THz S-D single ion transition", Optics Lett. **21** (11), 1 June 1996, 824-826.

Mar96: (Sr^+) L. Marmet, A. A. Madej, K. J. Siemsen and J. E. Bernard, "Precision frequency measurement of the clock transition in a Sr^+ single atom with an ultra-stable diode laser", CPEM96 Digest, PTB, Germany 17-20 June 1996.

McG78: ($^{199}Hg^+$) M. D. McGuire, R. Petsch and G. Werth, "Precision determination of the ground-state hyperfine separation in $^{199}Hg^+$ using the ion-storage technique", Phys. Rev. A **17** (6), June 1978, 1999-2004.

Nak95: (HCN, C_2H_2, $^{13}C_2H_2$) K. Nakagawa, M. de Labachelerie, Y. Awaji and M. Kourogi, "Precise optical frequency atlas of acetylene lines in the 1.5 μm region", Proc. Laser Spectoscopy XII, Isle of Capri, Italy, 11-16 June 1995, Ed. M. Inguscio, M. Allegrini and A. Sasso (Singapore: World Scientific) 1996, 228-229.

Nak96: (C_2H_2, $^{13}C_2H_2$) K. Nakagawa, M. de Labachelerie, Y. Awaji and M. Kourogi, "Accurate optical frequency atlas of the 1.5 μm bands of acetylene", J. Opt. Soc. Am. **B 13** (12), December 1996, 2708-2714. Note: 36 lines in $^{12}C_2H_2$ and 54 lines in $^{13}C_2H_2$ are listed: the high and low extremes are given.

Nez93: F. Nez, M. D. Plimmer, S. Bourzeix, L. Julien, F. Biraben, R. Felder, Y. Millerioux and P. De Natale, "First pure frequency measurement of an optical transition in atomic hydrogen: better determination of the Rydberg constant", Europhys. Lett. **24** (8), 1993, 635-640.

Nez95: F. Nez, M. D. Plimmer, S. Bourzeix, L. Julien, F. Biraben, R. Felder, Y. Millerioux and P. De Natale, "Determination of the Rydberg constant by direct frequency measurement", IEEE Trans. Instrum. Meas. **44** (2), April 1995, 568-571.

Pa94: (^4He) F. S. Pavone, F. Marin, P. De Natale, M. Inguscio and F. Biraben, "First pure frequency measurement of an optical transition in helium: Lamb shift of the 2^3S_1 metastable level", Phys. Rev. Lett. **73** (1) 4 July 1994, 42-45.

Po83: (^{20}Ne) C. R. Pollock, D. A. Jennings, F. R. Petersen, J. S. Wells, R. E. Drullinger, E. C. Beaty and K. M. Evenson, "Direct frequency measurements of transitions at 520 THz (576 nm) in iodine and 260 THz (1.15 μm) in neon", Optics Lett. **8** (3), March 1983, 133-135.

Po283: (CO) C. R. Pollock, F. R. Petersen, D. A. Jennings and J. S. Wells, "Absolute frequency measurements on the 2-0 band of CO at 2.3 μm; calibration standard frequencies from high resolution color center laser spectroscopy", J. Mol. Spectrosc. **99**, 1983, 357-368.

Poi96: (^{199}Hg$^+$) M. E. Poitzsch, J. C. Bergquist, W. M. Itano and D. J. Wineland, "Cryogenic linear ion trap for accurate spectroscopy", Rev. Sci. Instrum. **67** (1), January 1996, 129-134.

Qu93: (CIPM 1992 revision of "Mise en pratique" of the definition of the metre) T. J. Quinn, "Mise en pratique of the definition of the metre (1992), Metrologia **30**, 1993/94, 523-541.

Ri96: (^{40}Ca) F. Rielhe, H. Schnatz, B. Lipphardt, G. Zinner, T. Trebst, T. Kurosu and J. Helmcke, "An atom interferometer used as optical frequency standard", Workshop on optics and interferometry with atoms, Elba, Italy, 10-12 June 1996, preprint P45 and H. Schnatz, B. Lipphardt, J. Helmke, F. Riehle and G. Zinner, " First phase-coherent frequency measurement of visible radiation", Phys. Rev. Lett. **76** (1), 1 January 1996, 18-21.

Ro97: (^{172}Yb$^+$) M. Roberts, P. Taylor, G. P. Barwood, P. Gill, H. A. Klein and W. R. C. Rowley, "Observation of an electric octupole transition in a single ion", Phys. Rev. Lett., in the press, February 1997.

Sas92: (^{20}Ne) H. Sasada and O. Kubota, "Lamb dip-stabilized 1.53 μm He-Ne lasers", Appl. Phys. **B55** (2), August 1992, 186-188. Note: the accurate lambda-meter measurements are described in:- J. Ishikawa and H. Wantanabe, "A fringe-counting wavemeter for infrared laser diodes, "IEEE Trans. Instrum. Meas., **42** (2), April 1993, 423-425.

Sch95: (muonium, μ^+e^-) W. Schwarz + 24, "Spectroscopy of the 1S-2S energy splitting in muonium", IEEE Trans. Instrum. Meas. **44** (2), April 1995, 505-509.

Se95: (^{171}Yb$^+$) M. J. Sellars, P. T. H. Fisk, M. A. Lawn and C. Coles, "Further investigation of a prototype microwave frequency standard based on trapped ^{171}Yb$^+$ ions", in Proc. 1995 IEEE International Symposium on Frequency Control (49th Annual Symposium) (New York: IEEE) 1995, 66-73.

Tan96: (^{113}Cd$^+$) U. Tanaka, H. Imajo, K. Hayasaka, R. Ohmukai, M. Watanabe and S. Urabe, "Laser microwave double-resonance experiment on trapped ^{113}Cd$^+$ ions", CPEM96 Digest, TH1B-4, 515-6.

Tay97: (^{172}Yb$^+$) P. Taylor et al., to be submitted to Phys. Rev. A, NPL, February 1997.

To97: (^{85}Rb two-photon) D. Touahri, O. Acef, A. Clairon, J.-J. Zondy, R. Felder, L. Hilico, B. de Beauvoir, F. Biraben and F. Nez, "Frequency measurement of the $5S_{1/2}(F=3)$-$5D_{5/2}(F=5)$ two-photon transition in rubidium", Optics Commun. **133**, 1 January 1997, 471-478.

Ud97: (H) Th. Udem, A Huber, M. Weitz, D. Leibfried, W. König, M. Prevedelli, S. Dimitriev, H. Geiger and T. W. Hänsch, "Phase-coherent measurement of the hydrogen 1S-2S frequency with an optical frequency interval divider chain", IEEE Trans. Instrum. Meas., in the press, February 1997.

We96: (Cs D_1) Martin Weitz, private communication, June 1996.

Ye96: (^{87}Rb) J. Ye, S. Swartz, P. Jungner and J. L. Hall, "Hyperfine structure and absolute frequency of the ^{87}Rb $5P_{3/2}$ state", Optics Lett., **21** (16), 15 August 1996, 1280-1282.

Ye296: (C_2HD) J. Ye, L-S. Ma and J. L. Hall, "Ultra-stable optical frequency reference at 1.064 μm using a C_2HD molecular overtone transition", CPEM96 Digest, WE2A-6, Braunschweig, Germany, 17-20 June 1996, Supplement, 19-20.

6. TEXT REFERENCES

1. D. J. E. Knight, "A tabulation of absolute laser-frequency measurements" Metrologia, **22**, 1986, 251-257.

2. T. J. Quinn, "*Mise en pratique* of the definition of the metre (1992)", Metrologia, **30**, 1993/94, 523-541.

3. A. Clairon, S. Ghezali, G. Santarelli, Ph. Laurent, E. Simon, S. Lea, M. Bahoura, S. Weiers and K. Szymaniec, "The LPTF preliminary accuracy evaluation of caesium fountain frequency standard", Proc. European Frequency and Time Forum, 5-7 March 1996, (London:IEE, Conference Publication no. 418) 1996, 218-223.

4. J. Vanier and C. Audoin, *The quantum physics of atomic frequency standards* (Bristol: Adam Hilger) 1989, Chapter 1.

5. A. G. Maki and J. S. Wells, *Wavenumber calibration tables from heterodyne frequency measurement.* (Washington D.C.: US Government Printing Office, NIST Special Publication 821) 1991.

6. A. G. Maki, "Infrared and far-infrared absorption frequency standards", in D. R. Lide and H. P. R. Frederikse, Eds., *CRC Handbook of Chemistry and Physics*, 75th Edition (Boca-Raton: CRC Press) 1994, pp **10**-319 - **10**-323.

7. D. J. E. Knight, "Requirements and practical possibilities for frequency standards for optical fiber communication bands", in M. Ohtsu, Ed., *Frequency control of semiconductor lasers* (New York: John Wiley and Sons) 1996, 173-206.

8. J. C. Bergquist, W. M. Itano and D. J. Wineland, "Laser stabilization to a single ion", in T. W. Hansch and M. Inguscio, Eds., *Frontiers in laser spectroscopy* (Amsterdam: North Holland) 1992, 359-376.

9. O. Acef, "Metrological properties of CO_2/OsO_4 optical frequency standard", Optics Commun. **134**, 15 January 1997, 479-486.

10. V. Bernard, C. Daussy, G. Nogues, L. Constantin, P. E. Durand, A. Amy-Klein, A. Van Lerberghe and Ch. Chardonnet, "CO_2 laser stabilization to 0.1 Hz level with external electrooptic modulations", preprint, received March 1997.

A noise contribution analysis of a Light Induced Microwave Oscillator including optical gain

S. Römisch and A. De Marchi

Dipartimento di Elettronica, Politecnico di Torino
e-mail romisch@polsik.polito.it

Abstract

In 1995 L. Maleki proposed an electrooptic oscillator made of a loop with a delay line realized by a simple optical fiber.
A new version of the same idea is illustrated in this paper. It includes a DFB laser and an Er^{+}-doped optical amplifier.
We found a phase noise power spectral density that shows f^{-2} and f^{-3} contributions instead of the simple f^{-2} measured and calculated by Maleki.
In consequence of this we made a more accurate analysis of all noise sources of the loop, with special attention to the optical components.
In fact optical FM noise (like laser noise or optical amplifier noise) can make FM or PM contributions to electrical signal noise.
A prototype of it was realized and characterized.

1 Introduction

In recent years a new kind of oscillator based on a hybrid electro-optical approach was presented. These Light Induced Microwave Oscillators (LIMO) use a delay line in the loop and in order to obtain the long delay times which are necessary for frequency discrimination, optical fibers with their low losses are used.
Experimental results [1] show noise performances comparable with DRO and quartz oscillators without working at cryogenic temperatures, and with a phase noise that seems to be indipendent from the oscillator frequency.
The presence of the fiber as a selective element allows high tunability: band limits come from devices that couple optical signals with electrical ones and they are from hundreds of kHz up to tens of GHz.
The translation between RF signal and amplitude modulation of the light can be done by an electrooptic modulator or by direct modulation of a semiconductor laser current; the conversion back into electric

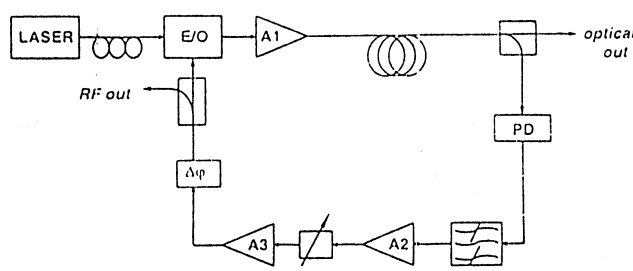


Figure 1: General experimental setup for LIMO

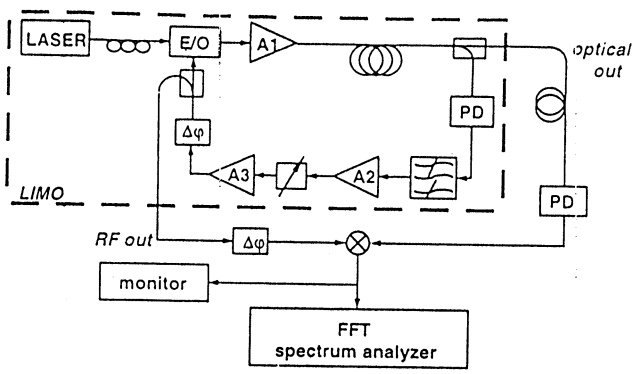


Figure 2: Measurement setup for oscillator's noise

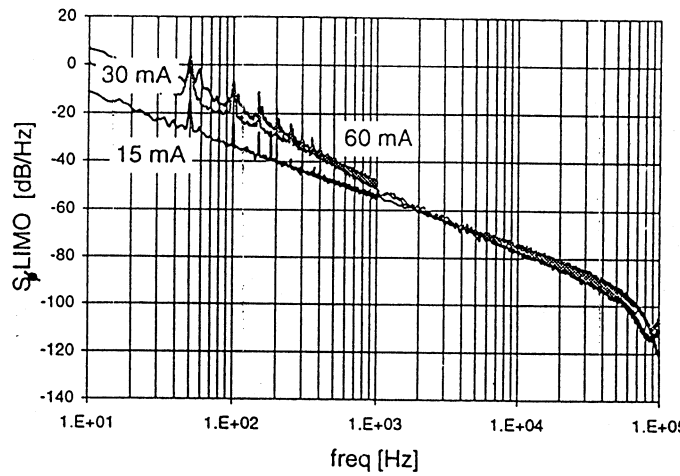


Figure 3: Phase noise of LIMO with optical amplifier and a 100 m. of loop fiber. DFB threshold current: 11.8 mA (ASTROTEC).

signal is made by a photodiode.

2 The LIMO

The experimental setup of the oscillator is shown in fig.1. A DFB laser in the 1.55 μm. band feeds a Mach-Zehnder electro-optic modulator, then the light continues through an optical amplifier followed by the fiber delay line and is converted into an RF signal by a photodiode.
The RF signal is then filtered, amplified and sent to the modulator's RF input to close the loop. A $\frac{\lambda}{4}$ polarization control is used to make amplitude modulator.
The microwave filter allows to choose among modes of the loop that are separated by tens or hundreds of kHz. In order to obtain this kind of selection a filter with a Q factor of about 1000 is sufficient.
Finally, for the fine tuning of oscillating frequency, there is an electrical delay line working as a phase shifter. A microwave amplifier with high gain is needed to reach unity loop gain.
The use of a DFB laser with a power level around 0 dBm, cause the presence of the optical amplifier trying to recover some power before downconversion to RF domain.

3 Oscillator's noise

To characterize the oscillator without a reference signal we use the delay line method [4], which measures frequency noise:

$$S_V(f) \propto (\tau_d \nu_0)^2 S_y(f) \quad for \quad f \leq \frac{0.7}{\pi \tau_d}$$

and we obtain phase noise by simple calculations:

$$S_\phi(f) = \left(\frac{\nu_0}{f}\right)^2 S_y(f)$$

A delay line 2.2 km long keeps the measurements valid up to 20 kHz.
The measurement setup is shown in fig.2: the delay line is an optical fiber 2.2 km long that takes via a coupler a fraction (about 30 %) of the optical power of the oscillator. By a photodiode we obtain the reference signal that is compared with a fraction of the RF oscillating power, taken by a directional coupler just before the electrical input of the modulator. In fact, at that point of the loop there is the maximum available RF power (about 20 dBm).
Comparison of the two signals is made by a double balanced mixer working as a phase detector.
The power spectrum $S_V(f)$ is then measured with an FFT dynamic signal analyzer.
We measured the frequency noise of the LIMO with different bias currents of the laser and we found changes of slope in $S_\phi(f)$: by increasing the current we saw more frequency flicker noise.
Possible reasons for this behavior: different lineshapes of the laser and/or increasing noise level of the optical amplifier due to working in saturation with higher power levels of the laser.
In order to understand better we made different measurements.

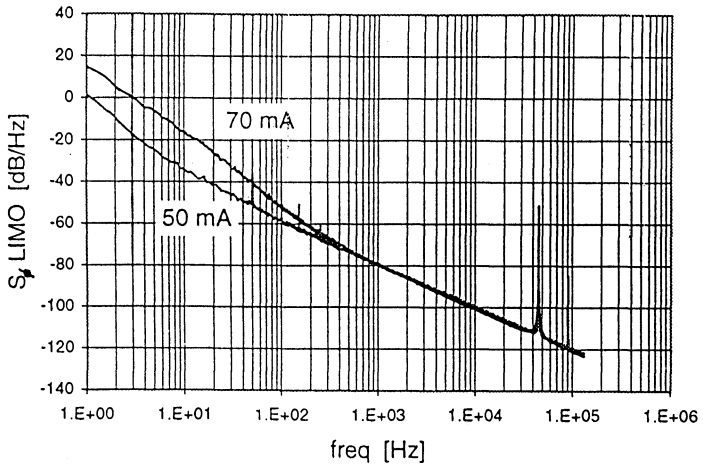


Figure 4: Phase noise of LIMO without optical amplifier and a 4.4 km. loop fiber. DFB threshold current:30 mA (ASTROTEC).

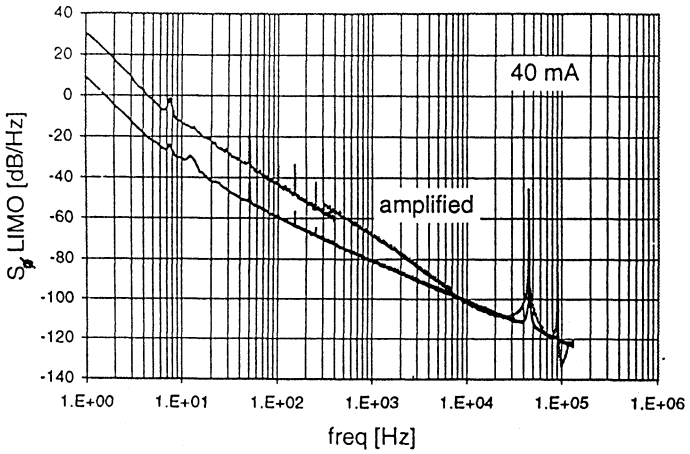


Figure 5: Phase noise comparison between LIMOs with and without optical amplifier and a 4.4 km. fiber. DFB threshold current: 12 mA (MITSUBISHI).

In fig.3 the phase noise of a LIMO with a power optical amplifier and a fiber 100 m. long in the loop is shown. The DFB laser used has a threshold current of 11.8 mA (ASTROTEC).
In fig.4 noise of a different setup with longer loop fiber (4.4 km.) and without the optical amplifier is shown.In that case a DFB laser with threshold current of 21 mA is used (ASTROTEC).
In fig.5 the comparison between measured noise of setups with and without optical amplifier is shown , keeping the same bias current for the same laser (MITSUBISHI).

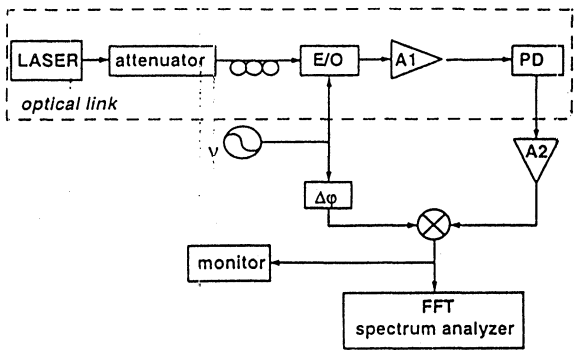


Figure 6: Experimental setup for noise measurements on optical link

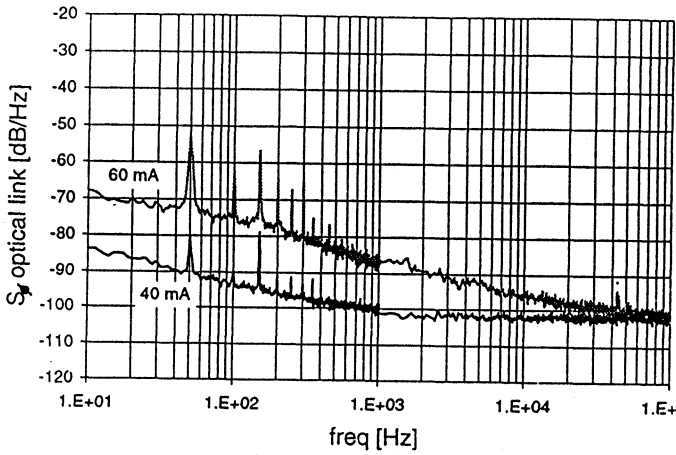


Figure 7: Optical link's phase noise without optical amplifier.DFB threshold current: 21 mA (ASTROTEC).

4 The optical link

The general setup used for all measurements of phase noise on optical link is shown in fig. 6. This kind of measurements implies absence of the optical fiber into the optical link to avoid measurement of source phase noise.
The modulator is biased in the middle of its transmission characteristic and the microwave power at its RF input is about 20 dBm.
The first group of measurements is without optical amplifier for different bias currents of the laser, and we obtain the curves shown in fig.7.
For the second group of measurements, we added the optical amplifier. Results of comparisons between the

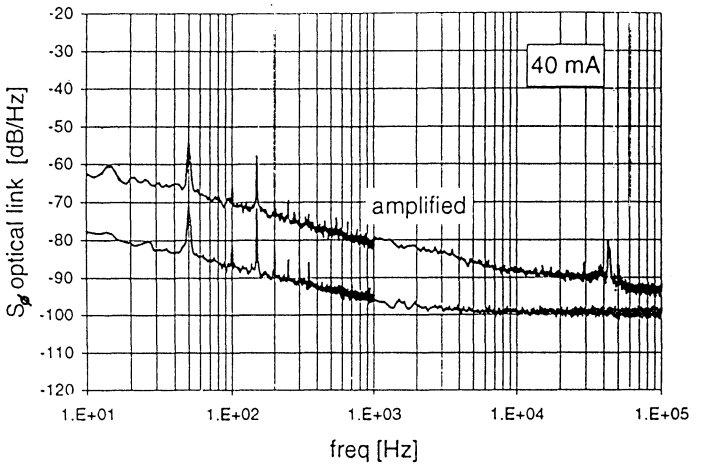


Figure 8: Phase noise comparison between optical links with and without optical amplifier. DFB threshold current: 21 mA (ASTROTEC)

two cases with the same current are shown in fig.8. To verify that different optical power levels on the photodiode cannot make any significant difference, we made measurements with insertion of an optical attenuator that keeps the same quantity of light on the photodiode for different laser currents. Results of these measurements don't differ from these shown in this paper.

5 Conclusions

A preliminary analysis of influence of FM optical noise on microwave phase noise in electro-optical oscillators has been made.
The noise contribution of the optical amplifier makes its presence in the loop not positive.
Comparisons between different lasers and different bias currents indicate the necessity of further investigation about lineshape of lasers used to build this oscillator.

Acknowledgments

The authors wish to acknowledge the cooperation of the Centro Studi per l'Elettronica delle Telecomunicazioni (CSELT) and the helpful discussions with many scientist of the place, in particular R. Girardi.

References

[1] X.S.Yao and L. Maleki,"Light Induced Microwave Oscillator", TDA Progress Report 42-123, Jet Propulsion Laboratory, pp. 32-42 (1995) http://tda.jpl.nasa.gov/progress_report

[2] R.Logan, L.Maleki and M.Shadaram, "Stabilization of oscillator phase using a fiber optic delay-line", Proceedings of 45th Annual Frequency Control Symposium, May 29,31, Los Angeles, CA, 1991

[3] S.Yao and L.Maleki, "Characteristics and performance of a novel photonic oscillator", Proceedings of the 49th Annual Frequency Control Symposium, 1995

[4] Hewlett-Packard Company"Phase noise characterization of microwave oscillators-frequency discriminator method", Product Note 11729C-2

Piezoelectric effect in coaxial cables

E. Rubiola† M. Bruno‡ G. Lorino† A. De Marchi†

†Politecnico di Torino ‡SPIN Electronic Instruments

Abstract

This paper reports preliminary results of some experiments in which the effects of vibrations on cables, including both piezoelectricity and triboelectricity, are measured. Cables with different shapes and dielectric materials are tested.

1 Introduction

The next generation of supercloks, as well as some applications of precision electronics, require an improvement of knowledge on the behaviour and the stability of basic components, taking into account effects that are traditionally neglected. Some recent papers, such as [1], deal with the stability problem. Working in this direction, we made an experimental study of the piezoelectric effect in coaxial cables, which could affect the stability of many circuits.

PTFE coaxial cables are widely used when high stability is to be achieved. PTFE is an excellent dielectric from many points of view: it is suitable for high frequency applications in the microwave range because of its low dielectric losses, it is chemically stable, and it shows extremely high DC resistivity, which can be 10^{16} Ωm or more. Other dielectric materials are also frequently used, mainly PE and PE foam, and sometimes alumina.

PTFE shows both piezoelectric and triboelectric effects, because of the strong distortion of the molecular electric field due to the presence of fluorine atoms. This means that a PTFE cable converts stress and vibrations into electric signals that can perturb electronic circuits either directly or by modulating some parameter. Other dielectrics show piezoelectricity and triboelectricity, although the effect in our experience is less relevant than that of PTFE. In order to give a practical idea of the relevance of the phenomena here considered, we made a simple experiment in which a small mass falls with a kinetic energy of about 1 mJ on a 0.5 m PTFE UT-85 semirigid cable connected at the input of an oscilloscope. This shock induces a 300 mV peak to peak signal, as shown in Fig. 1.

Although people from the DC precision measurement community are aware of these phenomena [2], we found it only superficially and qualitatively mentioned in the literature.

The main purpose of this study is to give some practical hints to be used as guidelines for designing precision electronics and clocks. We have measured a number of cables (see Tab. I) of different origin, including some from old instruments. In the following sections the measurement method will be illustrated and the experimental results will be discussed.

2 Measurement method

Piezoelectricity is formally described by the costitutive equations which couple the mechanical stress T

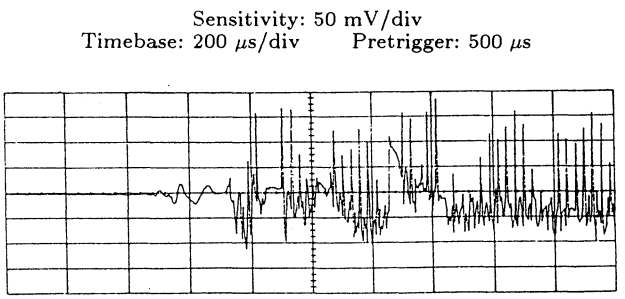


Figure 1: electric effect of a shock on a UT-85 cable.

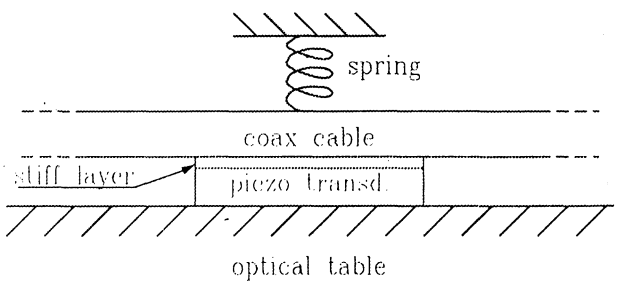


Figure 2: mechanical arrangement of the test cell.

TABLE I

cable no.	military standard	general properties	dielectric	Z Ω	cap. pF	len. m	diam. mm	factory and type
1	RG 142 B/U	flexible	PTFE	50	181.0	1.81	4.6	Thermatics inc.
2	RG 188 A	flexible	PTFE	50	23.5	0.14	2.5	
3		rigid tin coat.	alumina	50	5.0	0.13	3.6	Uniform Tubes
4		rigid tin coat.	alumina	50	27.9	0.20	2.3	Uniform Tubes
5		flexible	PTFE	50	76.2	1.06	3.5	
6	RG 405	semirigid	PTFE	50	41.9	0.34	2.2	
7	RG 402	semirigid	PTFE	50	56.2	0.50	3.6	
8	RG 402	semirigid	PTFE	50	49.5	0.46	3.6	
9	RG 405 like	semir. tin coat.	PTFE	50	43.9	0.38	2.2	
10	RG 401	semirigid	PTFE	50	55.2	0.53	6.4	
11		flexible	PE foam	75	73.4	1.31	6.6	ITC art.333BS
12		flexible	PE foam	75	68.2	0.53	6.6	Cavel SAT703B
13	RG 58 C/U	flexible	PE	50	45.9	0.43	4.9	Suhner
14		flexible	PE	75	35.5	0.50	5.8	Siemens
15	RG 213 U	flexible	PE	50	103.8	0.99	0.0	FMC
16	RG 8/U	flexible	PE	50	51.4	0.46	0.2	Amphenol
17	RG 58	flexible	PE	50	155.6	1.53	4.8	Pirelli
18	RG 58/U	flexible	PE	50	135.1	1.36	4.8	Amphenol

and strain S to the electric vectors E and D

$$T = cS - eE$$
$$D = eS + \epsilon E$$

(vector and matrix subscripts are omitted) through the elastic stiffness c, the piezoelectric constants e and the permettivity ϵ, as described in [3].

Unfortunately, the piezoelectric constants are not so useful to describe the behaviour of a cable in a circuit. For instance, in a typical situation we have many coax cables linking the parts of a system: each cable follows its own bended path, which is not well defined; the mechanical coupling of the cables to the environment are not sufficiently defined as regards the vibration problem and the mechanical resonances of the cables are not known at all; finally, in our case piezoelectricity can not be divided from triboelectricity. This last effect is due to electric charges generated by rubbing a dielectric surface; unfortunately, we haven't been able to find in the literature any specification for polymers like PE or PTFE. Since in the described scenario we can only have a basic knowledge of the vibration spectrum, either formal or numerical analysis are out of reach in any reasonable resource budget of a project. Finally, we got to the conclusion that the electric phenomena and the mechanical coupling of the cable to the environment must be considered together. Consequently, we have identified a typical situation, intended to be representative of many practical cases, in which the cable is mounted on a vibrating plate; between the cable and the plate there is one or more stiff layer, i.e., enamel, glue, the insulation cladding of the cable itself, or anything else. Then, we have designed a mechanical arrangement for a test cell imitating this situation, as shown in Fig. 2.

The displacement x produced by the piezo transducer has been measured with the scheme of Fig. 3. The effect of the signal sent to the transducer is a parallel AC translation of the laser beam, which is measured by the split detector and compared to the DC translation manually imposed through the micrometer; the detectors and its electronics are described in [4]. The response of the transducer is roughly flat, $x_{\text{rms}} \simeq 8.8$ μm for a driving voltage of 8.5 V, in the 100 Hz to 100 kHz region.

The complete experimental setup is shown in Fig. 4. The computer sets the drving frequency scanning the 100 Hz - 100 kHz range and reads the piezoelectric signal through a lock-in amplifier which has a bandwidth of 10 mHz or 1 Hz, depending on the needs. The transresistance amplifier shows negligible input impedance (virtual ground), which shorts the cable; moreover, there is no external electric field applied to the cable ($E = 0$). Hence the amplfier output voltage V is proportional to the piezoelectric current, which is the time derivative of the electric displacement D integrated over the *stimulated region* of the cable. This scheme makes V independent on the cable capacitance, which is proportional to the cable length.

In principle, our experimental arrangement can't distinguish between the piezoelectric and the triboelectric effects. Therefore, we believe that this latter is less relevant because the mechanical stimulus

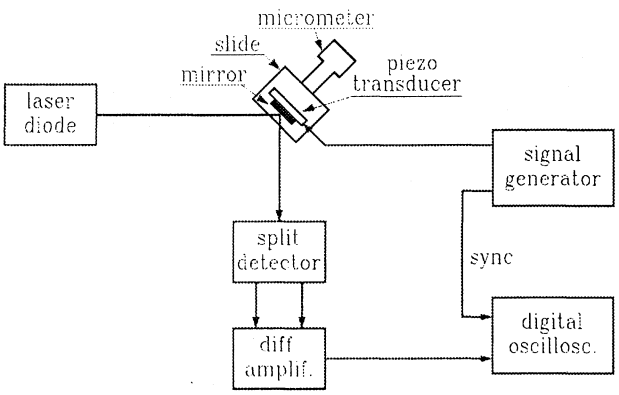


Figure 3: calibration of the piezo transducer.

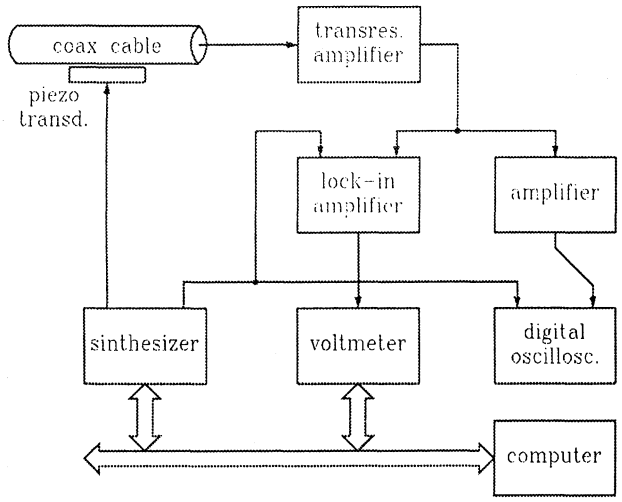


Figure 4: Experimental setup for the cables piezoelectricity measurement.

is orthogonal to the cable surface; in this condition, rubbing the dielectric surface is a second order effect.

Some measurement runs have been made after setting the synthesizer output to zero amplitude. In this condition the measured signal represents the system noise, both electric and acoustic.

3 Experimental results

The available cables were measured with the above described equipment. Using these data, we obtained a series of plots in which the piezoelectric current is reported as a function of the vibrating frequency. Fig. 5 reports three examples of cable responses measured with an equivalent bandwidth of 1 Hz and a frequency resolution of 300 points per decade; in the left part of these plots the equipment noise, which is about 30 pA at 100 Hz with slope f^{-1}, is evident.

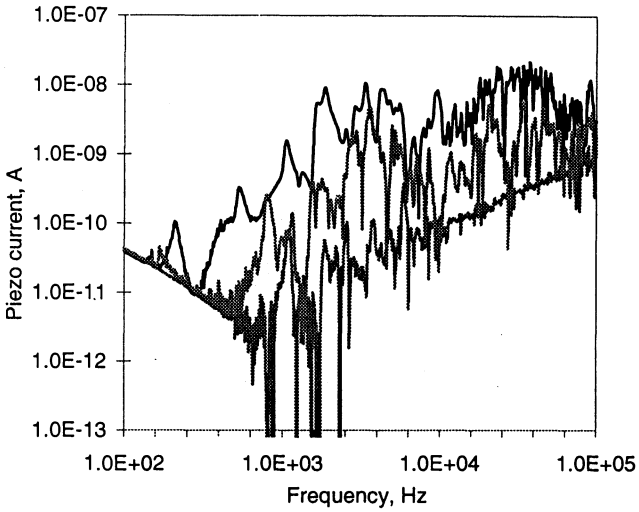


Figure 5: examples of the measured piezoelectric behaviour: cable #2 (top), #5 (middle) and #4 (bottom).

The response plots are rather uneven becase of the presence of many resonant modes. Therefore, fitting the experimental data with integer-slope straight lines we recognize two types of asymptotic behaviour:

1. slope 1 (see cable 4, Fig. 5),
2. slopes 2 and 0 (see cables 2 and 5, Fig. 5).

The results of the straight line fitting are reported in Fig. 6. This picture should be considered carefully because of the complexity of the reported information. Most of measurements (solid lines) are taken with a frequency resolution of 300 points per decade and an equivalent bandwidth of 1 Hz. Others (dashed lines) are taken with 5–10 points per decade and a bandwidth of 10 mHz. These latter ones were taken after realizing that the poor S/N made doubtful the previous results (1 Hz bw).

Observing the results of Fig. 6, it turns out that most of semirigid and rigid cables show type 1 behaviour, while type 2 is rather frequent in the flexible ones. The physical origin of the experimentally observed slopes will be discussed below.

Let's first consider the semirigid and rigid cables, which show a +1 slope. The displacement x of the unloaded piezo transducer is independent of frequency, as pointed out in Section 2 ($x_{\rm rms} \simeq 8.8\ \mu$m).

At high frequencies the transverse string mode can't be excited. Instead, we expect that the piezo transducer excites volume vibrations inside the cable. In these conditions, the inertial mass turns the acceleration into a force $F \propto {\rm d}^2x/{\rm d}t^2$, which is proportional to f^2. Since we measure the current $I = {\rm d}Q/{\rm d}t$, we would expect $I \propto f^3$. However, a slope +3 has never been observed. We believe that

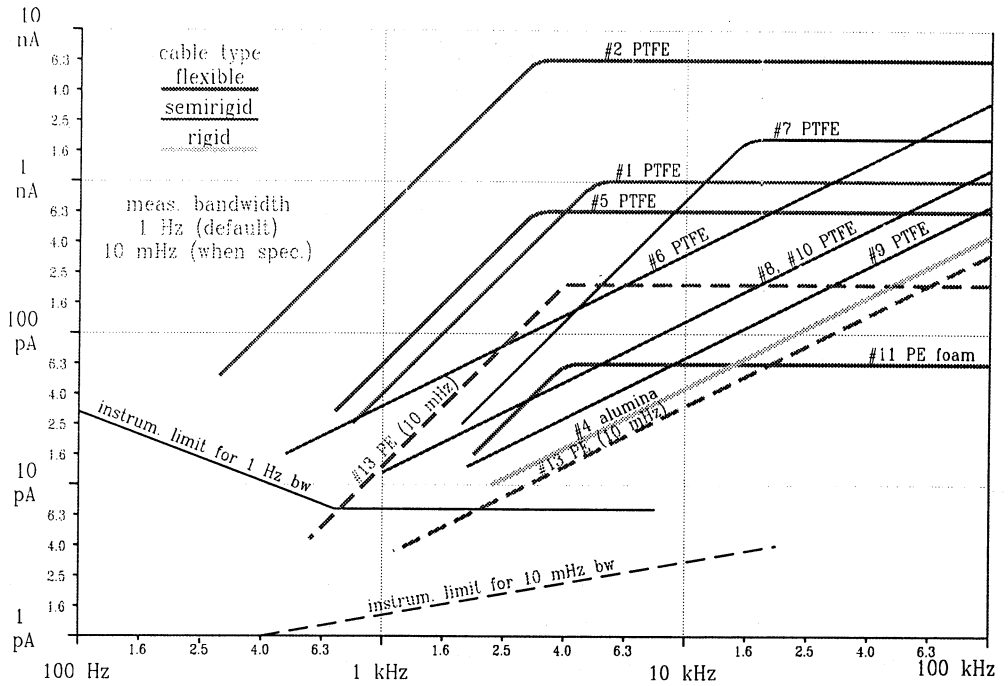


Figure 6: asymptotic straight-line approximation of the measurement results.

this is due to the enamel coating of the piezo transducer, which is a thin stiff layer (see Fig. 2). The mechanical coupling of the piezo transducer to the cable is equivament to a mass-spring damped resonator that acts as a second order low pass filter with slope -2. This hypothesis explains the observed slope, which is still $+1$.

The other type of asymptotic behaviour is not easy to explain. Its modelization probably requires a more complicated analysis of the mechanical system.

4 Conclusions

Experimental evidence of piezoelectric effects in commonly used coaxial cables was collected.

It was shown that care must be taken in selecting coaxial cables when extra low noise measurements are being planned.

References

[1] L. M. Nelson, F. L. Walls, "Environmental effects in mixers and frequency distribution systems", *Proc. 1992 IEEE Frequency Control Symposium* pp. 831–837, Hershey (Pennsylvania, USA), 27–29 May 1992.

[2] *Low level measurements*, 4th edition, ©Keithley Instruments, Inc, 1992.

[3] "IEEE Standard on Piezoelectricity", ANSI / IEEE Std. 176-1987 (under revision). Reproduced in *IEEE Transactions on Ultrasonic, Ferroelectrics, and Frequency Control*, vol. 43 no. 5, sept. 1996.

[4] A. De Marchi, M. Ortolano, F. Periale, E. Rubiola "A measurement of the period stability of a free pendulum in vacuum", *Proc. 50th Annual Frequency Control Symposium* pp. 665–673, Honolulu (Hawaii, USA), 5–7 Jun 1996.

Two dimensional Raman velocity selection and prospects for laser cooled atomic fountain frequency standards.

S. Ghezali, K. Szymaniec, *L. Cognet, **J. Reichel, S. Bize and A. Clairon.

BNM/LPTF, Observatoire de Paris, 61 Avenue de l'Observatoire. Paris 75014. France.
*: IOTA, ESO, Université Paris 11. Orsay. France.
**: LKB, ENS, 24 Rue Lhomond. Paris 75005. France.

Abstract :

Colds collisions constitute a crucial effect at high densities in a laser cooled atomic fountain [1]. That is because the collisional cross sections become sufficiently large at the μK temperature regime to induce a relative spin exchange frequency shift of the order of 10^{-15} measured against a Hydrogen maser reference [2].
On the other hand, because of the thermal expansion of the atomic cloud, only a fraction of the atoms, few percent, passes a second time through the 1cm microwave cavity aperture and undergoes a second Ramsey interaction.
Whereas all the atoms contribute to the collisional shift, the major part of them are lost for the detection signal and can be eliminated from the fountain with transverse velocity selective pulses via Raman transitions [3].
We will present our experimental set-up using an electro-optical modulator to generate the two Raman beams with a frequency difference of 9.19...GHz and the 2D velocity selection results. We will discuss the possibility of applying a Raman cooling to reduce the transverse temperature, the technique being limited by the time that launched atoms spend in the cooling region.
The potential improvements for atomic fountain frequency standards will be given.

References :

[1] : A. Clairon et al, IEEE Trans. Instrum. Meas., Vol. **44**, 128, April 1995.
[2] : S. Ghezali et al, EuroPhys. Lett., Vol. **36**, 31, October 1996.
[3] : S.Lea et al, Phys. Scr., Vol. **T51**, 78, 1994.

WORKSHOP

NEW GENERATION OF SPACE CLOCKS

SESSION 1

Chairman: G. Busca

HIGH PRECISION FREQUENCY AND TIME TRANSFER IN SPACE

ROBERT F. C. VESSOT AND EDWARD M. MATTISON

SMITHSONIAN ASTROPHYSICAL OBSERVATORY
CAMBRIDGE MASSACHUSETTS 02138, USA

1. ABSTRACT

A review of techniques available for making high precision time and frequency transfer between high stability atomic frequency standards. Methods for cancelling of first-order Doppler shifts owing to relative motion and for removing propagation disturbances in coherent microwave systems for frequency comparisons and in phase-modulated microwave systems for time transfer. Newly developed high precision timing equipment for operating pulsed laser time transfer from space to earth will be described.

Keywords: high precision time transfer, atomic clocks

2. INTRODUCTION

One of the most significant scientific improvements in our era is the uparalleled precision of measurement made possible by atomic clocks. In making frequency and time transfer. Use of electromagnetic signals (micro wave and laserlight) for high precision range, range-rate and angle measurements requires us to include appropriate corrections for gravitational and relativistic effects.

This discussion is meant to provide a picture of where we stand in the present state of techniques for clock-related space experiments involving high precision frequency and time transfer. Please note that some of the figures referred to appear in the EFTF paper "Applications of Highly Stable Clocks in Space" (AHSCS)

3. PRESENT STATE OF CLOCK TECHNOLOGY AND SYSTEMS FOR SPACE EXPERMENTS

The most commonly used definition of frequency stability is the Allan (standard) deviation, $\sigma_y(\tau)$. This is the one-sigma expectation of the fractional frequency difference $\Delta f/f$ (designated by the subscript y), between time-adjacent frequency measurements, each made over time intervals of duration, τ. The functional relationship of $\sigma_y(\tau)$ versus (τ) depends on the Fourier spectrum of of the phase variations.[1]

An estimate of the time dispersion of a clock or oscillator for a future time interval, τ, can be obtained from the relation

$$\sigma_{\Delta\tau}(\tau) \sim \tau\,\sigma_y(\tau) \qquad (1)$$

Representation of oscillator performance and of propagation disturbances by $s_y(t)$ makes possible statistical estimates of the limits imposed by clock performance on measurements made with electromagnetic signals.

Figure 1 shows $\sigma(\tau)$ versus τ plots for stored ion devices[2], atomic hydrogen masers[3], and for the binary pulsar[4]. This figure also includes the Allan deviation of disturbances caused by the earth's troposphere and ionosphere on S-Band (2GHz) signals traversing vertically. In the following discussions of experimental techniques, the H maser performance data in this figure will be used as a basis for numerical examples.

4. EFFECTS OF OSCILLATOR INSTABILITY ON MEASUREMENTS OF DISTANCE, AND RANGE-RATE USING DOPPLER DATA

In the case of distance measurements made from the one-way propagation of light we can obtain an estimate of range dispersion from time dispersion given in Equation 1, by writing

$$\sigma_{\Delta r}(\tau) = c\tau\,\sigma_y(\tau) \qquad (2)$$

where c is the velocity of light.

The one-way Doppler frequency shift of signals from an oscillator transmitting at a frequency, f, moving with velocity v_r toward the receiver is

$$\frac{\Delta f}{f} = \frac{v_r}{c}$$

The contribution of the oscillator to the imprecision of determining range rate, v_r, during measurement intervals, τ, is given by

$$\sigma_{Vr}(\tau) = c\,\sigma_y(\tau) \qquad (3)$$

Figure 2 is a nomograph of range-rate error and range distance error based on the H-maser data in Figure 1. On the left hand axes are scales for $\sigma_y(\tau)$ and the corresponding one-way range-rate measurement error, $\sigma_{Vr}(\tau)$. On the right hand axes are the scales for time dispersion, $\sigma_{\Delta\tau}(\tau)$, and the corresponding one-way range measurement error.

5. EFFECTS OF OSCILLATOR INSTABILITY ON MEASUREMENTS OF ANGLE USING VLBI

High precision measurement of the angle between the propagation vector of a signal and the direction of a baseline, defined as the line between the phase centers of two widely separated radio telescope antennas, can be made with VLBI techniques[5] shown in Figure 6 AHSCS Here, two radio telescopes, separated by a distance, L, each detect the arrival of radio noise signals from a distant radio star. After heterodyning to a lower frequency the noise signals are recorded as a function of time and the two sets of noise data are subsequently time-correlated. The observable quantities from the correlation process are the correlated amplitude and the relative phase of the signals detected at the widely separated points on the wavefront. VLBI measurements have been used in light deflection tests of relativistic gravitation.[6]

The stability limit on the successive measurements of angle imposed by the oscillator instability on successive measurements of angle taken τ seconds apart is

$$\sigma_{\Delta\theta}(\tau) \sim \frac{c\,\tau\,\sigma_y(\tau)}{L \sin\theta}, \qquad (4)$$

where θ is the angle between the propagation vector and the baseline. The result of correlating the noise data obtained from a common source by the two stations is the production of fringes analogous to those observed from two-slit optical diffraction. The spacing between the fringes is λ/L Sinθ, where λ is the average wavelength of the signals arriving at the antennas. The visibility of the fringes depends on the extent to which the signals arriving at the antennas are correlated.

The angular resolution of the interferometer is given by the change of fringe phase, φ, with θ

$$\frac{d\phi}{d\theta} = \frac{2\pi L}{\lambda} \qquad (5)$$

The error in successive angular measurements owing to the instability of the clocks in a terrestrial system with L = 6000 km, assuming σ_y (10^3 sec) = 1 x 10^{-15}, and θ = π/2, is given by

$\sigma_{\Delta\theta}$ (10^3 sec) = 5 x 10^{-11} rad or 2 μ arcsec.

This is far smaller than the present actual limit of 100 μ arcsec level from terrestrial stations with an 8000-km baseline operating at 7 mm wavelength.[7] The effects of tropospheric and ionospheric fluctuations impose limits that are far more serious than clock instability.

By operating VLBI stations in space, limits in angular resolution owing to tropospheric and ionospheric propagation and baseline distance, imposed by the size of the earth, can be overcome .

As an example of the limits that a spaceborne two station system could achieve, let us consider a spaceborne system where L = 5 x10^6 km, σ_y (10^4 sec) = 4 x 10^{-16}, and θ = π/2. In this case,

$\sigma_{\Delta\theta}$ (10^4) = 2 x 10^{-13} rad or 0.05 μarcsec .

For λ = 1 mm we have λ/L = 2 x 10^{-13} rad and we see that the limit imposed by clock stability with 10^4 sec integration time is capable of resolving fringes at 1-mm wavelengths in a spaceborne system with baseline distances of 5 x 10^6 km.

6. SYSTEMS FOR CANCELLING FIRST-ORDER DOPPLER AND SIGNAL PROPAGATION EFFECTS

The one-way and two-way Doppler measurement technique was used in a three-link Doppler-cancelling system developed for the 1976 SAO/NASA Gravity Probe A test of the gravitational redshift.[8,9] This "Doppler cancellation" scheme was pivotal to the success of the experiment By measuring the Doppler effects in a *separate* two-way system and subtracting one-half the number of the two-way cycles from the phase of the received signal from the phase of the one-way microwave link connecting the space vehicle clock to the earth station, the propagation effects were *systematically* removed.

Figure 1 AHSCS describes the phase-coherent analog system that was used in the 1976 SAO-NASA test of the gravitational redshift. The fractional output frequency is

$$\frac{f_S - f_e}{f_0} = \frac{(\phi_S - \phi_e)}{c^2} - \frac{|\vec{v}_e - \vec{v}_s|^2}{2c^2} - \frac{\vec{r}_{se} \cdot \vec{a}_e}{c^2} \qquad (6)$$

where f_0 is the clock downlink frequency. The term $(\phi_S - \phi_e)$ is the Newtonian potential difference between the spacecraft and earth station, $\vec{v}_e$ and $\vec{v}_s$ are the velocities of the earth station and the spacecraft, $\vec{r}_{se}$ is the vector distance between the spacecraft and earth station, and $\vec{a}_e$ is the acceleration of the earth station in an inertial frame.

The first term is the gravitational redshift resulting from the difference in the Newtonian gravitational potential between the two clocks, the second term is the second-order Doppler effect of special relativity, and the third term is the result of the acceleration of the earth station during the light time, r/c, owing to the earth's rotation. During the two-hour near-vertical flight, the first-order Doppler shifts were as large as ±2 x 10^{-5} and the noise from ionospheric and tropospheric propagation effects was about 1 x 10^{-12} at τ ~100 sec, as shown in the top left curve of Figure 1. After the frequency variations predicted in Equation 6 were fitted to the data, the error in the fit of the data was within (+2.5 ± 70) x 10^{-6} of Einstein's prediction.[10] The residuals were analyzed after subtracting the predicted frequency variation over the time of the mission. The resulting Allan standard deviation of the frequency residuals is shown in Figure 3. Here we see that the stability of the frequency comparison made through the three-link system over signal paths of 10,000 km, in the presence of Doppler shifts of magnitude ± 44kHz, plus the ionospheric and tropospheric noise shown in Figure 1, is comparable to the frequency comparison made between the two reference masers in the same room, reaching 6 x 10^{-15} stability at about 10^3 sec.

Figure 1 AHSCS shows how phase coherence throughout the system was provided by ratio frequency synthesizers. As there was a considerable difference in the S-band frequencies, owing to the transponder's turn-around frequency ratio, 240/221, the ionospheric electron content in the signal path, could have caused serious problems from phase delay in the three microwave links[11] This error was removed by choosing the frequency ratios so that the ionospheric effects in the three signal paths were cancelled at the output of mixer M3. [12,13]

6.1 A Symmetrical Four-Link System to Provide Time-Correlated Doppler Data

By providing a transponded signal back to the spacecraft, as shown in Figure 2 AHSCS, we can obtain one-way, two-way, and Doppler-cancelled data at both stations of the system in terms of the proper time scale kept by the station's clock. When the light time between stations is long compared, for example, to the intervals required for measurements, a dominant, spatially localized noise process can be cancelled *systematically* by time

correlating the data obtained from the two stations. Figure 3 AHSCS shows the space-time paths of the four signals in Figure 2 AHSCS. Here the dots signify the clocks, and the arrows, $E_1(t)$ and $E_2(t)$, signify signal outputs representing earth-based one-and two-way data at a particular epoch in the continuum. $S_1(t)$ and $S_2(t)$ represent one- and two-way data recorded in space. By time-correlating the Doppler responses we can systematically cancel a strong localized noise source such as from the earth's troposphere and ionosphere.[14,15] We see that the iono-tropo noise pattern received from the the spacecraft transmitted at time t_i and received at earth at time t_i+R/c is the same as the noise from earth and received at the spacecraft at time (t_i+2R/c). By advancing $E_1(t)$ by time R/c with respect to $S_1(t)$ and subtracting the two data sets we can systematically remove the noise in the $S_1(t_i+2R/c) - E_1(t_i+R/c)$ combined data set at the small expense of increasing the random noise in the data by $\sqrt{2}$. In situations where the localized dominant noise is substantially larger than the nonlocalized random noise, this process can be highly effective.

Note that in the case where there is a substantial light-time delay between earth and space, the Doppler cancellation, made by subtracting one-half of the two-way Doppler cycles from the one-way cycles, will be very effective in the spacecraft system. Here, the signal transponded from earth at time $t_{i+R/c}$ travels through the same noise-producing Iono-Tropo disturbances as the one-way earth-to-space signal, received in space at time t_i+2R/c. However, this is not the case for two-way signals transmitted to space and transponded back to earth. For tracking a close-approach solar probe, where there will be about 10^3 seconds of go-return transponder delay , the Iono-Tropo conditions could change and cause problems in the Doppler cancelled data.

7. CARRIER MODULATED SYSTEMS FOR MICROWAVE TIME TRANSFER

At SAO, during the early 1980s, we were involved in a project called STIFT (Satellite Time and Frequency Transfer) [16 17 18]. The objective was to compare both frequency and time information from the NASA shuttle using the same basic microwave system as in Figure 1 AHSCS .This system used PRN phase modulation of the uplink-to transponder signal that was phase-coherently transponded to earth and correlated to determine tha two-way signal delay. A different PRN modulation was applied to the one wat time transfer signal from the space clock. Figure 4 shows the light-time diagram of the modulated signals and Figure 5 shows a block diagram of the proposed system. We expected that, for each pass of the shuttle, precision of better than 1 nanosec would be possible in time transfer and that 1 part in 10^{14} in frequency comparison should be possible. With the present positioning capability made possible by GPS, that is comparable to the signal wavelength, it is not out of the question that the accuracy of frequency comparison could be extended by recovering, from orbit to orbit, the phase of the signal from the microwave system.

8. PULSED LASER TIME TRANSFER

Pulsed laser time transfer between the earth and a moving vehicle is not a new idea. The first transfer I am aware of is a test relativistic gravitation that was done by with clocks on propellor driven aircraft.[19] In the late 1970s, development began at SAO on a Satellite Time and Frequency Transfer (STIFT) [18] in which pulsed laser time transfer would compare of the timing accuracy of laser and microwave techniques. The first actual test of pulsed laser timing from space to earth I know of was done by the LASSO[20] project.

The HMC retroreflector array, shown in Figure 6, consists of 20 fused silica cube corners, each 1 cm in diameter, mounted in a 65 mm diameter hemispherical base.[21] This shape provides a hemispherical field of view, permitting reflections independent of spacecraft attitude. Laser pulses impinging on the retroreflector array are brought to a photomultiplier tube by an omnidirectional fiber optic light collector. The collector consists of a 22-cm long bundle of 127 optical fibers, each 100 μm in diameter. At one end of the bundle the fibers are splayed out into a hemispherical pattern and inserted through holes drilled in a 1.5-cm diameter hemispherical shell that is mounted at the apex of the retroreflector array. The hemispherical fiber array ensures that at least one fiber is illuminated by light coming from anywhere in a hemispherical field of view. The other end of the fiber bundle connects to a photomultiplier tube (PMT) that detects the laser pulses.

The time transfer system's ability to resolve sub-nanosecond intervals results from a high-precision space-qualified event timer developed at the Los Alamos National Laboratory that is capable of timing with a resolution of 10 ps.[22] The output pulse from the PMT's preamplifier is sent to a constant-fraction discriminator (CFD), which produces a pulse whose shape is largely independent of the input pulse amplitude. [23-25] The CFD output pulse goes to the timeinterpolator,[26,27] which is a combined digital and analog hybrid circuit that has the effect of subdividing the period of a 100 MHz clock signal by a factor of 1000. Measurements of the interpolator's integral linearity, which measures the total timing error for any bin compared to a perfectly linear interpolator, are shown in Figure 7. The maximum excursion of the integral linearity is less than 10 ps, with test-to-test repeatability of less than 1 ps, and variation of less than 5 ps in any bin from 10°C to 30°C. The resulting systematic variations in measured time are repeatable and can be calibrated to a few picoseconds.

9. CONCLUSION

The time keeping accuracy of newly developed primary frequency standards in various nation's standards labs appears to be beyond our present time comparison capibility . It is clear that presently available technology using space techniques, and spaceborne high stability clocks, once implemented, would permit global synchronization at levels of a few picoseconds.

The question is how to support the realization of these technologies. While at present, there are few actually stated requirements, as in many previous situations, once a new technology becomes accessible, we can expect a strong effort toward implementing it in many new ways.

10 REFERENCES

[1] J. Ruttman and F.L. Walls, "Characterization of frequency stability in precision frequency sources," Proc. IEEE, vol. 79, 1991.

[2] D. J. Wineland, J.C. Bergquist, J.J. Bollinger, W.M. Itano, D.J. Heinzen, S. L. Gilbert, C.H. Manney, and M.G. Raizen, "Progress at NIST toward absolute frequency standards using stored ions," *IEEE Trans. on Ultrasonics, Ferroelectrics and Frequency Control,* vol. 37, no. 6, Nov. 1990.

[3] R.F.C.Vessot, E.M. Mattison, W.J. Klepczynski, G.M.R. Winkler, I.F. Silvera, H.P. Godfried, and R.L. Walsworth, "Present clock stability and realistic prospects for the future," in *Proc. Fourth Marcel Grossman Meeting on General Relativity* (R. Ruffini, Ed.) Rome, Italy, June 17-21, 1985

[4] L.A. Rawley, J.H. Taylor, M.M. Davis, and D.W. Allan, "Millisecond pulsar PSR 1937 + 21: A highly stable clock," *Science,* vol. 238.

[5] A.E.E. Rogers and J.M. Moran, Chapter 5, "Interferometers and Arrays," Chapter 5 in *Methods of Experimental Physics,* Astrophysics, M.L. Meeks, Ed., vol. 12. New York: Academic Press, 1976.

[6] E. B. Formalont and R.A. Sramek, "Measurement of the solar gravitational deflection of radiowaves in agreement with general relativity," *Phys. Rev. Lett.,* vol. 36, p. 1475, 1976.

[7] N. Bartel, V. Dhawan, T. Krichbaum, D.A. Graham, I.I.K. Pauliny-Toth, A.E.E. Rogers, B.O. Rönnäng, J.H. Spencer, H. Hirabayashi, M. Inoue, C.R. Lawrence, I.I. Shapiro, B.F. Burke, J.M. Marcaide, K.J. Johnston, R.S. Booth, A. Witzel, M. Morimoto, and A.C.S. Readhead, "VLBI imaging with an angular resolution of 100 microarcseconds,"*Nature,* vol. 334, no. 6178, pp. 131-135, 1988.

[8] R.S. Badessa, R.S. Kent, R.L. Nowell, and C. L. Searle, "A Doppler-cancellation technique for determining the altitude dependence of gravitational redshift in an earth satellite," *Proc. IRE,* vol. 48, p. 758, 1960.

[9] D. Kleppner, R.F.C. Vessot, and N.F. Ramsey, "An orbiting clock experiment to determine the gravitational redshift," *Astrophys. Space Sci.* , vol. 6, pp. 13-32, 1970.

[10] R.F.C. Vessot, M.W. Levine, E.M. Mattison, E.L. Blomberg, T. E. Hoffmann, G.U. Nystrom, B.F. Farrell, R. Decher, P.B. Eby, C.R. Baugher, J.W. Watts, D.L. Teuber, and F.D. Wills, "Tests of relativistic gravitation with a space-borne hydrogen maser," *Phys. Rev. Lett.,* vol. 45, pp. 2081-2084, Dec. 1980.

[11] A.J. Tucker and B.M. Fannin, "Analysis of ionospheric contributions to the Doppler shift of C.W. signals from artificial earth satellites," *J. Geophys. Res., Space Physics,* vol. 73, no. 13, pp. 4325-4334, 1968.

[12] R.F.C. Vessot and M.W. Levine, "Performance data of space and ground hydrogen masers and ionospheric studies for high-accuracy frequency comparison between space and ground clocks," in *Proc. Twenty-eighth Annual.Symp. Frequency Control,* U.S. Army Electronics Command, Ft. Monmouth, NJ, 1974, pp. 408-414.

[13] R.F.C. Vessot and M.W. Levine, NASA Experimental Final Redshift Report, GPA Project Report, Contract NAS8-27969.

[14] R.F.C.Vessot and M.W. Levine, "A time-correlated four-link Doppler tracking system," in A Closeup of the Sun (M. Neugebauer, and R.W.Davies, Eds.) JPL Publication 78-70, NASA, 1978.

[15] T. Piran, E. Reiter, W.G. Unruh, and R.F.C. Vessot, "Filtering of spacecraft Doppler tracking data and detection of gravitational radiation," *Phys. Rev. Lett., D,* 34, no. 4, p. 984, 1986.

[16] H. Penfield, E. Imbier and R.F.C. Vessot "Design of the Satellite Time and Frequency Transfer (STIFT) Microwave Ground Terminal" *Proceedings of the 14th Annual Precise Time and Time Interval Application and Planning Meeting,* Goddard Space Flight Center, Greenbelt, MD, Dec. 1–2, 1982, p. 223.

[17] D.W. Allan, C.O. Alley, N. Ashby, R. Decher, R.F.C. Vessot, and G.M.R. Winkler "Ultra-Accurate International Time and Frequency Comparison via an Orbiting Hydrogen Maser Clock," *Third Symposium on Frequency Standards and Metrology,,*Aussois, France; October 1981, Journal dePhysique, Colloque C8, Supplement au nToQ12, Tone 42, Decembre 1981, pp. 395–413.

[18] D.W. Allan, C.O. Alley, Jr., R. Decher, R.F.C. Vessot, G.M.R. Winkler "Shuttle experiment to demonstrate high accuracy global time and frequency transfer" *IEEE Trans. on Geoscience and Remote Sensing*, Vol. GE–20. No. 3, July 1982.

[19] C.O. Alley,"Time Experiments in Gravitational Fields with Atomic, Aircraft, and Laser Light Pulses", *Quantum Optics, Experimental Gravity, and Measurement Theory,* ed. P Meystre nad M Scully, Plenum Publishing Corp, 1983, pp. 363-427.

[20] P. Friedelance and C. Veillet, "Operation and Data Analysis in the LASSO Experiment", *Metrologia* vol 32, pp. 27-33, 1995.

[21] E.M. Mattison and R. F. C. Vessot, "High Precision Time Transfer in Space with a Hydrogen Maser on Mir"in Proceedings of the 1995 Precise Time and Time Interval Applications and Planning Meeting (San Diego, CA, 25 Nov-1 Dec 1995),pp 181-192.

[22] E. M. Mattison, D. Boyd, L.M. Coyle R.C. Smith, and R. F. C. Vessot, "A 10 Picosecond Event Timer for Precise Time Transfer in Space" ,in Proceedings of the 5th Symposium on Frequency Standards and Metrology (Woods Hole, MA, 16-20 Oct 1995), pp 475-477.

[23] R.C. Smith, "Optimal design of high speed analog APD receivers," *SPIE Proceedings* 987, Sept 1988; Los Alamos National Laboratory report LA-UR-88-3731.

[24] B.T. Turko and R.C. Smith, "A precision timing discriminator for high density detector systems," Lawrence Berkeley Laboratory report LBL-30602.

[25] B.T. Turko, W.F. Kolbe and R.C. Smith, "Ultra-fast voltage comparators for transient waveform analysis," *IEEE Trans. on Nucl Sci.* NS-37, 424 (1990).

[26] R. Nutt, "Digital time interval meter," *Rev. Sci. Instr.* 39, 1342 (1968).

[27] B. Turko, "Multichannel interval timer," *IEEE Trans. Nucl Sci.* NS-31, 167 (1984).

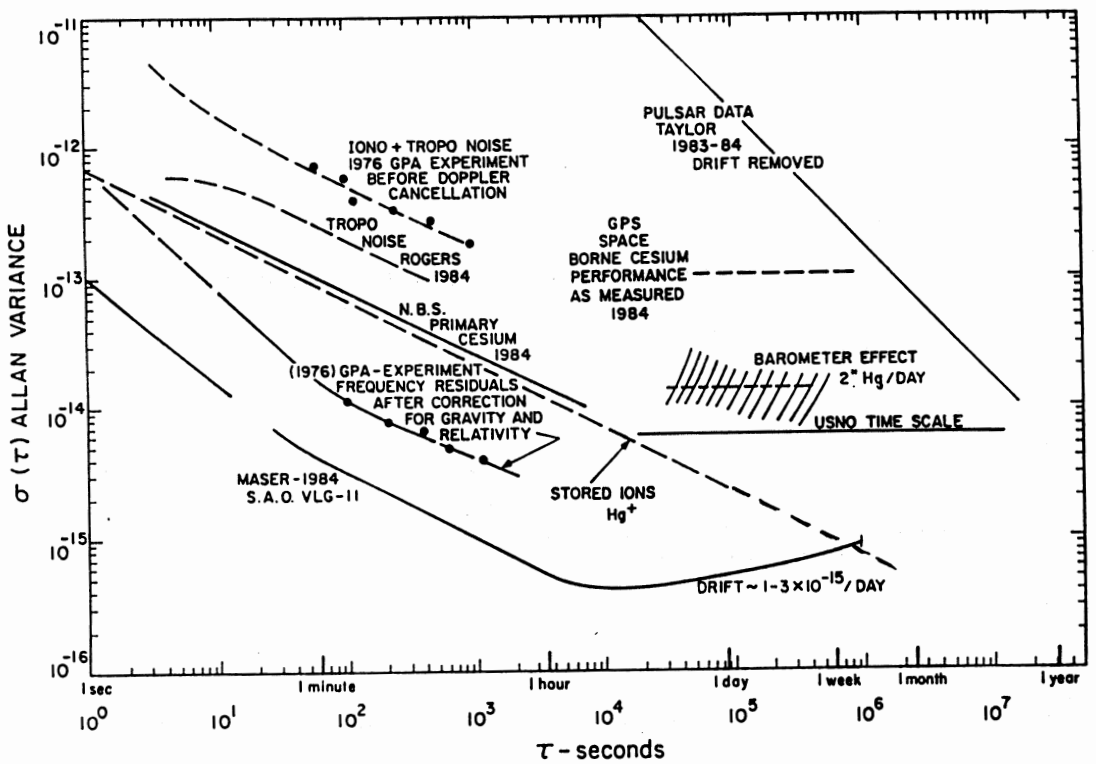


Figure 1
Allan Standard Deviation of Frequency Stability and of Propagation Effects

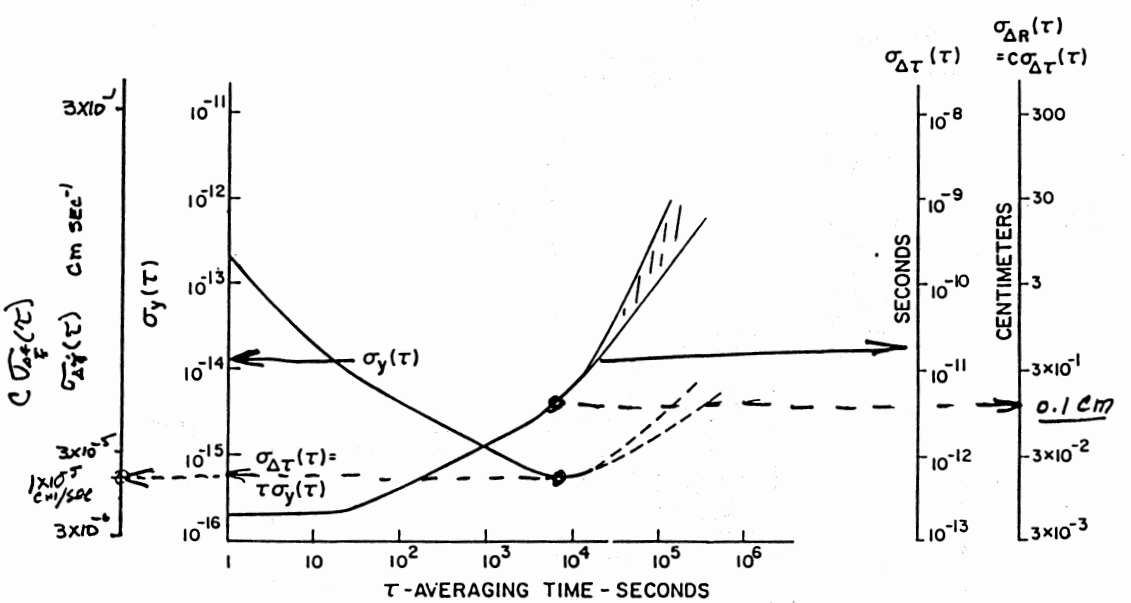


Figure 2
Time, Range Distance, and Range- Rate Dispersion from H-Maser Data in Figure 1

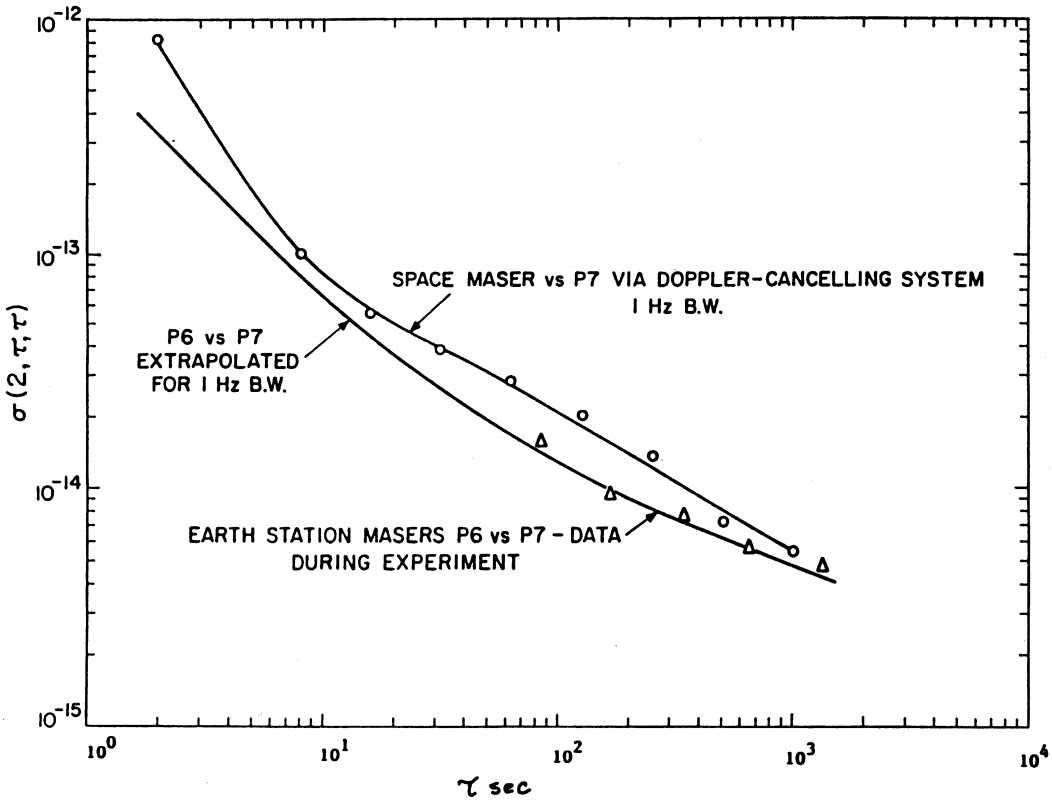


Figure 3
Frequency Stability of 1976 SAO/NASA GP-A Experiment after Removal of Relativistic Effects

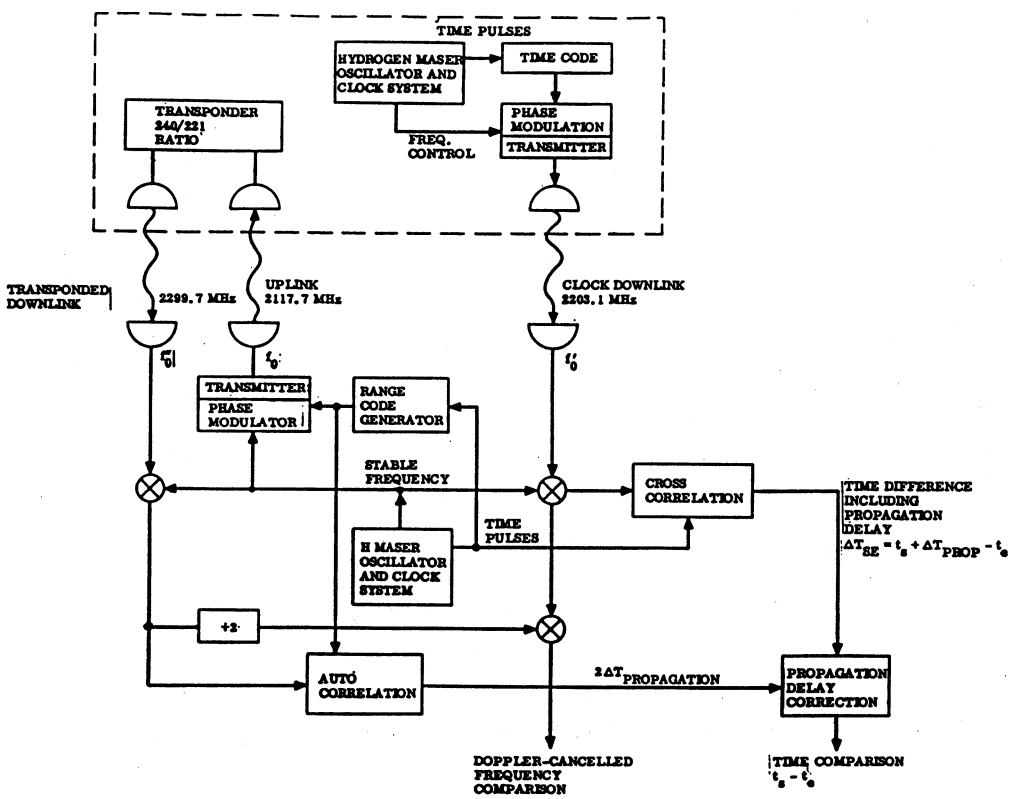


Figure 4
Time transfer by Phase Modulation of the Signals in Figure 1 of paper AHSCS

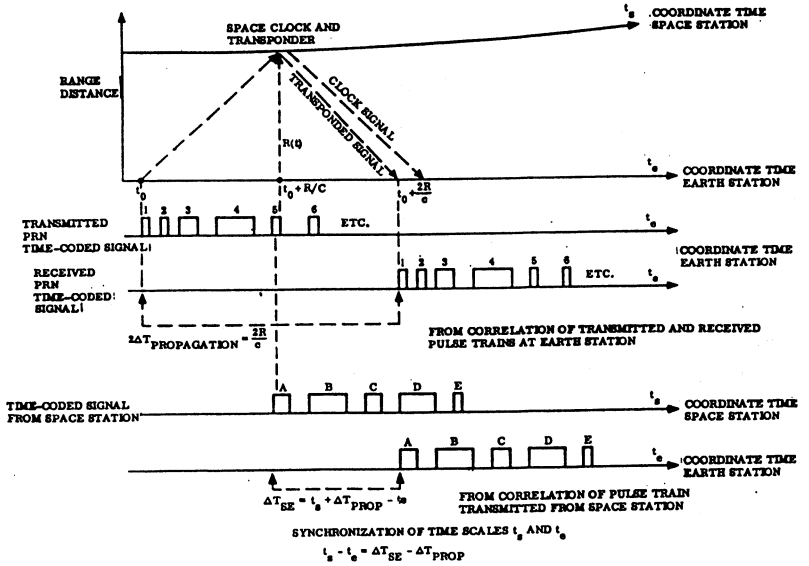


Figure 5
Light-Time Diagram of Signals in Figure 4

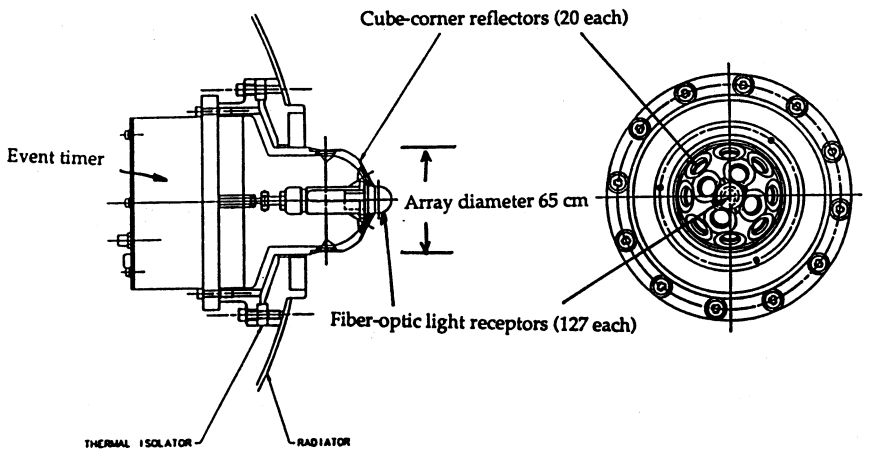


Figure 6
Retro reflector-Laser pulse Detector Designed for the SAO H-Maser C lock Test on Mir

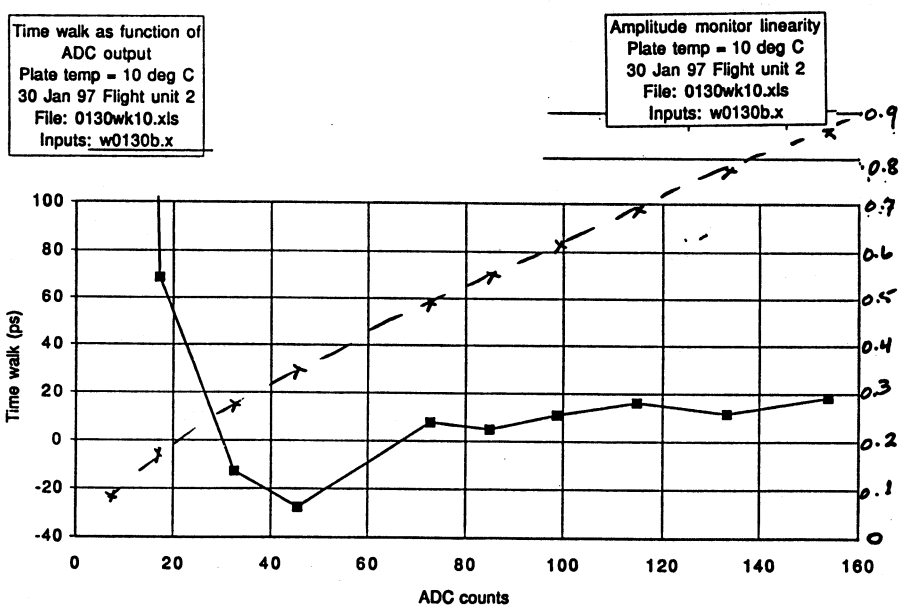


Figure 7
Time-Walk and Amplitude Monitor Linearity for the Event timer in Figure 6

PHARAO : A compact clock using laser cooled cesium atoms.

P. Lemonde[2], P. Laurent[1], E. Simon[1], G. Santarelli[1], A. Clairon[1], C. Salomon[2], N. Dimarcq[3], P. Petit[3], C. Audoin[3], F. Jamin Changeart[4], F. Gonzales[4]

(1)Laboratoire Primaire du Temps et des Fréquences, 61,av. de l'Observatoire 75014 PARIS, FRANCE
(2)Laboratoire Kastler Brossel, 24, rue Lhomond 75231 PARIS, FRANCE
(3)Laboratoire de l'Horloge Atomique, Bat.221, Université Paris-Sud, 91405 ORSAY, FRANCE
(4)Centre National d'Etudes Spatiales, 18, av. Edouard Belin, 31055 TOULOUSE, FRANCE

ABSTRACT

We describe a compact cold atom cesium frequency standard designed as a prototype of a space device. It will be tested in the reduced gravity of aircraft parabolic flights. On earth, the experimental resonance linewidth is 14 Hz with a preliminary signal to noise ratio around 300. The main advantages of this clock are its transportability and reliability. This is a major step toward a space clock using cold atoms. Minor changes would optimize its performances in presence of gravity and lead to performances comparable to those of the BNM-LPTF atomic fountain. It would then be an ideal tool the comparison of distant ground based clocks.

I Introduction

Laser cooling of neutral atoms has drastically improved cesium clocks in the last few years. The very low atomic velocities easily obtained with this technique (~1 cm/s for cesium [1]) allow much longer interrogation times than in classical devices. The resonance is proportionally narrowed and hence the short term stability can be improved. Furthermore, most systematic effects which shift the atomic resonance are reduced with the atomic velocity. The first frequency standard operating with laser cooled atoms, the BNM-LPTF atomic fountain, indeed shows the best performances ever obtained with cesium frequency standards. The stability is 2×10^{-15} on a few hours of integration time [4]. For longer integration time, the stability is limited by the H-maser used as a reference oscillator. The accuracy is presently 2×10^{-15} [2]. In the fountain, the interrogation time approaches one second and is limited by gravity. It is predicted that microgravity conditions should enable a further factor of ten improvement in the interaction time with a simple and compact device [3]. The objective of the PHARAO project is to develop such a space clock using cold cesium atoms.

II PHARAO project

In 1994, the French space agency (CNES) decided to support a preliminary research program on a space frequency standard using cold atoms. Three laboratories, the BNM-LPTF, the ENS-LKB and the LHA are cooperating for the construction of a prototype. Simultaneously, studies on local oscillators, frequency synthesis chains, microwave cavity modeling and time-frequency transfer are being performed.

The aircraft prototype

The main progress in laser cooling were achieved in the last few years. Although well understood and quite easily reproduced in laboratories with the recent development of diode lasers, these experiments are far from being compatible with space requirements. The weight of the atomic fountain is about 2 tons, with some hundreds of optical components. The device has to operate in a quiet environment.

We constructed a much smaller and reliable prototype. We want to record the

microwave resonance during the reduced gravity of aircraft parabolic flights. This device is also designed to be transformed to a high performance transportable frequency standard.

The experimental set-up is shown on figure 1. As the atomic fountain [4], the prototype operates in a pulsed mode. In a vacuum chamber, typically 10^7 atoms are first captured and cooled from a vapor at the intersection of the six beams of an optical molasses. This capture process takes between 100 and 500 ms, depending on the desired number of cold atoms and on the vapor pressure. The atoms are launched and cooled further in 3 ms. They experience a microwave interaction in a cylindrical TE_{013} cavity and finally, are optically detected. Due to the absence of gravity in the plane, only a single pass in the cavity is possible. All the laser beams are generated by a separated optical bench. A frequency chain synthesizes the 9.192...GHz microwave field from a 10 MHz ultra stable quartz oscillator. A computer generates the timing and processes the data.

Optical bench

The optical bench has been designed as a separated and closed box. Thus, no stray light disturbs the operation of the experiment. Its temperature is servo controlled at ~30 °C. 8 optical fibers link the bench to the vacuum tube : six for the optical molasses, two for the detection. They are monomode polarizing optical fibers. The output power of each fiber is servo controlled by means of voltage controlled retarder plates. The laser power is maintained within 10^{-3} with a response time of ten milliseconds. The dimensions of the optical bench, 65x65x15 cm, have been drastically reduced as compared to those of the atomic fountain.

Four diode lasers generate all the beams necessary for the experiment. The master laser is frequency locked 2 MHz below the F=4-F'=5 transition of the cesium D_2 line. Two different master diode laser will be tested in the aircraft. The first one is an extended cavity diode laser, the second one is a spectrally narrowed by weak optical feedback and fast electronic servo-lock DBR diode laser. Both have a linewidth in the 100 kHz range, a value much lower than the cesium D_2 natural linewidth (Γ=5 MHz).

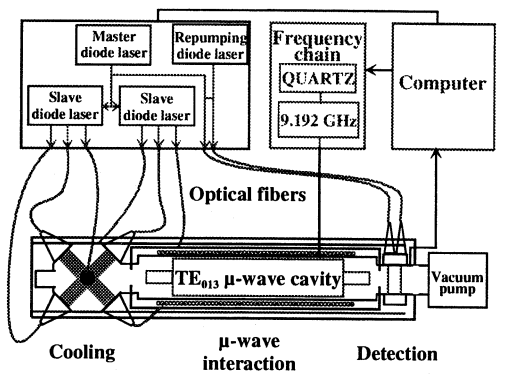


Figure 1 : the experimental set-up of the compact cold atom cesium frequency standard.

A part of the master laser beam is coupled into an optical fiber for the atomic detection. The output power can be adjusted up to 6 mW. The other part, after double-pass through an acousto optic modulator (AOM) which sets the molasses detuning up to 12 Γ, is used to inject lock two slave diode lasers which provide the cooling beams. Each delivers 150 mW of laser power. After a double pass in an AOM, each slave laser beam is divided in three and coupled into the fibers. The maximum laser power at the output of each fiber is ~13 mW. The frequency difference ν of the two AOM sets the atomic launching velocity V according to the relation :

$$V = \sqrt{3}\,\nu\,\lambda$$

where λ is the optical wavelength. The atomic velocity can be set from 0.5 m/s to 8 m/s. These AOM also allow a quick change of the optical power. To avoid any residual light shift of the atomic resonance during the microwave interaction, mechanical shutters insure a complete extinction of all the laser beams.

An additional DBR diode laser is locked on the F=3-F'=4 transition of the cesium D_2 line. A part of the beam (a few

mW) is splitted and mixed to four of the cooling beams to pump the atoms in the upper hyperfine state (F=4) of the ground state during the cooling process. The other part (100 μW) of the beam is used for the detection of the atoms in F=3 through a separated optical fiber.

Vacuum Chamber

The whole apparatus is 1 meter high and has a 300 mm external diameter. The temperature of the whole chamber is regulated by a heating coaxial wire. The cooling region contains a low pressure cesium vapor (10^{-6} Pa). The six optical fibers for the optical molasses are connected with FC type connectors to optical expanders screwed around the cooling region. The laser beams are expanded to a 8 mm waist radius with 100 mm lenses. They are distributed in three orthogonal pairs of counterpropagating beams. The polarization of each beam is linear and perpendicular to that of the counter-propagating beam [1]. The six laser powers are measured by using polarizing cubes and photodiodes integrated in the fiber collimators. The atoms are launched with the moving molasses technique [5], in the (1,1,1) direction of the trihedron defined by the beams [6]. With this geometry, only two frequencies are needed to launch the atoms and no beam passes through the cavity, which would limit the beam diameter to 1 cm. The cooling region is made of a titanium polyhedron whose angular tolerances are at the level of 10^{-4} radian. Thus, no adjustment of the beams direction is needed. The fiber-collimator systems can be removed and put back with a reproducibility of the beam direction at the level of a few 10^{-4} radian. The resulting launching direction is defined within 5×10^{-4} radian uncertainty. The overlap of counterpropagating beams is insured within 0.1 mm. The pressure in the interaction and detection regions is a few 10^{-8} Pa to avoid collisions and parasitic fluorescence at the detection. It is maintained by a 20 l/s ion pump and graphite tube cesium getters.

The microwave cavity is a 20 cm long TE_{013} cylindrical copper cavity. It has a loaded Q factor of several thousands. This cavity has been chosen for three reasons :

- It is long enough to allow a long interrogation time. The atoms spend half of their flight time inside the cavity.
- The holes at each end of the cavity can be large (1 cm) without disturbing the microwave field.
- It has a phase and amplitude distribution symmetry along the atoms pass.

In order to avoid microwave leakage, the length of the end cut-off are 8 cm.

A highly homogeneous static magnetic field (a few mG) is produced with a long solenoid and a magnetic shield around the cavity. Two compensation coils and three additional magnetic shields ensure the direction homogeneity of the magnetic field along the whole experiment. The total axial static magnetic attenuation is about 10^6 at the center of the cavity. The attenuation in the transverse direction is not measurable.

After the atoms pass through the cavity, the population of both hyperfine levels of the clock transition are independently measured by the fluorescence induced by two laser beams. Two optical fibers are connected to the detection region. Each beam is expanded to a 1 cm diameter by a fiber collimator. The F=4-F'=5 beam is divided into two parts : one to detect the atoms in the F=4 level, the second one, mixed with the F=3-F'=4 beam , to detect the atoms in the F=3 level. The two detection beams are stopped down to obtain rectangular beams of size 10x5 mm and are circularly polarized. They are parallel, separated by a few millimeters and their direction is perpendicular to the atomic velocity. A single mirror, fixed on the other side of the vacuum tube, retro-reflects the two beams to create standing waves. A small part of the F=4 beam is not reflected so that the atoms in this state undergo a traveling wave after they have been detected. The atoms in the F=4 level are first detected in the standing wave. They are then pushed away by the radiation pressure of the traveling wave. The remaining atoms, in the F=3 level, are

optically pumped to F=4 and detected in the second beam. Two condenser lenses collect about 3% of the fluorescence emitted by the atoms in each laser beam. The photodiode signals are digitized and processed by the computer.

Frequency chain

A 10 MHz reference quartz oscillator is multiplied to 100 MHz in a commercial multiplier. The 100 MHz signal is band pass filtered and then routed to a custom x92 multiplier. The output power is 10 dBm. The resulting 9.2 GHz and the 7.3...MHz of a RF synthesizer driven by the computer are mixed in a single side band mixer with better than 25 dB image and carrier rejection. The output signal amplitude is adjusted by a voltage controlled microwave attenuator. To reduce the effect of the magnetic field sensitivity of the quartz oscillator and of the microwave multipliers, the chain is enclosed in a mu-metal box. Its temperature is regulated.

During the aircraft flight, the experiment undergoes variations of acceleration : 20 s at 2 g, 20 s at 0 g, 20 s at 2 g, and 2 minutes at 1 g. This sequence is repeated 31 times. That sets strong constraints on the accelerometric sensitivity of the quartz oscillator. The projected atomic resonance linewidth is in the Hertz range. To scan this resonance without introducing errors greater than a few percents, the frequency retrace from one parabola to the other must be a few 10^{-12}. The typical static g-sensitivity of a high performance commercial quartz is a few 10^{-11}, with a frequency stability of $3\text{-}6\text{x}10^{-13}$ between 1 and 10 seconds. The LCEP [7] realized a special quartz oscillator for our application. It has a $5\text{x}10^{-12}$/g sensitivity and a 1 s stability of $1.8\text{x}10^{-13}$. The frequency chain has been used by the LPTF atomic fountain. The resulting stability was $1.5\text{x}10^{-13}\,\tau^{-1/2}$.

III Experimental results

The whole device is now operating in a laboratory environment. The number of detected atoms is a few 10^7 per shot. This number is obtained with an optical molasses and not with a magneto-optical trap. The relative fluctuations of the atom number is very low, less than 10^{-3} from shot to shot. we have deduced from the time of flight width an atomic temperature of less than 10 μK.

Figure 2 shows the microwave resonance (experimental and theoretical) of the atoms initially in F=4, m_F=0. The atoms in the other m_F substates of the F=4 state are not removed and thus contribute to the signal. It has been obtained for a launching velocity of 4 m/s which corresponds to an interrogation time of 58 ms. The agreement between theory and experiment is quite good.

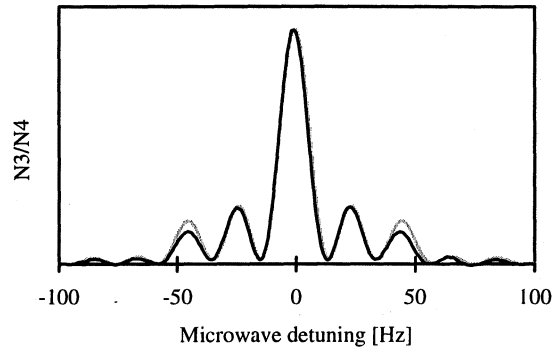


Figure 2 : atomic micro-wave resonance : we plotted the ratio N3/N4 of the populations of the two hyperfine levels. the gray curve is a numerical simulation. The black curve is the experimental resonance obtained with a one Hz step and no data averaging.

A slight disagreement only appears on the second sideband, which still has to be understood. The linewidth is 14 Hz, limited by gravity, with a contrast of 96 %. Unlike the Ramsey interrogation scheme, the shape of the resonance strongly depends on the microwave power.

To take into account all systematic effects on the atomic resonance, one needs to know very accurately the phase and amplitude of the field at every point inside the cavity. The best cavity choice for a space version of this frequency standard is still an open question.

The signal to noise ratio is 300 per shot, most probably limited by the fluctuations of the distribution of the atoms among the 9 Zeeman substates.

The experiment has been operated two weeks continuously.

III Conclusion

The prototype will be tested in parabolic flights by the middle of year 1997. We want to demonstrate its good operation in this hostile environment : vibrations of the plane, several °C variations of the temperature during the 3 hours of the flight, variations of gravity, rotation of the plane in the earth magnetic field.... We expect to increase the interaction time between the atoms and the microwave and record a resonance fringe in the Hz range. After that, the prototype will be fully evaluated. To study the effects due to the use of the TE_{013} cavity, this prototype will be transformed to operate in the two following modes : a single atomic interaction or a double interaction in an atomic fountain configuration. With a fountain geometry we expect a stability of about $10^{-13}\,\tau^{-1/2}$ and an accuracy better than 10^{-15}. This transportable frequency standard would then be an ideal tool for the comparison of distant ground based clocks at a level of a few 10^{-16}, which cannot be obtained with existing time transfer techniques.

Acknowledgments

The PHARAO project has the financial and logistic support of the French Space Agency (CNES). We acknowledge the assistance of M. Lours, M. Dequin, P. Aynié, A.H. Gerard, D. Guitard.

References

[1] C. Salomon et al., Europhys. Lett., **12** (8), 1990.

[2] E. Simon et al., Proc. of the 11th EFTF, Neuchâtel (Switzerland).

[3] Proc. of the 25th Moriond conference on dark matter, cosmology, ultra-stable clocks and fundamental tests.

[4] A. Clairon et al., IEEE Trans. on Instr. And Meas., **vol.44**, N°2, April 1995.

[5] A. Clairon et al., Europhys. Lett. **16**, 165 (1991).

[6] K. Gibble et al., Phys.rev.Lett., **70**, 1993.

[7] Laboratoire de Chronométrie Electronique et Piézoélectricité, 26 chemin de l'épitaphe, 25030 Besançon, France.

The Time Transfer by Laser Link experiment on MIR (T2L2)

SAMAIN Etienne & FRIDELANCE Patricia
Observatoire de la Côte d'Azur, CERGA
Avenue Copernic, 06130 GRASSE
FRANCE

The T2L2 experiment under development at OCA (Observatoire de la Côte d'Azur) and CNES (Centre National d'Etude Spatial) permits the synchronisation of remote clocks with an accuracy in the 50 ps range. We have the opportunity to put a T2L2 payload on the Russian space station MIR in 1999. T2L2 is based on the propagation of light pulses in space which is better controlled than the radio waves propagation. To perform a T2L2 time transfer, laser stations and a satellite are needed, both equipped with a clock and a time tagging unit. Some new optical timers and the definition of a new time origin allow direct accurate time transfer without external calibration. The ultimate stability of the payload equipment is 1 ps over 300 s and 10 ps over 10 days. If the ground equipment has the same kind of performances one can measure the stability of two remote ground clocks having a stability in the range of 3.10^{-15} over the station visibility period (around 300 s). One can also accumulate the data of several successive passages of MIR to reach a stability in the 3.10^{-17} range over 10 days. T2L2 on MIR will allow us to transfer time with an improvement of 2 orders of magnitudes compared to the radio-electric techniques and will be able to measure the performances of the new clocks generation (Atomic fountain, Trapped ion, H-Maser).

I. Introduction

The Time Transfer by Laser Link experiment T2L2 [1], which as been proposed on the Russian space station MIR [2], will allow us to synchronise and to compare the frequency stability of several remote ground clocks with some performances never reached before. T2L2 will synchronise remote ground clocks with an accuracy better than 50 ps. It will be able to measure the frequency stability of the new clocks generation (Atomic fountain, Trapped ion, H-Maser) at their ultimate performances level. The stability and the accuracy of the T2L2 time transfer will be improved by two orders of magnitude as compared to the actual techniques (GPS, Two-Way) [3, 4]. This kind of time transfer experiment is the basis of many applications : fundamental physics experiments, clock comparison, time scale realisation... . For example, T2L2 on MIR permits to test the speed of light anisotropy with an improvement by a factor 7 compared to the best recent GPS measurements. T2L2 on MIR is also an indispensable stage to validate the optical time transfer technique which will be used in several future spatial applications (ACES : Atomic Clocks Ensemble in Space, SORT : Solar Orbit Relativity Test....)

II. Principle

T2L2 (Time Transfer by Laser Link) is a dedicated time transfer experiment under development at OCA (Observatoire de la Côte d'Azur) and CNES (Centre National d'Etudes Spatiales), France. The principle of this time transfer is based on the propagation of light pulses between the clocks to synchronise as in the LASSO experiment [5] but without external calibration. The light pulses carry the temporal information from one clock to another. A clock and a detection-timing system on board of the Russian station MIR are used as relay to allow time transfer between remote ground clocks. To perform a T2L2 time transfer, each ground clock to be synchronised has to be connected with a laser station (the time base of the laser station has to be the clock to be synchronised). The laser station emits some short light pulses (wavelength = 532 nm) in the satellite direction. An array of retroreflectors located near the payload returns a fraction of the received photons to the station. The station records the start time of the laser pulses and the return time after reflection from the satellite. The T2L2 payload records the arrival time on board. The synchronisation between ground clocks is calculated from all these data. One can determine accurately the offset between two ground clocks (one is able to measure the phase between the 10 MHz signals of the clocks) if the ground start times are measured accurately. We choose as time origin the 0 crossing of the clock sinusoidal output, at a given point

along the cable which transports this signal (for instance at the output plug of clock signal). Accurate synchronisation between several remote clocks will be then possible if one disposes of a system able to time a light pulse with respect to the origin of the sinusoidal signal. This will be performed with an optical timer under development at OCA [6, 7]. The arrival time on board of the satellite can be known with an arbitrary origin, since this origin will be the same for light pulses coming from the two stations. The stability specification of the payload equipment is 1 ps over 300 s and 10 ps over 10 days. If the ground equipment stability has the same kind of performances, one can measure the stability of two remote ground clocks having a stability in the range of 3.10^{-15} over the station visibility period (around 300 s). One can also accumulate the data of several successive passages of the station to reach a stability in the 3.10^{-17} range over 10 days.

III. Ground Segment.

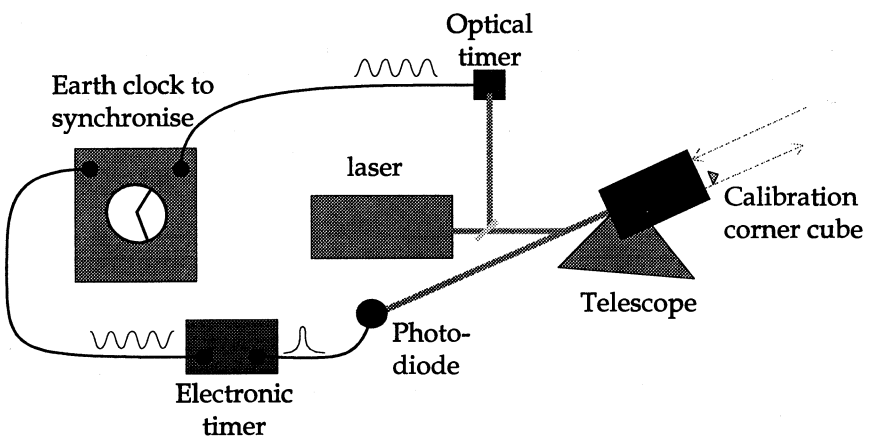


Figure 1 : *Ground segment configuration*

The ground segment is composed of the following elements (See figure 1) :

- Clock to synchronise
- Laser : a neodyme YAG laser emits short pulses at a 10 Hz rate. The output wave length is 532 nm (YAG doubled) and the full half-width is comprise between 50 ps and 200 ps.
- Telescope
- Optical timer : the principle is to pass the light pulse through an electro-optic crystal submitted to an oscillating electric field. Depending on the light pulse arrival time into the crystal, the polarisation of the latter is more or less changed. Conversely, analysing the polarisation at the output of the crystal allows us to deduce the arrival time. By synchronising the electric field oscillation with the clock output oscillating signal, one obtains an exact datation of the light pulse. Some experiments permit to envision datations with respect to a clock in the 30 ps range with a direct coupling of the 10 MHz clock signal to the timer (no frequency multiplication). Such an accuracy is obtained because there is no photon-electron transformation and because the device works with some sinusoidal signal.
- Return detector : it is a photodiode working in the Geiger mode. This mode allows us to detect single photons with a precision of 35 ps and to detect 100 photons or more with a precision below 10 ps. In this mode the time walk versus the photon number is about 150 ps/decade. A parallel measurement of the photon number is necessary to subtract this time walk. An other solution, degrading the precision of the measure but simplifying the set-up experiment, is to maintain the return level in the single photon mode.
- Electronic timer : its purpose is to time the photodiode output electrical pulse. Its time base comes from the clock to be synchronised. A timer designed by Dassault Electronic (D. E.) (France) permits to obtain datation with a precision of 5 ps and a time stability of 8 fs over 1000 s (see figure 2) . The origin of the datation is arbitrary but constant as long as the timer is powered.

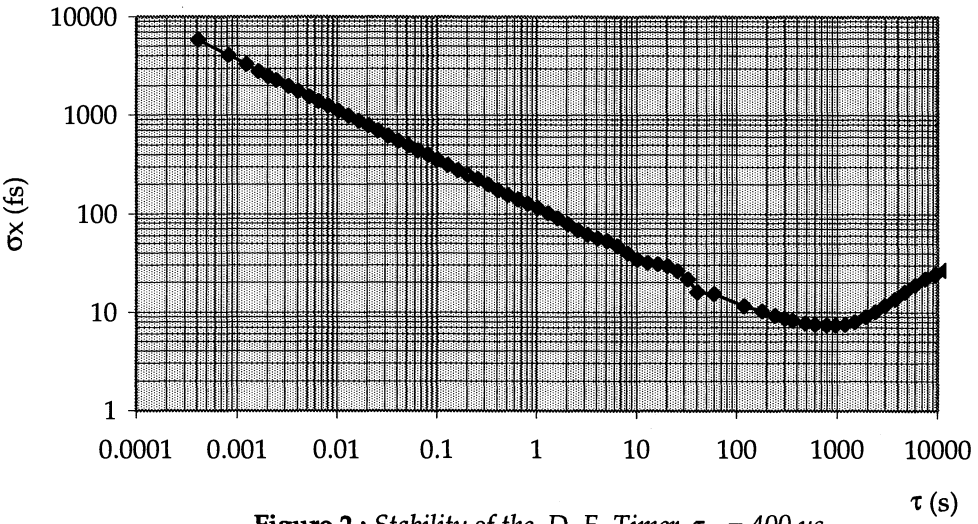


Figure 2 : *Stability of the D. E. Timer* $\tau_0 = 400\ \mu s$

IV. The satellite Segment.

The satellite segment is composed of the following elements (See figure 3) :

- An ultra stable oscillator (quartz)
- An electronic timer
- A detection unit
- A detection optic
- A retroreflector array

The ultra-stable oscillator (USO) which is the time base of the payload will be a DORIS USO or an USO developed by the LCEP laboratory (Besançon, France) for the PHARAO project (Projet d'Horloge Atomique par Refroidissement d'Atomes en Orbite) [8]. Respectively, their time stabilities (square root of TVAR) are 5.10^{-11} and 5.10^{-12} over a time $\tau = 200$ s, see figure 4. The detection unit, is designed to permit precision timing in the 30 ps level. It includes a photodetector and a device able to transform the electrical pulse coming from the photodetector for the timer. This performance is preserved even if the photon number is not constant or if the pulse width varies in the range : 50 ps - 200 ps. The timer will be equivalent to the ground timer described before. The timer and the detection unit will be designed and built by Dassault Electronique (D. E.). The optical package (detection optic an retro-reflectors), shown on the left side of figure 3, will include three (or more) corner cubes and some optical fibers connected to the point (extremity) of each corner cube. The optical fibers are put together and plugged into the detection unit described before. We can also imagine a configuration without any optical fiber but with a photodetector directly coupled on each corner cube. The orientations of the corner cubes are chosen so that there is no superposition between the corner cube fields of views. This way, only one corner cube at a given instant is able to reflect a light pulse coming from a laser station and only one optical fiber (the one which is connected to the corresponding corner cube) will carry to the detection unit the light pulse coming from this station. One eliminates then the error which would be induced if the attitude of the MIR station is not known with a sufficient accuracy.

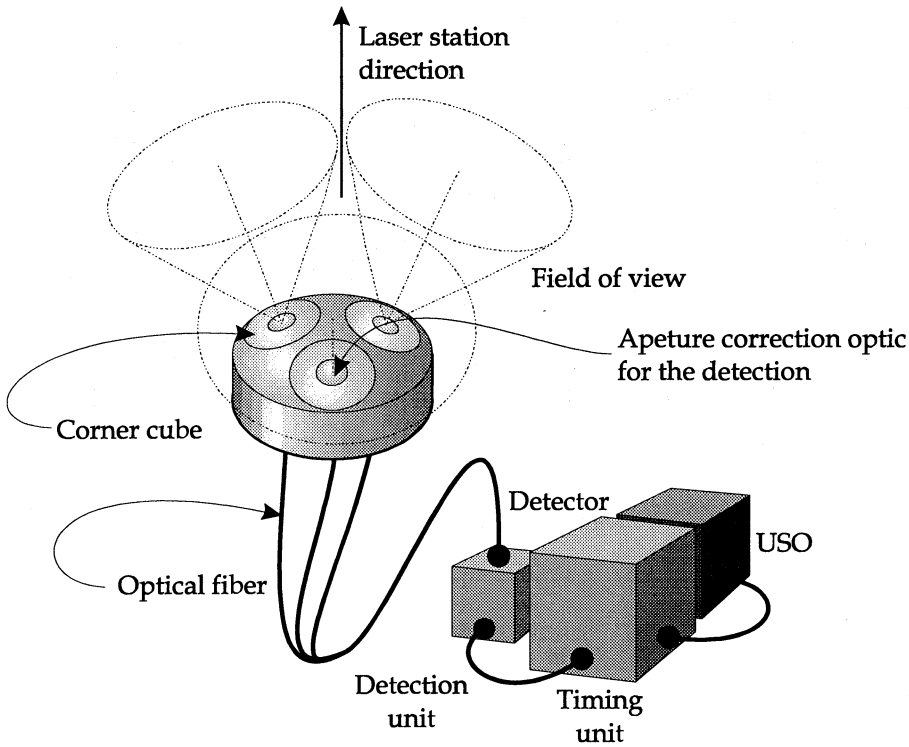


Figure 3 : *T2L2 Payload*

The total weight of the T2L2 payload should be around 15 kg. A prototype of the T2L2 experiment has been built by Dassault Electronique. This prototype includes the photodetector and the associated electronic but does not include the optical part (corner cube, optical fiber ...). Its purpose is to validate the payload timing performances and to simulate a real time transfer between two laser station clocks located at the same place via this payload prototype.

The first results extracted from the experiments lead by D.E. with some optical pulses sent onto the prototype are :

- Precision : 10 ps
- Time stability over 200 s : < 0,4 ps
- Time stability over 10 000 s : < 0,2 ps
- Drift : < 5 ps over 2 days

This experimental results are in accordance with the T2L2 specification.

V. Error Budget

One considers first the error budget for a laser link between a ground clock and the clock on board of MIR. The error budget applied to a time transfer between two ground clocks is then computed from the error budget between the ground clocks and the board clock. To deduce the error budget from two ground stations two cases are considered :

- a common view case : two ground stations are seeing the MIR station during a common period. The relatively low altitude of the MIR station implies a maximum distance between ground clocks smaller than 1000 km.
- a non common view case : a first laser station is operating on MIR during a period τ_1 and a second station during a period τ_2 after a dead time τ_3. This configuration will permit to transfer time over very long distances.

The clock board frequency has to be measured at least by one station. This can be performed by accumulating data over several passages of the MIR station. One assumes then that there is no effect induced on the error budget described below by the frequency accuracy of the USO.

V.1 Error budget between a ground clock A and the MIR station clock S

- The time transfer precision σ_{AS} between a ground clock A and the MIR station S is given by :

$$\sigma_{AS}^2 = \frac{\sigma_{Start}^2}{4} + \frac{\sigma_{Return}^2}{4} + \sigma_{Board}^2 + \sigma_{Relativity}^2 + \sigma_{Atmosphere}^2 + \frac{\sigma_{Calibration}^2}{4} + \frac{\sigma_{Geometric\_Ground}^2}{4} + \frac{\sigma_{Geometric\_Board}^2}{4}$$

From the first experimental results already obtained, one gets : σ_{AS} = 30 ps.

- The time transfer stability $\sigma_{X.AS}(\tau)$ between a ground clock A and the MIR station S is :

$$\sigma_{X.AS}^2(\tau) = \frac{\sigma_{X.Start}^2(\tau)}{4} + \frac{\sigma_{X.Return}^2(\tau)}{4} + \sigma_{X.Board}^2(\tau) + \sigma_{X.Atmosphere}^2(\tau)$$

The stability $\sigma_{X.AS}(\tau)$ is computed from the square root of the time variance TVAR [9]. All the stabilities written in the latter equation are given with an interval between each successive events $\tau_0 = 1/F_{laser} = 0{,}1$ s, where F_{laser} = 10 Hz is the laser rate. From some measurements and some extrapolations one can deduce the time stability. The figure 4 shows the T2L2 Ground_Board stability. The stability curves of the LCEP and DORIS quartz oscillator are also plotted. The fourth curve represents the stability of the PHARAO clock. PHARAO is a cold atomic clock project in space which should have a stability in the 10^{-16} range over 10 days and an accuracy in the same order.

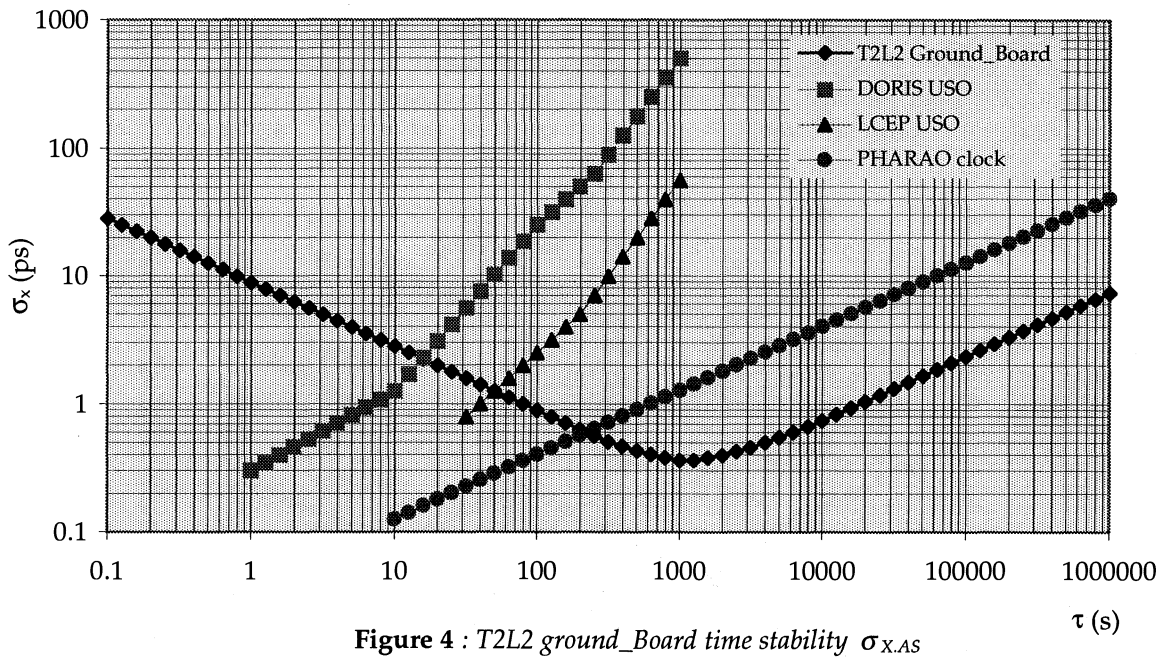


Figure 4 : *T2L2 ground_Board time stability* $\sigma_{X.AS}$

- Time transfer accuracy $\varepsilon_{X.AS}$ between a ground clock A and the MIR station clock S :

The time accuracy $\varepsilon_{X.AS}$ over 10 days is estimated to 50 ps considering that the ground and satellite contributions are equal. It must be stressed that the time origin at the satellite is arbitrary so the phase between the ground clock and the satellite clock will be known with an uncertain delay which will remain constant (in the 50 ps range) during 10 days as long as the T2L2 payload is powered. This mean that the frequencies of the ground clock and the satellite clock can be compared with an accuracy in the range of 5.10^{-17}. This frequency accuracy performance will be useful for some futur projects when an ultra-stable clock in orbit like PHARAO will be connected to the T2L2 payload.

V.2 Error budget between two ground clocks A and B : common view case

In this case, the stability of the board USO must be taken into account over a time τ_{common}, where τ_{common} is the maximum time interval between two successive pulses.

- Time transfer precision σ_{AB} between two ground clocks A and B in common view case :

$$\sigma_{AB}^2 = \sigma_{AS}^2 + \sigma_{BS}^2 + \left(\frac{\tau_{common}}{\sqrt{3}} M\sigma_y(\tau_{common}) \right)^2$$

One assumes that the USO frequency is known and does not introduce any supplementary error. σ_{AS} and σ_{BS} are respectively the time transfer precision between the ground clock A and and B and the MIR Station clock S. $M\sigma_y(\tau_{common})$ is the modified Allan variance square root of the board USO taken over the time τ_{common} which is the maximum time interval between two laser shoots. If the laser rate is 10 Hz, roughly one has : τ_{common} = 0,1 s. For both USO envisaged, the latter term : $\frac{\tau_{common}}{\sqrt{3}} M\sigma_y(\tau_{common})$, introduced by the board oscillator imperfection, can be neglected as long as τ_{common} is smaller than a few seconds.

- Time transfer stability $\sigma_{X.AB}(\tau)$ between two ground clocks A and B in common view case :

$$\sigma_{X.AB}^2(\tau) = \sigma_{X.AS}^2(\tau) + \sigma_{X.BS}^2(\tau) + \left(\frac{\tau_{common}}{\sqrt{3}} M\sigma_y(\tau_{common}) \right)^2 \cdot \frac{\tau_{common}}{\tau}$$

As previous, the latter term, which add a white phase noise can be neglected as long as τ_{common} is smaller than a few seconds. This time stability $\sigma_{X.AB}(\tau)$ is shown on figure 5. The dead zone is due to the visibility of the space station MIR from the ground stations : the laser stations will be able to shoot on MIR during about 300 s per passage, and the station will be visible from a given station twice per day. The heavy curve represents the stability of the best clock of the world : the fountain clock developed in France.

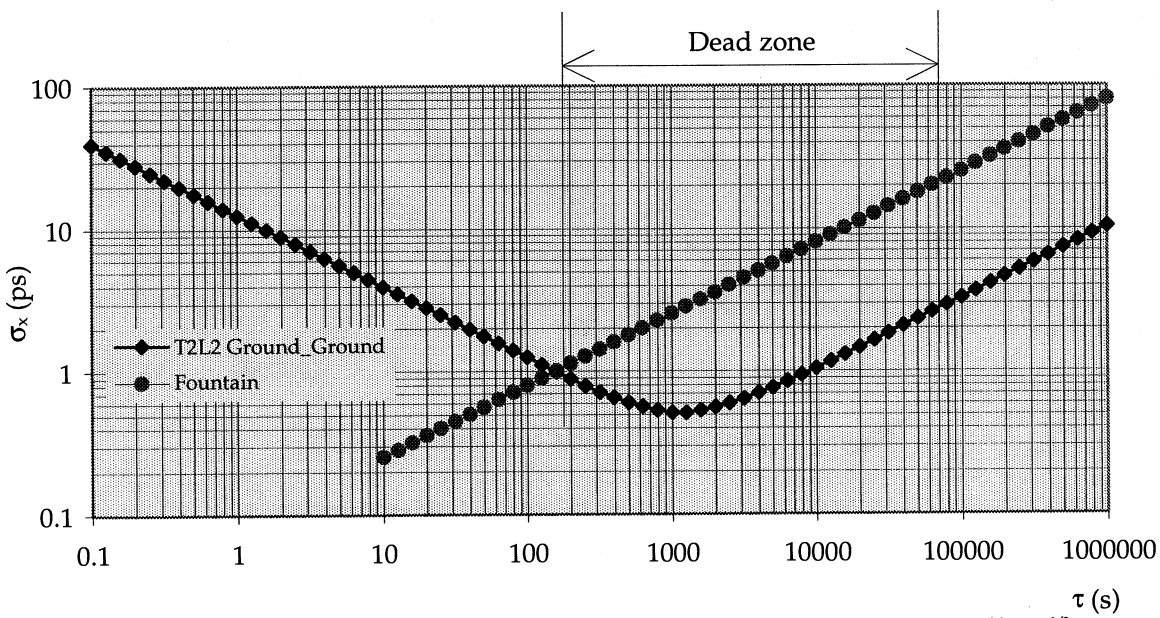


Figure 5 : *T2L2 ground_Ground time stability* $\sigma_{X.AB}$ $\sigma_{X.Fountain} = 8.10^{-14}\ \tau^{+1/2}$

- Time transfer accuracy $\varepsilon_{X.AB}$ between two ground clocks A and B in common view case :

The laser shoots have to be integrated during a time τ_i so that : $\sigma_{X.AS}(\tau_i) << \varepsilon_{X.AS}$ and $\sigma_{X.BS}(\tau_i) << \varepsilon_{X.BS}$ where $\varepsilon_{X.AS}$ and $\varepsilon_{X.BS}$ are respectively the time transfer accuracy between the ground clocks A and B and the board USO. One gets then :

$$\varepsilon_{X.AB}^2 = \varepsilon_{X.AS}^2 + \varepsilon_{X.BS}^2 + \left(\frac{\tau_{common}}{\sqrt{3}} M\sigma_y(\tau_{common}) \right)^2$$

As in the precision budget the latter term can be neglected as long as τ_{common} is smaller than a few seconds. Considering that the ground stations are using the optical timer described before, one gets : $\varepsilon_{x.AB} = 70$ ps. The start times are known accurately so the phase of the clock signals of frequency ν_0 can be compared with an accuracy of $2\pi\nu_0\varepsilon_{x.AB}$. The clock frequencies of the ground station A and B can be compared with an accuracy $\varepsilon_{x.AB}/T_{obs}$, where T_{obs} is the measurements duration. With $T_{obs} = 10$ days one has : $\varepsilon_{x.AB}/T_{obs} = 7.10^{-17}$.

V.3 Error budget between two ground clocks A and B : non-common view case

A first laser station is operating on MIR during a period τ_1 and a second station during a period τ_2 after a dead time τ_3. In this configuration, which will permit time transfer over very long distances (for instance between Europe and the USA), the board USO stability must be taken into acount over the period τ_3 and cannot be negligated even with the LCEP oscillator. The time transfer performances will be degraded with respect to the common view case.

As in the previous common view case the laser shoots have to be integrated during a time τ_i so that : $\sigma_{X.AS}(\tau_i) << \varepsilon_{X.AS}$ and $\sigma_{X.BS}(\tau_i) << \varepsilon_{X.BS}$ By taking into acount the noise generated by the board USO, one gets :

$$\varepsilon_{X.AB}^2 = \varepsilon_{X.AS}^2 + \varepsilon_{X.BS}^2 + \left(\frac{\tau_3}{\sqrt{3}} M\sigma_y(\tau_3) \right)^2$$

One assumes that the USO frequency has been measured by integrating data form several MIR passages The figure 6 illustrates the T2L2 time transfer accuracy versus the distance L_{AB} separating the ground stations involved, for the DORIS and the LCEP USO.

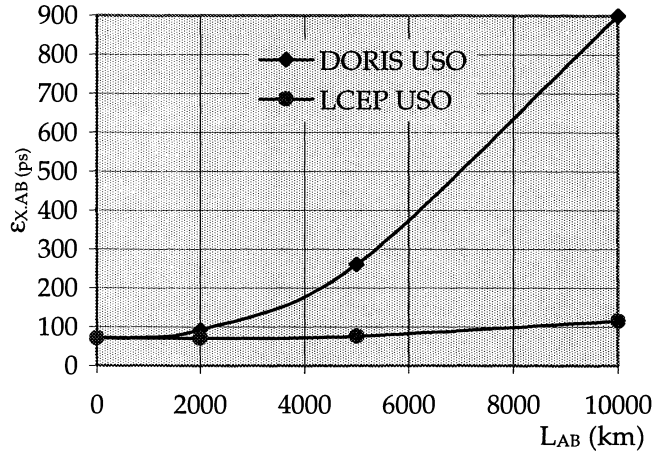


figure 6 *:time transfer accuracy version the distance L_{AB} between the ground clocks*

VI. Conclusion

The T2L2 on MIR experiment will permit ground time transfer with a stability and an accuracy which are consistent with the best clock available today. The recent improvement obtained on the quartz oscillators will allow us to perform a time transfer between remote clocks in the non-common view mode. T2L2 on MIR is an important stage for the future space mission ACES on the International Space Station ISS. The experimental environment between ISS and MIR will be similar (altitude, radiation,...) so the MIR experiment is a good exercise for the future. Many laboratories coming all over the world will work together. Some interesting collaborations will be then created. T2L2 on MIR will permit also to perform some scientific experiment such as VLBI interferometry and the test of anisotropy of the one way speed of light [10].

---

[1] P. Fridelance, E. Samain & C. Veillet, *T2L2 - Time Transfer by Laser Link : A new optical time transfer generation*, to be published in Experimental Astronomy, 1997.

[2] P. Fridelance & E. Samain, *Vol longue durée sur MIR réponse à l'appel à idées. Projet T2L2, Transfert de temps par lien laser*, Doc CNES, 1996

[3] W. Lewandowski, C. Thomas, *GPS Time Transfer*, proceedings of the IEEE, **79**, No. 7, 991-1000, 1991

[4] D. Kirchner, *Two-way satellite time transfer via telecommunication satellites*, in proceedings of the IEEE, **79**, 983-990, 1991.

[5] P. Fridelance : *L'expérience LASSO*, Thèse de doctorat de Physique de l'Université Paris 6, 1994.

[6] E. Samain, *Le laser Lune millimétrique et nouvelles méthodes de datation optique*, thèse de doctorat de l'université de Nice Sophia Antipolis, 1995.

[7] E. Samain, *Optical Statistical Timer*, submitted to IEEE.

[8] C. Salomon, P. Lemonde, P. Laurent, E. Simon, G. Santarelli, A. Clairon, N. Dimarcq, P. Petit, C. Audoin, F. Gonzalez, F. Jamin Changeart, *PHARAO : A cold atom clock in microgravity*, in proc. of the Space Station Utilisation Symposium, SP 385, ESA Darmstadt, Germany, 1996.

[9] D.W. Allan, M.A. Weiss, J.L. Jespersen, *A frequency domain view of time domain characterisation of clocks and time and frequency distribution systems*, 45^{e} annual symposium on frequency control, 1991.

[10] P. Fridelance, E. Samain & T. Melliti, *Scientific applications of the Time Transfer by Laser Link (T2L2)*, these proceedings.

Scientific applications of the Time Transfer by Laser Link (T2L2)

P. Fridelance*, E. Samain* & T. Melliti**

* Observatoire de la Côte d'Azur, CERGA, Grasse, France

** Laboratoire de Gravitation et Cosmologie Relativiste, Paris, France

Abstract The new generation of optical time transfer (T2L2 : Time Transfer by Laser Link) under development at OCA and CNES allows the synchronisation of remote ultra-stable clocks and the determination of their performances over intercontinental distances. This synchronisation method associated with ultra-stable clocks, on Earth or in orbit, has many scientific applications that are presented here. Moreover, T2L2 is a basic element for the future missions in the solar system (SORT) intending to test some aspects of gravitation.

1. Introduction

The Time Transfer by Laser Link (T2L2) [1] under development at OCA (Observatoire de la Côte d'Azur) and CNES (Centre National d'Etudes Spatiales) allows the synchronisation of remote clocks and the measurement of their stability over many thousand of kilometres. This optical time transfer, associated with ultra-stable clocks on Earth or in orbit, has many scientific applications in fundamental physics and astronomy. Some applications are presented here in the framework of three space missions : T2L2 on MIR [2,3], ACES (Atomic Clock Ensemble in Space) on the International Space Station (ISS) [4], and SORT (Solar Orbit Relativity Test).

2. Applications in fundamental physics

Test of the isotropy of the one-way speed of light

A test of the isotropy of the one-way speed of light can be achieved with a laser link between a ground clock and an orbiting clock. The principle, similar to that performed by Krisher et all [5], is to compare the propagation time of laser pulses, along the same station-satellite path, but for different spatial orientations. The sensitivity of the experiment is estimated by $\delta c/c = \varepsilon/[2T(1-\cos\theta)]$ where ε represents the measurement uncertainty, T is a typical transmission time, and θ is the angle between the two laser links [6]. In our case, the measurement uncertainty is estimated by

$$\varepsilon^2 = 2\left[\sigma^2_{x.\mathrm{TimeTransfer}}(\tau)+\sigma^2_{x.\mathrm{GroundClock}}(\tau)+\sigma^2_{x.\mathrm{SpaceClock}}(\tau)\right]$$

where τ is the duration of each link, $\sigma_x(\tau)$ is the time stability over the time τ. One assume that the time transfer drifts are negligible, and that the clocks are syntonized. The stability of the time transfer is given in [3]. The ground clock is assumed to be a caesium fountain [7] having a time stability $\sigma_{x.\mathrm{GroudClock}}(\tau) = 8\ 10^{-14}\tau^{1/2}$; $100\ \mathrm{s} < \tau < 10^6\ \mathrm{s}$. For T2L2 on MIR with a DORIS USO [8], the measurement uncertainty is $\varepsilon \sim 6$ ps for $\tau \sim 13$ s, and the sensitivity is $\delta c/c \approx 4\ 10^{-10}$ (T ~ 4 ms, $\theta \sim 150°$). For T2L2 on MIR with a PHARAO quartz developed at LCEP (Laboratoire de Chronométrie, Electronique et Piézoélectricité) [9], the measurement uncertainty is $\varepsilon \sim 3.5$ ps for $\tau \sim 25$ s, and the sensitivity is $\delta c/c \approx 2.5\ 10^{-10}$ (T ~ 4 ms, $\theta \sim 150°$). For ACES, with a PHARAO clock [7], the measurement uncertainty is $\varepsilon \sim 2.5$ ps for $\tau = 65$ s, and the sensitivity is $\delta c/c \approx 2\ 10^{-10}$ (T ~ 3.5 ms, $\theta \sim 145°$). In summary, the best direct determination of the one-way isotropy of light yet published [10] is expected to be improved by a factor > 1000 with T2L2 on MIR (PHARAO quartz) or with ACES, and the recent direct determination using GPS data [11, 12] is expected to be improved by a factor ~ 7. The sensitivity presented here corresponds to the

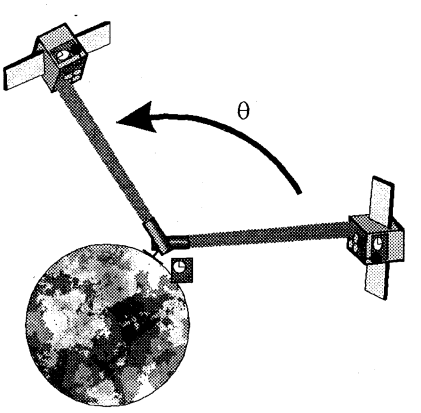


Figure 1 *Principle of the test of the isotropy of the one-way speed of light.*

case where the anisotropy lies in the orbital plane of the satellite, and $\delta c/c$ is determined for one satellite crossing. It should be possible to determine the anisotropy for many spatial orientations and to improve the sensitivity by using many crossings and the data from many stations.

Test of ultra-stable clocks for the future space missions

Ultra-stable clocks are under development in different laboratories : cold atoms clocks at LPTF-ENS-LHA (Laboratoire Primaire des Temps et Fréquences, Ecole Normale Supérieure, Laboratoire de l'Horloge Atomique, France), PTB (Physikalisch Technische Bundesanstalt, Germany) and NPL (National Physical Laboratory, UK) ; trapped ion clock at JPL (Jet Propulsion Laboratory, USA) and CRL (Communication Research Laboratory, Japan) ; and hydrogen masers at Neuchâtel Observatory (Switzerland), SAO (Smithonian Astrophysical Laboratory, USA), and Institute for Metrology for Time and Space (Moscow, Russia). To measure the performances of these clocks, it is necessary to compare them to clocks having the same kind or better performances. That is actually not possible because the clocks are developed in distant laboratories and they are not easily transportable. The optical time transfer permits the comparison of ground clocks with T2L2 on MIR [3], for example between the French cold atoms clock and the Swiss maser using a common view configuration, and between this clocks and the Americana ultra-stable clock using a non-common view configuration. The clocks can be compared in space with ACES, and their performances can be determined. Considering the performance improvement due to the micro-gravity (cold atoms clock only), the best space clock available can be selected. This clock will then be used in the future space missions doing very long base interferometry, or studying some aspects of gravitation (SORT).

Test of some aspects of gravitation

Ultra-stable clocks associated with an optical time transfer allow the study of some aspects of gravitation. The PHARAO experiment gives the opportunity to measure the Einstein effect at an accuracy never reached. If the ultra-stable clock is orbiting in an eccentric Earth orbit. The comparison of the measured frequency shift $(\Delta\nu/\nu)_{mes}$ and the theoretical value can be achieved at the level of 10^{-16}. As this frequency shift is proportional to $(1+\alpha)$, where α is the PPN (Post Newtonian Parametrised) parameter characterising the universality of the gravitational redshift, α can be determined with an accuracy of the order of 10^{-6} improving the Vessot-Levine determination by a factor 20 [13]. This parameter is very important to discriminate the gravitation theories, because if α is null, the theory should be metric and on the contrary, it should be non-metric. In fact, as stated by Will [14], "Every non-metric theory of gravity predicts a gravitational redshift which depends on the nature of the clock whose redshift is being measured."

3. Applications in astronomy

Very Long Base Interferometry

In radioastronomy, antenna are combined to obtain a better angular resolution, forming a Very Long Base Interferometer (VLBI). The signal of the astronomical object is recorded at each antenna with a clock signal. The signals are then correlated in deferred time. For that purpose, the clock signal must be very stable during the observation to correlated correctly the signals. That necessitates an hydrogen maser. The period during which the phase coherence is preserved between the signals issued from the different antennas depends on the observation frequency, and is typically a few minutes. To extract the more information, or to observe a faint source, the coherence time must be increased. That way, a known radiosource (in position) can be observed in the same time than the studied object to allow the samples adjustment [15]. That is only possible when a standard source lays at less than a few degrees of the object, that is not often the case. The optical time transfer can then be used to compare the phases of the masers connected to the antennas.

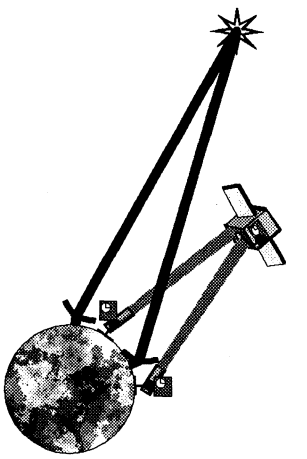


Figure 2 *VLBI observation combined with a laser link synchronisation of the masers.*

The integration time will then be longer, improving the observation sensitivity of at least a factor 2 in the case of T2L2 on MIR.

Geodesy

Orbit determination from the laser measurements will be improved permitting a better study of the forces acting upon the satellite. That is of great interest for the geodesy community, for example for the Earth gravity field determination. On the other hand, the VLBI observation of a radiosource well known in position, combined with a synchronisation of the masers linked to the antennas, allows to determine the distance between the antennas at the millimetre level. Such a measurement may give new insight into Earth crust movements and deformations that may be important for the prediction of earthquakes.

4. SORT : Solar Orbit Relativity Test

The T2L2 experiment associated with an ultra-stable clock orbiting around the Sun, as proposed by the SORT experiment, allows the measurement of the Shapiro delay [16]. This delay $\propto (1+\gamma)$ where γ is the PPN parameter characterising the space curvature produced by unit rest mass, will be maximum near the occultation with a variation as large as 100 µs in a few days allowing an estimation of γ with an accuracy of 10^{-7}. That corresponds to an improvement of four orders of magnitudes compared to the radio-electric deflection estimations [17] limited by the solar corona noise. This γ estimation will bring a new confrontation of the different gravitation theories at this accuracy level. For example, the general relativity predicts a γ value equal to 1, and the scalar tensor theories express this PPN parameter as : $\gamma = \frac{1+\omega}{2+\omega}$ where ω is an arbitrary coupling function that determines the strength of the scalar field. In the Bergmann-Wagoner-Nordtvedt formulation [18, 19, 20], the ω function could be very large so that the theory prediction are almost identical to general relativity today. But, for the past or the future values of the scalar field, ω could take on values that would lead to significant differences in cosmological models.

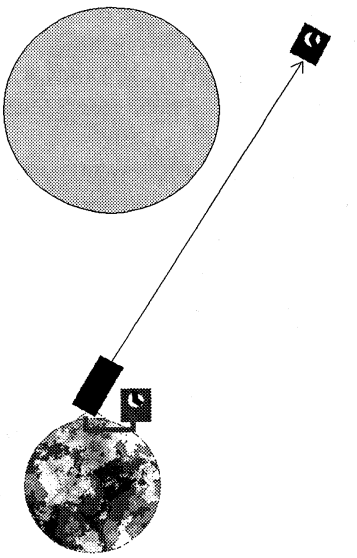


Figure 3 *Principle of the SORT experiment.*

5. Conclusion

The Time Transfer by Laser Link (T2L2), associated with ultra-stable clocks on Earth or in orbit, has many scientific applications in fundamental physics and astronomy. It should allow, for example, the test of the isotropy of the one-way speed of light, the measurements of the new ultra-stable clocks performance through their comparison on Earth or in orbit, the improvement of the VLBI observation, some studies in geodesy, and the improvement of the time scale taking into account the performances of the new ultra-stable clocks. T2L2 is also a basic element for future missions in the solar system (SORT) intending to test some aspects of gravitation.

Acknowledgements

The work of P. Fridelance was supported by the Société de Secours des Amis des Sciences, Paris, and the work of E. Samain was supported by the Centre National d'Etudes Spatiales (CNES).

References

[1] P. Fridelance, E. Samain, C. Veillet, *T2L2 - Time Transfer by Laser Link : A new optical time transfer generation*, to be published in Experimental Astronomy, 1997.

[2] P. Fridelance & E. Samain, *Réponse à l'appel à idées du CNES pour le vol longue durée sur MIR 99*, 1996.
[3] E. Samain & P. Fridelance, *The Time Transfer by Laser Link experiment (T2L2) on MIR*, these proceedings.
[4] C. Salomon & C. Veillet, *ACES : Atomic Clock Ensemble in Space*, proc. of the first symposium on the utilisation of the International Space Station, Darmstadt (Germany), 295-297, 1996.
[5] T. P. Krisher, L. Maleki, L. E. Primas, G. F. Lutes, J. D. Anderson, C. M. Will, *Results of a new test of relativity*, proc. of the XX^{e} Precise Time and Time Intervals Meeting, 251-260, 1988.
[6] P. Wolf : *Proposed satellite test of special relativity*, Phys. Rev. A, 51, No. 6, 5016-5018, 1995.
[7] A. Clairon, P. Laurent, G. Santarelli, S. Ghezali, S. N. Lea, M. Bahoura, *A cesium fountain frequency standard : preliminary results*, IEEE transactions on instrumentation and measurements, **44**, 128-131, 1995.
[8] M. Brunet, *DORIS precise orbit determination and location system performances of ultra-stable oscillators*, proc. of the 6th EFTF, Noordwijk (The Netherlands),125-136, 1992.
[9] R. Besson, LCEP/ENSMM, Besançon, France (personnal communication, 1996).
[10] T.P. Krisher, L. Maleki, G.F. Lutes, L.E. Primas, R.T. Logan, J.D. Anderson, *Test of the isotropy of the one-way speed of light using hydrogen-maser frequency standards*, Phys. Rev. D, **42**, No. 2, 731-734, 1990.
[11] P. Wolf , *Ph. D. thesis*, Queen Mary and Westfield College, London, in press.
[12] P. Wolf & G. Petit, *A satellite test of special relativity using GPS*, in these proceedings.
[13] R. Vessot, M. Levine, *Test of relativistic gravitation with a space-born hydrogen maser*, Phys. Rev. Lett., **45**, No. 26, 2081-2084, 1980.
[14] C.M. Will, *Gravitational redshift measurements as tests of non-metric theories of gravity*, Phys. Rev. D., **10**, 2330-2337, 1974.
[15] J.F. Lestrade, A.E.E. Rogers, A.R. Whitney, A.E. Niell, R.B. Phillips, R.A. Preston : *Phase-referenced VLBI observations of weak radio sources. Milliarcsecond position of Algol*, The Astronomical Journal, **99** (5), 1663-1672, 1990.
[16] C. Veillet, *SORT, Solar Orbit Relativity Test*, Proposal in the discipline area of fundamental physics in response to ESA's call for mission concepts for the follow-up to horizon 2000, 1994.
[17] D.E. Lebach, B.E. Corey, I.I. Shapiro, M.I. Ratner, J.C. Webber, A.E.E. Rogers, J.L. Davis, T.A. Herring, *Measurement of the solar gravitational deflection of radio waves using very-long-baseline interferometry*, Physical Review Letters, **75**, No. 8, 1995.
[18] P.G. Bergmann, *Comments on the scalar-tensor theory*, Int. J. Theor. Phys., **1**, 25-36, 1968.
[19] R.V. Wagoner, *Scalar-tensor theory and gravitational waves*, Phys. Rev., **D1**, 3209-3216, 1970.
[20] K. Nordtvedt, *Post-Newtonian metric for a general class of scalar-tensor gravitational theories and observational consequences*, Astrophys. J., **161**, 1059-1067, 1970.

High Q sapphire resonator oscillator development at LPMO

Vincent Giordano
Laboratoire de Physique et Metrologie des Oscillateurs
Universite de Franche-Comte
Besancon, France

Very high Q factors could be obtained in the microwave frequency range using monocrystal sapphire resonators associated with a special field configuration: the whispering gallery mode. Since 3 years, LPMO has engaged in research work about such high-Q resonators in the aim to design an ultra-low noise microwave source operating near 10 GHz.
A whole series of theoretical and experimental results have been yet accumulated. In the paper we will present the most significant of them:

- precise modelisation of the whispering gallery mode resonator with matching and 2-D finite element methods
- 2E5 Qo resonators at 300K
- 3E7 Qo resonators at 77K
- effective spurious mode suppression
- preliminary evaluation of phase noise performances

THE SHM SPACE BORNE HYDROGEN MASER FIRST EVALUATION OF THE PEM PHYSICS PACKAGE

Bernier L.G., Busca G., Jornod A., Schweda H.

Observatoire Cantonal, Neuchâtel, Switzerland

1. ABSTRACT

A compact Space borne Hydrogen Maser (SHM) frequency standard is being developed by the Observatory of Neuchâtel (ON). The development is sponsored by the Swiss PRODEX program of ESA. The SHM instrument will be flown on the Radioastron I (RA-I) mission in 1999 to be used as an alternate reference clock for orbital Very Long Base Interferometry (VLBI) and for the red-shift experiment CRONOS.

The Prototype Engineering Model (PEM) of the Physics Package (PP) is now under test. The first atomic signal was obtained in December 1996. This paper reports the first tests made on the PEM-PP for the verification of the design.

Keywords: Hydrogen Maser, Frequency Standards, Space Applications, Sapphire loaded microwave cavity.

2. THE SHM INSTRUMENT

2.1. Introduction

The ON is developing a compact hydrogen maser for space applications, based on a miniature sapphire loaded microwave cavity. The hydrogen maser is a high performance frequency standard capable of a frequency stability of about $1x10^{-15}$ over averaging intervals of 1'000s to 10'000 s. If used as a clock, this is equivalent to a time stability of 0.7 ps over a 1000 s time interval or 7 ps over a 10'000 s time interval. The SHM instrument will fly on the international RA-I orbital VLBI mission to be used as the alternate master clock. The SHM instrument will also be used for an accurate measurement of the red-shift effect, i.e. the time dilatation due to the gravitational field. This is the purpose of the CRONOS experiment proposed by ON and sponsored by the Swiss PRODEX program of ESA.

2.2. The Radioastron I Mission

The RA-I payload is a space borne VLBI radiotelescope. The achievable baseline is determined by the apogee of the RA-1 orbit which is 78'000 km. This is to be compared with the continental baseline, of the order of 10'000 km, of ground based VLBI. The RA-I mission was designed around a phase transfer scheme in which a ground based hydrogen maser remotely controls the embarked reference clock using a two-way microwave link [1], [2]. The remotely controlled clock is used to drive the VLBI experiments. A space borne rubidium frequency standard is used as the reference clock during periods of non-visibility of the spacecraft from the tracking ground stations. The addition of the SHM instrument to the original RA-I payload offers the alternative to drive directly the VLBI experiment with a local hydrogen maser, yielding the best performance at all times, including during non-visibility periods.

2.3. The CRONOS Experiment

In addition to its use as the VLBI reference clock, the SHM instrument will be used to perform the CRONOS experiment proposed by ON. CRONOS is a general relativity experiment whose goal is to measure the red-shift effect with a 10^{-6} accuracy. The red-shift effect is the time dilatation produced by the gravitation field. The expected accuracy of 10^{-6} is made possible by the highly elliptic orbit of RA-I and by the high stability of the SHM clock.

2.4. Future Applications

The technological demonstration of the successful operation of a hydrogen maser frequency standard in orbit also opens the way to many other future applications in VLBI, navigation, ranging, and in precise time transfer and dissemination.

2.5. Development Steps

The SHM compact hydrogen maser frequency standard is based on a Miniature sapphire loaded Microwave Cavity (MMC) designed by ON in 1993 under the first ESTEC contract for the development of the SHM instrument [3], [4].

Originally the foreseen first mission for the SHM instrument was the EXTRAS experiment on a Meteor 3M Russian satellite. This experiment was proposed by the Earth Observation Preparatory Programme (EOPP) of ESA.

The design of the SHM instrument was started and the first tests of a MMC were performed in 1994. Breadboarding activities around the hydrogen getters were started [5].

After the cancellation of EXTRAS, an agreement was made with the Astro Space Centre (ASC) in Moscow to fly the SHM instrument on the Radioastron I mission. The accommodation study of the SHM instrument was performed by ASC in 1995. In January 1995 the CRONOS experiment was submitted by ON to the Swiss PRODEX program of ESA and subsequently accepted.

In 1995 the design of the SHM instrument was continued. Various bread-boarding activities were continued or started around the sapphire loaded microwave cavity, the hydrogen getters, the magnetic shields, the receiver and other critical elements of the design. The main design parameters of the maser physics (cavity and atomic quality factors) were experimentally confirmed. Vibration tests of critical sub-assemblies were performed.

During the second half of 1995 and the first half of 1996 all the contractual and technical requirement documents were prepared for the industrial part of the CRONOS project. The industrial contracts cover the engineering analysis activities and the electronics package. ON is responsible for the PP.

The development by ON of the SHM-PP for the CRONOS experiment on Radioastron is financed under a PRODEX institute agreement since July 1995.

In 1996 the detailed design of the PP was finished. The bread-boarding and the vibration tests of critical sub-assemblies were continued. The PEM-PP was manufactured during the first half of 1996 and integrated in October and November. The first atomic signal from the integrated PEM-PP has been obtained in December 1996.

Present plans are to fine-tune the PP and to perform all the design verification tests during the first half of 1997. This will be done in parallel with the industrial engineering analysis contract which was kicked-off in January 1997. The electronics package contract will be performed after the final tests on the PEM-PP.

3. SHM INSTRUMENT DESIGN

3.1. Condensed Specifications

- long-term drift : $< 3 \times 10^{-12}$ year^{-1}
- output levels : 0.4 Vrms in 50 Ω
- frequency stability

τ [s]	Allan deviation
1	$< 1.5 \times 10^{-13}$
10	$< 2.1 \times 10^{-14}$
100	$< 5.1 \times 10^{-15}$
1'000	$< 2.1 \times 10^{-15}$
10'000	$< 1.5 \times 10^{-15}$

Table 1
Short-Term Stability

- phase noise

	14.71 MHz output	5 MHz output
f [Hz]	L (f) [dBc]	L (f) [dBc]
1	< -100	< -110
10	< -122	< -132
100	< -132	< -143
1'000	< -141	< -151
10'000	< -145	< -155

Table 2
Phase Noise

- operating temperature range: 10°C to 35°C
- thermal sensitivity: $\leq 3 \times 10^{-15}$ K^{-1}
- magnetic field range: ± 1 Gauss
- magnetic sensitivity: $\leq 2 \times 10^{-14}$ Gauss^{-1}
- DC input voltage: 22 to 50 V
- power consumption: < 70W
- largest diameter (horizontal) 460 mm
- full height (vertical) 600 mm
- mass 50 kg

3.2. Maser Physics Design

The heart of the PP design concept is the sapphire loaded MMC. The internal volume of the MMC is 4.4 litre. This is to be compared with the 20 litre of a conventional circular microwave cavity tuned to the hydrogen hyperfine frequency. The storage volume for hydrogen is 1.7 litre which is comparable to a conventional design based on a full size cavity. It is basically the size reduction of the microwave cavity that has made possible the realisation of a 50 kg SHM instrument. The sapphire cylinder of the MMC has both a microwave function, the dielectric loading of the cavity, and a maser physics function since, together with the bonded titanium covers, it also constitutes the hydrogen storage bulb.

State selection is performed by a conventional quadrupole magnetic state selector. The dissociator is a small fused quartz bulb. The hydrogen supply is stored in the form of a solid-state hydride material. Magnetic shielding is performed by a set of four 0.5 mm magnetic shields around the MMC and by a fifth 0.5 mm magnetic shield that surrounds the whole instrument.

There are two vacuum systems : the hydrogen vacuum enclosure and the thermal vacuum enclosure. Each enclosure is pumped by a set of getters and by a set of two 2 l/s ion pumps. For each enclosure, one ion pumps would be sufficient. The second pump is there for redundancy. The sealed part of the PP is closed by an aluminium bell jar that closes the thermal vacuum enclosure.

3.3. Mechanical Design

The design load for structural design is 20 *g*. The SHM instrument is designed to pass the Radioastron qualification vibration levels.

3.4. Thermal Design

In orbit the instrument is covered by a MLI thermal insulation blanket. All thermal exchanges are performed by conduction to a dedicated thermally controlled base plate which is part of the spacecraft.

The temperature of the MMC is controlled with a design goal thermal stability of $\sigma_T(10'000s) \leq 7mK$. This is achieved by three concentric layers of active thermal control around the MMC.

The thermal coefficient of the MMC TE_{011} resonant frequency is 65 kHz/°C. With such a thermal coefficient it is not possible to achieve the specified thermal coefficient of the instrument with thermal control of the MMC resonant frequency only. An Automatic Cavity Tuning (ACT) system in necessary in order to reduce the cavity pulling effect by an additional electronic frequency control of the MMC.

3.5. Electronics Design

The ACT system is the most important performance critical element of the electronics. Its purpose is to stabilise the resonant frequency of the microwave cavity in order to minimise the cavity pulling effect on the atomic signal frequency.

The design goal thermal coefficient of the PP is $6x10^{-12}$ K^{-1} with only thermal control of the MMC resonant frequency and $3x10^{-15}$ K^{-1} with the ACT system in operation, i.e. with electronic frequency control of the MMC. In other words the ACT system must improve the thermal coefficient by a factor of 2000 with respect to thermal control alone.

4. EVALUATION OF PEM PHYSICS PACKAGE

The figures reported below are the first results obtained this January, right after integration of the PEM-PP. Debugging and optimisation of the bread-board electronics will be necessary before it is possible to perform the final design verification tests of the PEM-PP. The full set of design verification tests will be performed during the first half of 1997.

4.1. Microwave Cavity

The measured loaded quality factor of the MMC is 34'000. The thermal coefficient of the TE_{011} mode is -65 kHz/K.

4.2. Atomic Quality Factor

The operating atomic quality factor at -104 dBm is $1.5x10^9$.

4.3. Thermal Control

The measured thermal stability of the MMC is $\sigma_T(10'000s) \leq 0.2mK$, i.e. 35 times better than the design goal.

4.4. Automatic Cavity Tuning

With only a thermal control of the MMC the measured thermal coefficient is $5x10^{-12}$ K^{-1}. On the other hand, the thermal coefficient becomes $3x10^{-13}$ K^{-1} with the bread-board ACT in operation. This is far from the specification and further debugging of the ACT is necessary.

The measured frequency stability with the bread-board ACT in operation is $\sigma_T(1'000s) \leq 3\times10^{-15}$. This is already very close to the specification.

4.5. Magnetic Shielding

The measured shielding factor with only the four cavity shields is 20'000. A global shielding factor of 200'000 is expected with 5 shields.

4.6. Getters & Ion Pumps

It was verified that the vacuum system has an autonomy of at least 10 days without electrical power to the ion pumps (getters only). This requirement is specific to the Radioastron I mission.

5. CONCLUSION

The evaluation of the PEM SHM-PP reported in this paper is the first step of the full set of design verification tests to be performed during the first half of 1997. The measured parameters are close to the design goal values. Only minor changes of the SHM instrument specifications are expected following the full testing of the PEM-PP.

We are confident that once the demonstration is made that the SHM instrument is able to deliver in orbit the same level of performance as a conventional ground maser but with a mass of only 50 kg, other space mission opportunities will open beyond the Radioastron first mission.

6. REFERENCES

[1] Valtaoja E., "The Radioastron Space VLBI project", *Proc. 4th European Frequency and Time Forum*, Neuchâtel, 13-15 March 1990, pp. 707-710.

[2] Andreyanov V.V., "Time-Frequency Synchronization in Radioastron OVLBI Project", *Proc. 7th European Frequency and Time Forum*, Neuchâtel, 16-18 March 1993, pp. 33-38.

[3] Bernier L.G., "Preliminary Design and Breadboarding of a Compact Space Qualified Hydrogen Maser Based on a Sapphire Loaded Microwave Cavity", *Proc. 8th European Frequency and Time Forum*, Munich, March 9-11 1994, pp. 965-980.

[4] Couplet C., Bernier L.G., "Proton Irradiation Test of the Teflon Wall Coating in the Atomic Hydrogen Maser in Preparation to Space Applications", *Proc. 8th European Frequency and Time Forum*, Munich, March 9-11 1994, pp. 952-958.

[5] Bernier L.G., "Preliminary Design and Breadboarding of a Compact Space Qualified Hydrogen Maser Based on a Sapphire Loaded Microwave Cavity", *Proc. 5th Time & Space Metrology Symposium*, Mendeleevo, Russia, 11-13 October 1994, p. 412-417.

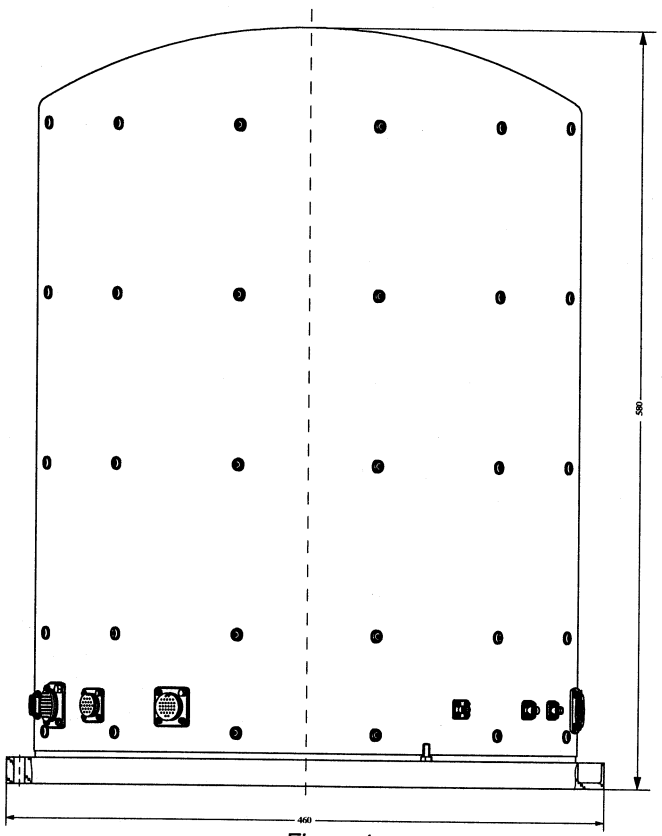


Figure 1
Mechanical Interface Control Drawing

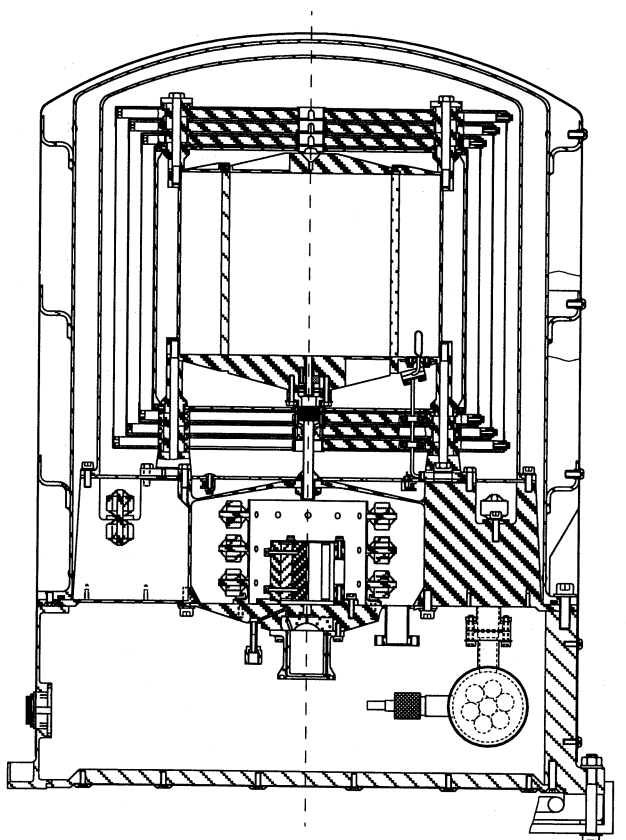


Figure 2
Top Assembly Drawing

H-MASER FOR SPACE APPLICATIONS

Gaigerov B.A., Domnin Y.S., and Yunoshev L.S.

Institute of Metrology for Time and Space (IMVP GP VNIIFTRI) Russia

ABSTRACT

H-maser of a small and rigid design is considered as a prototype of a spaceborn clock for global time comparison. The dimensions of the maser is mainly determined by a cavity. Partially loaded with a sapphire the cavity of the maser could be ten times less in the volume than a full-size cavity. The performances of the small maser are approximately the same as those of the full-size ones. Two points distanced by 5000 km and equipped with ground sets of masers and laser stations can be used as a reference for timing.

Keywords: maser, cavity, sapphire, accuracy, stability, laser link.

1. INTRODUCTION

If one relies on H-maser current development one can carry out an experiment on using a spacecraft portable clock to synchronise time scales in a global scale. Flying space clock (let it be FLOCK) can be compared with terrestrial clocks practically everywhere in the world many times per day. Typical procedure of time comparison between two terrestrial points via an old-fashioned portable clock takes a few days. The accuracy of the FLOCK time comparisons should be much higher than classical portable clock comparisons, which in it's turn are comparable with the best results of time comparisons via GPS or GLONASS navigation systems. There is no need to use common mode view to get an accuracy 0.1 ns by using FLOCK compared to GPS. (GPS common view mode lets to get accuracy in a few nanoseconds). To realise a common view mode the ground station should receive simultaneously the signals from one and the same satellite. This demand restricts essentially an availability of this method. In case of FLOCK there is no need to know the co-ordinates of a satellite and hence any user can extract all potential ability of a comparison equipment and one can carry out independent measurements and make evaluations of various effects on the time scale synchronisation.

The most precise channel of comparison is a laser link, but one should match the laser link ability with H-maser performances. The laser link does not allow to measure H-maser short-term stability (for time intervals of 1-100 sec) with the proper accuracy but it has adequate or better resolution for time intervals of the order of 3000 sec or more. If we accept that we can reliably measure laser pulse with the resolution of 10 picosecond we shall have relative resolution $10^{-11}/\tau$. On the other hand our portable H-maser "Sapphire" has the stability $\approx 3\times10^{-13}/\sqrt{\tau}$. So the performances of the laser channel and these of the H-maser are met at averaging time about 1000 sec. The accuracy is limited by the laser link at averaging times less than 1000 sec, and it is limited by the H-maser with greater averaging times.

2. H-MASER

For the last thirty years the hydrogen masers serve as the most stable oscillators. The stability of the best H-masers is characterised by variations value as $(3\text{-}5)\times10^{-13}$ for 1 s and $(1\text{-}3)\times10^{-15}$ for time intervals of one day and more.

The dimensions and the weight of the H-masers are essentially determined by the cavity dimensions and this restricts their applications as portable oscillators of high stability.

One of the ways to decrease the dimensions of the H-maser is to use the cavity, partially filled with a dielectric of a high dielectric permitivity and small losses. The cavity can have any shape. It is essential, that the oscillation modes should have a closed electrical line of force. If one fills a region of high frequency field by dielectric, separated from the cavity walls, one can decrease essentially the dimensions of the cavity. The Q-factor of the cavity can be as high as for full-size one or even higher and that ensures stable H-maser oscillations. This kind of the H-maser was proposed in (Ref. 1,2). The most simple design is the cylindrical axially symmetrical one. In this design a hollow dielectric cylinder covered by teflon inside is used as a storage valve. We managed to succeed in getting stable H-maser oscillations with the cavity which has ten times less volume than the full-size one. It is important, that the mechanical properties of the new cavity are much higher than those for a full-size one. One can see it easily from Fig.1, where a full-size and compact designs are presented on a scale. Even if one does not go into details of a fixation of the storage valve in the full-size cavity one may come to the conclusion that the mechanical properties of the compact cavity are much higher than those for the full-size cavity. The thickwall storage valve of the compact cavity is manufactured of the monocrystal of sapphire. It is pressed between the covers of the cylinder cavity and this improves the mechanical properties too

The final choice of the cavity dimensions is determined by a trade-off between the physical dimensions and Q-factor of the line. The Q-factor of the line depends on the volume of the storage valve,

so the storage valve determines the cavity dimensions. In our practical design we have chosen a typical storage valve with the volume of 0.9 l. In this case we can realise the line Q-factor as high as 1.5×10^9.

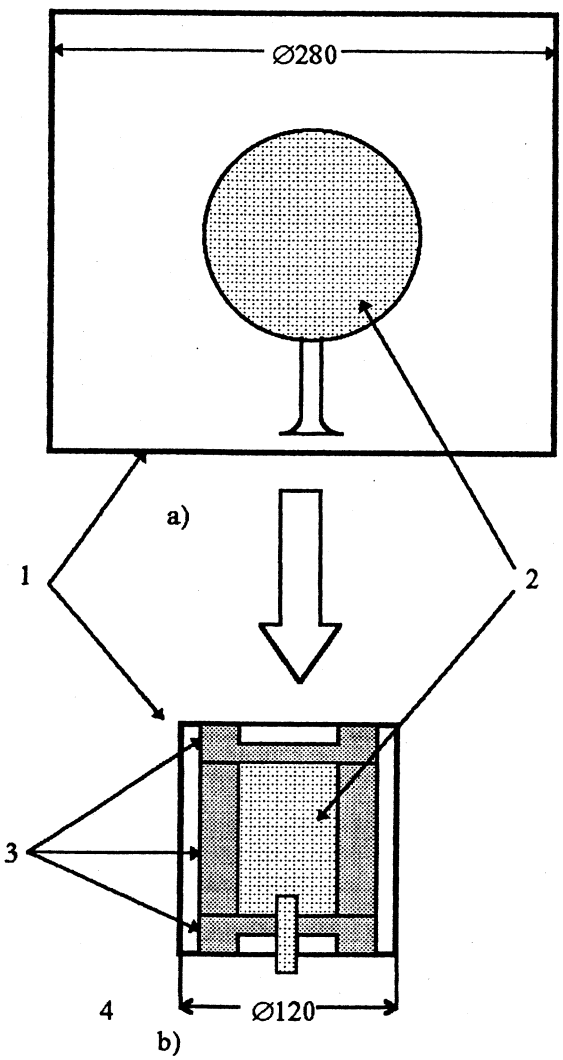


Fig.1
a) - Full-size cavity, b)- compact cavity partially filled with sapphire
1-cavity, 2 -storage valve, 3-sapphire

The smallest cavity was manufactured and the maser oscillations took place. This cavity has the following parameters. The cavity dimensions were Ø120x160 mm, the sapphire storage valve height was 140 mm and external and internal diameters were 66 and 38 mm correspondingly. The calculated Q-factor was 61 000 and experimentally measured one was 55 000.

In a practical design we use the cavity made of titanium with an internal diameter of 174 mm and a height of 180 mm. At the centre of the cylinder a sapphire storage valve of 0.9 l is placed. Its internal surface is plated by a teflon. The flexible joint is used to connect a storage valve neck to a vacuum chamber. The basement of the cavity is fixed to a basement of a thermostat and electrically isolated from it. Inside the cavity the surface is plated by a silver. The cavity Q-factor is about of $(45\text{-}50) \times 10^3$.

To stabilise the temperature we use a two-stage thermostat. The internal stage is placed on the external surface of the cavity with a thread for heater. It maintains the temperature within the limit of ± 0.01 °C. The working temperature is of 50°C. The external stage maintains the temperature of an aluminium cylinder within the limit of ± 0.1°C. The temperature of the cylinder and its upper and bottom covers is controlled separately.

To protect from an external magnetic field four magnetic shieldings are used. Their total gain is about of 1×10^5. An operating field is of 1×10^{-2}A/m. Besides an active magnetic field stabilisation is used with a gain of not less than 100. The external magnetic screen is used as a vacuum cover for the cavity. Its bottom basement is used as the main basement of the maser.

To prevent from a frequency shift due to a magnetic field inhomogeneity we use a periodical inversion of a population of the F=1, m_F=1 and F=1, m_F=-1 levels. It is done by an inversion coil which creates an alternative magnetic field.

A vacuum system consists of a source of hydrogen atoms, a chamber with a selecting magnet and a chamber with a getter and ion pump to get the necessary vacuum in the region of the storage valve. The getter pump is made of cylinder with a porous titanium. It effectively works to pump the hydrogen at the temperature (0-40)°C during five years. The ion pump removes the other gases.

One of the serious drawbacks of this portable maser is a high temperature dependence of the cavity frequency. It consists of 50 kHz/K due to the temperature performance of the sapphire. In order to maintain H-maser frequency within 1×10^{-14} it is necessary to maintain the cavity temperature within 2×10^{-6} K. It seems practically unrealisable. One may stabilise the cavity frequency with the help of AFC system. Our system of automatic cavity tuning enables us to get the stability as high as 1×10^{-15} (Ref.3).

Investigations of a maser frequency change due to the cavity frequency change were carried out. It was shown, that for small variations maser frequency comes to an equilibrium with a time constant $\delta = 2Q_c/\omega_o \approx 2 \times 10^{-6}$s. We can evaluate vibrations influence on the maser frequency in quasistatic approximation. Experimentally shown, that a static force equivalent 1g acceleration leads to the cavity frequency change of 100 -150 Hz. So if we have vibrations with 0.1g we shall have the

maser frequency modulation with a depth of $(\Delta f/f_0) = (\Delta f_c/fc)x(Qc/Q_l) \approx 2x10^{-13}$. For an averaging time of 1 s the frequency instability will be of $2x10^{-14}$ and for 10 s - $2x10^{-15}$ correspondingly.

The quantum clock developed on the base of the H-maser have the following main characteristics:

Dimensions (mm)	480x330x530
Mass (kg)	60
Life time (hours)	50 000
Power consumption (W)	50
Operating temperature	(0÷40)°C
Temperature frequency shift	$3x10^{-15}$/K
Magnetic field frequency shift	$4x10^{-15}$A/m
Frequency stability	$\sigma(\tau)=3x10^{-13}\sqrt{\tau}$
The highest stability	$1x10^{-15}$

3. GROUND SEGMENT FOR TIMING

IMVP is in charge not only for time and frequency in Russia, but for the determination of the Earth rotation parameters too. In particular, state T&F service has a sets of H-masers and Laser Ranging Stations (LRS) at Mendeleevo and Irqutsk, At Mendeleevo we have national T&F standard and at Irkutsk a spare T&F standard. They are distanced approximately by 5000 km. Because they are equipped with the time standards and LRS they can be compared not only with the high resolution, but with the highest accuracy too. Now we have LRS with a ranging accuracy of 15 cm, but in 1997 it is planned to modify LRS to get an accuracy of 2-3 cm. That means that we may have potential accuracy of 0.1 ns.

Mendeleevo and Irqutsk time scales are compared continuously via GPS and GLONASS. Sometimes they are calibrated with the help of a portable clock. At present the uncertainty of one day comparisons is near of 3 ns. So. Mendeleevo-Irqutsk can serve as "the time baseline" to test the comparison accuracy via FLOCK. In the nearest future we intend to use phase measurements and we hope to decrease the comparison uncertainty to the level of 0.1ns. If the one day stability of the FLOCK H-maser is of $1x10^{-15}$, the accuracy of the one day comparisons via FLOCK will also be of 0.1 ns.

4. Ifs, POSSIBILITIES AND THE MEANING OF THE EXPERIMENT

There are a lot of **ifs** which restrict or expand experimental facilities. Even if we have a space clock and a well equipped ground station we cannot make any experiment **if** we do not expand the spacecraft equipment. First of all it is necessary to provide a link between the space and ground clock. So the board should be equipped with a microwave antenna and a microwave set to receive, process and transmit the signals. A laser corner reflector and a pulse counter should be installed to realise the laser link.

If we add the abovementioned equipment we can carry out the timing experiments. A successful experiment will demonstrate the readiness of new technology to manufacture a new type of the space clock, based on the H-maser, one of the most stable oscillators. In terrestrial environments our compact design H-maser has a lifetime of 50 000 - 100 000 hours. In space environments new factors may appear, which will restrict the lifetime or deteriorate the stability.

If it works we can make time scale comparison on a global scale and with the highest accuracy. It will be an independent instrument to test the accuracy of the satellite navigation systems. The users can be supplied with the highest time synchronisation and dissemination to carry out experiments in metrology, geodesy, navigation, communication and for fundamental physical experiments.

If we add one more space clock and on board comparison equipment we can study a clock behaviour for short and medium time intervals and we can separate the technical frequency shifts from the frequency shifts due to external factors.

If we add space GPS or GLONASS receiver we can study ionosphere effects.

If we add a broadcasting transmitter we can develop universal time scale and synchronize all clock all over the world with an accuracy of a few milliseconds.

Ifs can stop the experiment or they can permit to carry out a lot of attractive experiments.

5. REFERENCES

1.Gaigerov B.A., Pushkin S.B., and Rusin F.S., Patent №1238184, "Vodorodniy generator", Priority of March 5, 1984.

2.Gaigerov B.A., "Hydrogen maser with a compact resonator", Izmeritelnaya Technica, №6, p.8, 1986

3.Gaigerov B.A., Elkin G.A. "Automatic tuning of a hydrogen maser", Izmeritelnaya Technica, №6, p.90, 1968.

SESSION 2

Chairman: Ph. Laurent

International Space Station Microgravity Applications Promotion Programme

H.U. Walter, ESA - ESTEC/ Directorate of Manned Spaceflight and Microgravity

The overall strategic objective for the Microgravity Applications Promotion is to generate a European activity using the ISS as a facility for industrial R & D. The ISS has the advantage of large resources (pressurised laboratory volume, external mounting surface, power, data transmission, up- and download), easy and relatively high frequency access, furthermore the continuous presence of astronauts for observations, experimentation, repair, servicing and exchange/upgrading of modular facilities. Laboratory-type experimentation is feasible, furthermore one may employ teleoperations and monitoring from the ground, as well as fully automated investigations.

"Applications" of microgravity may be understood as the exploitation of the ISS for applied research and as a testbed for the development of technology and processes useful on Earth and for long-duration space flight. The objective is therefore to develop projects in order to:

- optimise applied ground-based physical processes
- generate data and materials samples relevant for industrial R & D
- investigate space-related physiological changes relevant for long-duration space flight, which may also be relevant for clinical applications on Earth.

The Microgravity Applications Promotion Programme has the objective to foster and to develop a first generation of projects with involvement from industry, which should demonstrate that microgravity is indeed a useful tool for industrial R & D. These projects will be prepared by precursor experiments using traditional carriers and opportunities in the early utilisation of the ISS.

The overall budget available for Microgravity Applications Promotion up to 2003 is as follows:

* Workshops/Expert Studies	2.5 MAU
* Application Projects/User Teams	14.1 MAU
* Applic. Payl. Studies & Develop.	33.5 MAU
* Prep. Exper. & Early Utilis.	17.1 MAU
	67.2 MAU

This budget breakdown is tentative, spending will depend on the response by the users. Substantial funding is available for payload studies and developments and also for preparatory experiments during the Early Utilisation of the ISS and eventually using other microgravity carriers.

In principle the Atomic Clock Ensemble in Space (ACES) Project would qualify for this Programme. There are perspectives, both for research and terrestrial applications, the European dimension can certainly be developed.

FUNDAMENTAL RELATIVISTIC METROLOGY ON THE SPACE STATION

Alessandro D.A.M. Spallicci *
Univ. Salerno, sede di Benevento, Fac. Ingegneria, Gruppo Gravitazione †
Leiden Rijksuniversiteit, Fac. Natuurkunde & Sterrenkunde, Kamerlingh Onnes Lab. ‡

Abstract

The Space Station circular orbit may provide valid measurements of the first order Doppler asymmetry term, the Sagnac effect and the gravitation null shift. Special and general relativistic effects upon atomic clocks are to be taken into account also for time generation distribution and navigation, when dealing with the GPS system.

Keywords: Special and General Relativity, Time & Frequency, Space Station

1. Introduction

Beside the in-house research work, activities on fundamental physics via T&F measurements on the Space Station (henceforth SS) started in 1990 [1] via the establishment of the Columbus Metrology and Gravitation Science Team (CMGST) [2]. Later ESA initiated an industrial study with the participation of several institutes [3]. Several papers were published [4-6]. About 12 experiments in fundamental physics were identified, but including also mechanical measurements of small forces.

This workshop aims at the scientific applications for T&F with SS and this paper deals solely with the fundamental physics measurements that can be carried out with SS in the next years. For applications in other disciplines or future use of clocks, the reader is referred to other papers in these proceedings.

2. First order Doppler asymmetry term

SS circular orbit is not optimal for measurements such as time dilation and gravitation shift, but other measurements of interest can be carried out [2,5]. We consider SS at a position $\vec{r}_1$, with a velocity $\vec{v}_1$, emitting a signal towards the Earth. This signal is first transponded by a station at a point $\vec{r}_g$ on the Earth, moving at velocity $\vec{v}_g$ and then received by SS which has reached the position $\vec{r}_2$ and has the instantaneous velocity $\vec{v}_2$. The frequency of the signal measured in the local reference frame at the positions $\vec{r}_1$, $\vec{r}_2$ and $\vec{r}_g$ is f_1, f_2 and f_{1g}, respectively. Working in a non rotating Earth centred coordinate system, the measured shift frequency is given by[1]

$$\Delta f = (f_{1g} - f_0) - \frac{1}{2}(f_2 - f_1) \tag{1}$$

where f_0 is the frequency of a terrestrial clock located at the Earth station (Es). The term $f_{1g} - f_0$ is a measurement made at the Es and $f_2 - f_1$ represents the 2-way Doppler shift measured on SS and transmitted in real time by data link to the Es. In presence of an offset δf between the SS clock and the one on the Es $f_0 = f_1 + \delta f$ the relative frequency shift is:

$$\frac{\Delta f}{f_1} = \frac{f_{1g}}{f_1} - \frac{\delta f}{f_1} - \frac{1}{2}\frac{f_2}{f_1} - \frac{1}{2} \tag{2}$$

The frequency f_{1g} of the SS signal at the Es is:

$$\frac{f_{1g}}{f_1} = \frac{1 - \vec{\beta}_g \cdot \vec{r}_{1g} + \frac{\phi_g}{c^2} + \frac{1}{2}\beta_g^2}{1 - \vec{\beta}_1 \cdot \vec{r}_{1g} + \frac{\phi_1}{c^2} + \frac{1}{2}\beta_1^2} \tag{3}$$

where $\beta_i = \vec{v}_i/c$, ϕ is the gravitational potential, and $\vec{r}_{1g} = \vec{r}_1 - \vec{r}_g$. The ratio f_2/f_1 is obtained from:

$$\frac{f_2}{f_1} = \frac{f_2}{f_{1g}}\frac{f_{1g}}{f_1} = \left(\frac{1 - \vec{\beta}_2 \cdot \vec{r}_{g2} + \frac{\phi_2}{c^2} + \frac{1}{2}\beta_2^2}{1 - \vec{\beta}_g \cdot \vec{r}_{g2} + \frac{\phi_g}{c^2} + \frac{1}{2}\beta_g^2}\right) \cdot \left(\frac{1 - \vec{\beta}_g \cdot \vec{r}_{1g} + \frac{\phi_g}{c^2} + \frac{1}{2}\beta_g^2}{1 - \vec{\beta}_1 \cdot \vec{r}_{1g} + \frac{\phi_1}{c^2} + \frac{1}{2}\beta_1^2}\right) \tag{4}$$

With the assumptions of negligible change of altitude $\phi_1 \simeq \phi_2$ and negligible change of velocity $\mid \beta_1 \mid \simeq \mid \beta_2 \mid$, *up to second order*, we finally get:

$$\frac{\Delta f}{f_1} = D^2 + DA^1 + DA^2 + GS - \frac{\delta f}{f_1} \tag{5}$$

where the time dilation is given by

$$D^2 = \frac{1}{2}\left(\beta_g^2 - \beta_1^2\right) \tag{6}$$

*This research work has been developed at the European Space Research & Technology Centre, ESTEC, Noordwijk, the Netherlands. Part of the results originated in the frame of the research cooperation with G. Busca, Obs. Neuchâtel, N. Ashby Univ. of Colorado at Boulder. Financial support from ESTEC for participation to this conference is acknowledged.

†Prof. Analisi Numerica, aspallic@ingbn.unisa.it
‡Visiting scientist, aspallic@rulkol.leidenuniv.nl

[1]This expression makes the difference between the 1-way measurement $f_{1g} - f_0$ and the 2-way measurement $f_2 - f_1$. The former uses independent clocks, while the latter makes use of a transponder which relays coherently the signal.

First and second order term in $1/c$ Doppler asymmetries (DA) are given by

$$DA^1 = -\frac{1}{2}\left[\left(\vec{\beta}_1 - \vec{\beta}_g\right)\cdot\vec{r}_{1g} + \left(\vec{\beta}_g - \vec{\beta}_2\right)\cdot\vec{r}_{g2}\right] \quad (7)$$

$$DA^2 = \left(\vec{\beta}_1\cdot\vec{r}_{1g}\right)^2 - \left(\vec{\beta}_1\cdot\vec{r}_{1g}\right)\left(\vec{\beta}_g\cdot\vec{r}_{1g}\right) - \left[\left(\vec{\beta}_g - \vec{\beta}_2\right)\vec{r}_{g2}\right]^2 \quad (8)$$

The gravitation shift is expressed by

$$GS = \frac{\phi_g - \phi_1}{c^2} \quad (9)$$

For a circular orbit at 350 km, the first order Doppler asymmetry term shows a high dynamic change and reaches peaks of the order of 10^{-9}. It can be measured with an accuracy of e.g. $1\times10^{-15}/1\times10^{-9} = 1\times10^{-6}$ or even one order of magnitude higher with better clocks. The gravitational shift and the time dilation are nearly constants due to the circular orbit; this situation differs greatly from the GP-A-1976 experiment of Vessot and Levine where these effects changed of orders of magnitude due to the ballistic orbit, while DA^1 was minimal. The second order Doppler asymmetry term is too small for measurements. A specific study should be financed to analyze the potential for a test on the isotropy of velocity of light, the Lense-Thirring effect, the test of the Local Position Invariance and other reference frames effects, with implications on the Equivalence Principle. Further considerations are to be made for the correspondance of the first order Doppler asymmetry term with the "acceleration" term identified by Vessot.

3. Sagnac effect

The Sagnac effect is a special relativistic effect due to finite light velocity in rotating frames. It is proportional to the area swept out by the radius vector from the Earth centre to the electromagnetic pulse as it propagates from the point of transmission to the point of reception (A, B respectively). It is expressed by

$$\Delta t = \frac{2}{c^2}\omega\cdot\vec{A} = \frac{2}{c^2}\omega\cdot\left(\frac{1}{2}\vec{r}_A\times\vec{r}_B\right) \quad (10)$$

An Earth station, at a defined time, i.e. the occurrence of the zero of the 1-way Doppler shift, makes a measurement of the time offset and frequency offset between its local maser and the space maser via the T&F and applies the relativistic corrections considered above. One orbital period later, again at the zero 1-way Doppler shift, the same comparison takes place. The Sagnac effect expected is shown in the table.

Light time Es-SS	2×10^{-3} s
Light time around the Earth	1.46×10^{-1} s
Total light time	1.5×10^{-1} s
Motion of Es during light time	69.9 m
Sagnac effect	233.1 ns

For a time error of 30 ps, the Sagnac effect could be measured with an accuracy of $3\times10^{-11}/2.33\ 10^{-7} \simeq 10^{-4}$ which is about two orders of magnitude better than currently obtained by near Earth satellites.

4. Gravitation null shift test

Soffel [3] suggested to make a gravitation null shift test, making advantage of the ACES proposal. The gravitation shift is defined as

$$\frac{\Delta f}{f} = 1 + \alpha\frac{\Delta\phi}{c^2} \quad (11)$$

where $\alpha = 1$ for general relativity. Measurements with two different clocks would lead to a formulation allowing a null test:

$$\frac{\Delta f_a}{f_a} - \frac{\Delta f_b}{f_b} = \Delta\alpha\frac{\Delta\phi}{c^2} \quad (12)$$

5. Relativistic corrections on a GPS space receiver

Relativistic corrections on a GPS space receiver have been analyzed [6]. Each GPS unit's clock is slowed by 4.465×10^{-10} for time dilation and gravitational frequency shift before launch. This cancels the main constant relativistic effects for terrestrial, including maritime and aeronautical users, who have only to process the GPS units'orbital eccentricities, a sinusoidal type function, and possibly the Sagnac effect. For a spacecraft, its velocity, especially if in lEo, its location in the Earth gravitational field, the effect of drag and the orbital eccentricity of the spacecraft determine a largely different effect than for terrestrial users. Therefore the need for proper corrections for time distribution and navigation[2] for GPS space receivers is evident. There are four different times, the coordinate time and the proper times at the Earth station, on board the SS, and on board the GPS unit. The final simplified expression relating the time on SS with respect to coordinate time is (ECI coordinates) is:

$$dt_{SS} \sim \left(1 - \frac{v_{SS}^2}{2c^2} + \frac{GM_\oplus}{R_\oplus c^2} - \frac{GM_\oplus}{r_{SS}\,c^2}\right)dt_c \quad (13)$$

where the second term on the right hand side represents the time dilation, while the third and the fourth the gravitation shift. In particular, r_{SS} and v_{SS} are the SS position and velocity, G is the constant of gravitation, $M_\oplus$ and $R_\oplus$ are the Earth mass and radius, respectively. For ISS at about 350km, (13) provides

[2] 1 $msec$ error is about 7.7m, for a 350km satellite.

a time error of the order of -3×10^{-10}, i.e. 9 ms/y, while a clock on-board GPS gains about 14 ms/y, with respect to coordinate time. These quantities are not negligible, five orders of magnitude better than maser clocks stability, and greater even than best quartz clocks. Besides, the sign of the offset changes according to altitude, since time dilation and gravitational red-shift have opposite sign. A lEo is mainly influenced by time dilation due to its high speed, while a geostationary satellite would record an offset only due to gravitational red-shift[3]. The corrections for satellites due to receiver motion during signal propagation are expressed by

$$\Delta t = \frac{\vec{v}_{SAT}}{c^2}(\vec{r}_{SAT} - \vec{r}_{GPS}) \qquad (14)$$

(14) is of relevance when the SS receivers are synchronized directly to GPS, without making use of SS clocks.

6. Conclusions

It is a common satisfaction to all the participants of the scientific team of CMGST and industrial study to have contributed to the birth of experiments as ACES, and to the further definition of experiments as PHARAO, T^2L^2 and masers in space.
The activities on T&F have shown a large interest from the scientific community of several institutes of ESA member states.
Interesting measurements can be carried out for fundamental physics as outlined above.
ESA and national space agencies should support financially activities such as basic and applied sciences of T&F.
High level science experiments will promote the image of SS as a laboratory for fundamental discoveries.
In any case the test of clocks on board SS is essential for future missions in deep space or more accurate near Earth orbit trajectories.

7. References

[1] Spallicci A., 1993. *Advanced Series of Astrophysics and Cosmology*, 7, 505, Proc. I.W. Fairbank mtg on Relativistic Gravitational Experiments in Space, 10-14 September 1990 Roma, M. Demianski and C.F.W. Everitt Eds., World Scientific.

[2] *Proceedings of Columbus Metrology Science Team*, 26 October 1991 and of 2-3 September 1991, ESTEC, Noordwijk, K. Knott, I. Roxburgh, A. Spallicci Eds; *The use of Columbus for experiments in gravity science* 1992. Roxburgh I., Busca G., Fuligni F., Nobili A., Spallicci A., 31 August 1992.

[3] ESA-DASA, 1995. *Time & Frequency Science Utilization Study and Space Station*, Study manager E. Graf, Study scientist A. Spallicci. Prime: DASA. Institutes: Deutsche Geodätisches Forschungsinstitut, München; DLR, Deutsche Forschungsanstalt für Luft und Raumfahrt Oberpfaffenhofen; Techn. Univ. München; Ecole Normale Superieure, Paris; Techn. Univ. of Dresden & Lohrmann Obs.; Inst. of Navigation, Univ. Stuttgart; Observatoire de la Cote d'Azur, CERGA, Grasse. ESA Contract 11287/94/NL/VK, Executive summary and Final report; in submission to Aerotecnica Missili e Spazio.

[4] Spallicci A., 1991. *9th Proc. of It. Conf. on General Relativity and Gravitational Physics*, 611, 25-28 September 1990 Capri, R. Cianci, R. de Ritis, M. Francaviglia, G. Marmo, C. Rubano and P. Scudellaro Eds., World Scientific; Spallicci A., 1992. *6th European Frequency and Time forum*, 67, 17-19 March 1992 Noordwijk, ESA-SP-340; Spallicci A., Brillet A., Busca G., Fuligni F., Nobili A., Roxburgh I., 1993. *Class. Quantum Grav.*, 10, S259, special issue for Journees Relativistes Conf., 13-15 May 1992 Amsterdam; in Bertotti B., 1993. *Proc. 13th Internat. Conf. on General Relativity and Gravitation*, 416, 28 June - 4 July 1992 Huerta Grande, Cordoba (Argentina), R.J. Gleiser, C.N. Kozameh, O.M. Moreschi Eds., IOP; Spallicci A., 1992. *10th Abstr. Proc. of It. Conf. on General Relativity and Gravitational Physics*, 1-5 September 1992 Bardonecchia; Spallicci A., Brillet A., Busca G., Fuligni F., Nobili A., Roxburgh I., 1996. *STEP symposium*, ESA WPP 115, 382, 6-8 April 1993 Pisa; Spallicci A., 1995. *Abstr. Proc. of 14th Internat. Conf. on General Relativity and Gravitation*, 6-12 August 1995 Firenze, Workshop C2: Space experiments and other experiments on gravitation, C.53; Spallicci A., 1995. *Abstr. Proc. of Symp. on Fundamental Physics in Space*, 16-19 October 1995 London, 48.

[5] Spallicci A., Busca G., 1996. *11th Proc of It. Conf. on General Relativity and Gravitational Physics*, 423, 26-30 September 1994 Trieste, M. Carfora, M. Cavaglià, M. Francaviglia, P. Frè, G. Pizzella, C. Reina, A. Treves Eds., World Scientific.

[6] Spallicci A., Jimenez C., Prisco G., Ashby N., 1992. *6th European Frequency and Time forum*, 61, 17-19 March 1992 Noordwijk, ESA-SP-340; Spallicci A., 1995. *Aerot. Miss. Sp.*, 74, 1-2, 41.

[3] For (13), at $r_{SAT} = 1.5R_{\oplus}$, or 3189 km altitude, relativistic effects compensate each other.

TIME & FREQUENCY SCIENCE UTILIZATION AND SPACE STATION

A. Spallicci [a,b] *, H. Drewes [c], J. Hahn [d], K.H. Ilk [e], C. Salomon [f], M. Soffel [g], K.-H. Thiel [h], C. Veillet [k] †, A. Wehr [h], E. Graf [i]

[a] *Univ. Salerno, sede di Benevento, Fac. Ingegneria, Gruppo Gravitazione*
[b] *Leiden Rijksuniversiteit, Fac. Natuurkunde & Sterrenkunde, Kamerlingh Onnes Lab.*
[c] *Deutsche Geodätisches Forschungsinstitut, München*
[d] *DLR Deutsche Forschungsanstalt für Luft und Raumfahrt, Oberpfaffenhofen*
[e] *Techn. Univ. München*
[f] *Ecole Normale Superieure, Paris*
[g] *Techn. Univ. Dresden & Lohrmann Obs.*
[h] *Inst. of Navigation, Univ. Stuttgart*
[k] *Observatoire de la Cote d'Azur, CERGA, Grasse*
[i] *European Space Research & Technology Centre, ESTEC, Noordwijk*

Abstract

The results of the first ESA industrial study of T&F applications for the Space Station made by a team composed by an aerospace european industry and several institutes are reviewed. The study showed about 32 experiments in disciplines ranging from Metrology and Fundamental Physics to Earth Observation, Space Science and Technology.

Keywords: Time & Frequency, Space Station

1. Introduction

This is a very brief summary of an ESA study [1-2], under the prime contractorship of DASA, involving several institutes. The activities in Time & Frequency (henceforth T&F) with the Space Station (SS) initiated in 1990 by the establishment of a scientific team [3]. The budget spent by ESA SS Utilization Department on this discipline has been small with respect to other activities, amounting to a little more than 100 kecu; nevertheless the CMGST [3] and this study have defined promising experiments of high standard and in relative large number, opening opportunities to experimenters, otherwise not included in the original plans of SS utilization.
Time and frequency standards can be used for metrology and fundamental physics but their utilization falls also onto other disciplines, often in connection with positioning measurements with GPS, GNSS, GLONASS, PRARE, DORIS, Laser and Microwave ranging systems. The study stimulated the conceptual proposition of new experiments as ACES (Atomic Clock Ensemble in Space) and collected some of the European existing effort in this field for a total of about 32 experiments.

2. Metrology and General Physics

The impact on fundamental physics tests is covered elsewhere in these proceedings as the T^2L^2, ACES, PHARAO experiments, also subject of this study. For the cooled atom clock, the main objective of the experiment on SS is the demonstration of its operations in microgravity conditions and the determination of performances (stability better than $10^{-13}\ \tau^{-1/2}$ and for one day of operation better than 3×10^{-16}, corresponding to an absolute time fluctuation of 30 *ps*).
The main concern derives from the large fluctutations on SS from the microgravity level ($10^{-6}g$) due to the crew and the machinery on board.

3. Earth Observation, Space Science, Technology

For Earth Observation, the measurement of the Earth's gravitational field, global and local point positioning, Earth rotation determination, oceanography, the electromagnetic behaviour of ionosphere and troposphere, have been considered. SS appears most suitable for terrain model and crustal motion determination by Interferometric Synthetic Aperture Radar (INSAR), because of its large construction enabling the mounting of two antennas with a long interferometric baseline. Two SAR images are taken simultaneously from neighbouring positions and pixel differences are analyzed to extract height information, eliminating the limitations of repeat-pass interferometry. Precise timing, sub-microsecond level i.e. a stabilized quartz, is required for the returned signal to be correlated for

*This study has been developed at the European Space Research & Technology Centre, ESTEC, Noordwijk, the Netherlands. Financial support from ESTEC for participation to this conference is acknowledged.

†Now at Hawaii Observ.

the phasing of the travel time of the slant ranges.
For Space Sciences (namely VLBI and relation among reference frames), SS did not appear to offer specific advantages within the study.
Technological applications are the monitoring of SATNAV systems, orbit and attitude determination of SS, subsatellite tracking, microwave sounding of the atmosphere and the use of SS as measurement platform for deep space missions.

4. Accommodation requirements

Generally the accomodation of the instrumentation itself does not pose evident show-stoppers. But the system features of SS limit the scientific utilization.
Viewing is limited by solar panels, thereby determining reduction of Earth coverage or towards space (GPS, radioastronomy).
The SS structural motion will aggravate those measurements depending on differential ranging on different points on SS.
An other issue concerns the excessive accelerations on the SS which are far greater than the microgravity levels. Concerning PHARAO, residual acceleration can be divided in two components, one parallel to the atomic drift motion between the source and the microwave cavity and the other orthogonal to the beam tube. The latter amounts only to a loss of atoms in the detection region, i.e. it amounts to an amplitude fluctuation of the signal. Accelerations greater than $2 \times 10^{-6}\ g$ in the frequency range of $5 \times 10^{-2} - 10\ Hz$ may begin to degrade the clock performances by changing the number of transmitted atoms by more than 20%. The former affects the transit time of the atoms in the microwave cavity and therefore enters the linewidth of the cesium resonance: it is equivalent to frequency noise of the microwave source. For typical operating conditions (velocity 5 cm/s, transit time of 5 seconds in the cavity), the residual parallel acceleration must be less than 2×10^{-5} in the range $5 \times 10^{-2} - 10\ Hz$.

5. Conclusions

The ACES experiment was born within this study and it is emerging as one of the most important experiments under consideration for SS allowing the characterization and the comparison of at least three different clocks, H-Maser, cooled atom clock and trapped ion clock. ACES provides, through radio and optic links, of an unique flying time scale based on an ensemble of very stable clocks, on a worldwide basis.
Further, PHARAO has emerged as a suitable experiment to fly on SS. Its considerable scientific value and its close relation to the microgravity environment pose it as an outstanding experiment.
This study was not concerned on commercial applications of T&F, although there is a potential.
The activities on T&F have provided finally a channel for the community to be represented and to propose several applications, especially in the field of Earth Observation and Technology. The reader is referred to the original study for detailed descriptions of the applications.

6. Acknowledgements

Thanks are due to J. Baune, U. Pape, Dr K.-H. Weber of DASA Bremen, Dr.s W. Knabe consultant, S. Feltham and Ms V. Kayser (ESTEC). Communications with Dr G. Busca (Obs. Neuchâtel), Prof.s A. De Marchi and S. Leschiutta (Politecnico di Torino) are duly acknowledged. Financial support was from the General Studies budget allocated to SS Utilization.

7. References

[1] ESA-DASA, 1995. *Time & Frequency Science Utilization Study and Space Station*, Study manager E. Graf, Study scientist A. Spallicci. Prime: DASA. Institutes: Deutsche Geodätisches Forschungsinstitut, München; DLR, Deutsche Forschungsanstalt für Luft und Raumfahrt Oberpfaffenhofen; Techn. Univ. München; Ecole Normale Superieure, Paris; Techn. Univ. of Dresden & Lohrmann Obs.; Inst. of Navigation, Univ. Stuttgart; Observatoire de la Cote d'Azur, CERGA, Grasse. ESA Contract 11287/94/NL/VK, Executive summary and Final report.

[2] In submission to Aerotecnica Missili e Spazio.

[3] Spallicci A., 1993. *Advanced Series of Astrophysics and Cosmology*, 7, 505, Proc. I.W. Fairbank mtg on Relativistic Gravitational Experiments in Space, 10-14 September 1990 Roma, M. Demianski and C.F.W. Everitt Eds., World Scientific; see the other paper in these proceedings for other references.

PERFORMANCES OF NEW HYPERSTABLE BVA OSCILLATORS

M. MOUREY, R.J. BESSON

Laboratoire de Chronométrie, Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques (ENS2M)
26 Chemin de l'Epitaphe - 25030 BESANÇON CEDEX - FRANCE

Abstract

In this paper we essentially show first experimental stability measurements of new hyperstable oscillators whose principle desgin and construction will not be discussed here. The preliminary data exhibit an improvement of roughly **an order of magnitude in $\sigma_y(\tau)$.**

I - INTRODUCTION

We first show that modern ultrastable oscillators (U.S.O.) performances are essentially limited by the noise in electronic circuits. This is especially the case with BVA resonators whose noise performances are outstanding. In fact, usually when a BVA resonator is used we could almost never improve the $\sigma_y(\tau)$ performances by changing the resonator.

II - WHAT CAN A USER ASK FROM A VERY GOOD U.S.O. ?

Giving a general answer to this question is difficult. We will only give an answer in the case of BVA U.S.O. This answer is summarized by following figures :

At 5 MHz

Stabilities	: between 7.10^{-14} and 2.10^{-13} ($10\ s \leq \tau \leq 1000\ s$)
Spectral purity	: -130 dBc/Hz at 1 Hz, -160 dBc/Hz at 100 kHz,
Aging	: between 2.10^{-12}/day and 3.10^{-11}/day
Stability versus temperature	: $\leq 5.10^{-13}$/°C from -20°C to +60°C
g sensitivity	: 10^{-10}/g (typical)

At 10 MHz

Stabilities	: between 1.10^{-13} and 2.10^{-13} ($10\ s \leq \tau \leq 100\ s$)
Spectral purity	: -123 dBc/Hz at 1 Hz, -158 dBc/Hz at 100 kHz,
Aging	: between 5.10-12/day and 3.10^{-11}/day
Stability versus temperature	: $\leq 5.10^{-12}$/°C from -20°C to +60°C
g sensitivity	: 10^{-10}/g (typical)

In the case of 10 MHz units we give here results measured by us and some other laboratories (including LPTF in Paris) on two units built in LCEP.

Regular extreme USO (10 MHz)

USO#	σ_y(1s)	σ_y(2s)	4s	5s	8s	10s	16s	32s	64s	unit 10^{-13}
3013	1.52	1.42	1.23	1.134	1.135	1.136	1.33	1.74	2.49	10-13
3003	1.58	1.45	1.28	1.180	1,181	1.182	1.38	1.81	2.59	10-13

II - HYPERSTABLE OSCILLATORS

A new type of oscillator has been developped for use in new generation of space clocks.
The $\sigma_y(\tau)$ characteristics has been measured for already 6 months including against hydrogen maser. At that point two sample units have been built. Oscillator's frequency is 10 MHz. Preliminary results on first prototype are as follows :

τ	1s	2s	4s	8s	16s	32s
$\sigma_y(\tau)$	1.5×10^{-13}	8×10^{-14}	5×10^{-14}	4×10^{-14}	4×10^{-14}	6×10^{-14}

The second prototype has yielded results down to 6×10^{-14} ; unfortunately, the quartz resonator was almost destroyed by accident.

IV - CONCLUSION

A new type of oscillator seems to allow an improvement of roughly one order of magnitude in $\sigma_y(\tau)$.

V - ACKNOWLEDGMENT

Authors wish to thank French authorities DRET (in Délégation Générale de l'Armement) and CNES (Centre National d'Etudes Spatiales) for support in this work. We specially thank MM. Duchaussoy, Petit, Klein, Brunet, Guillemot and Gonzales for support and interesting discussions.

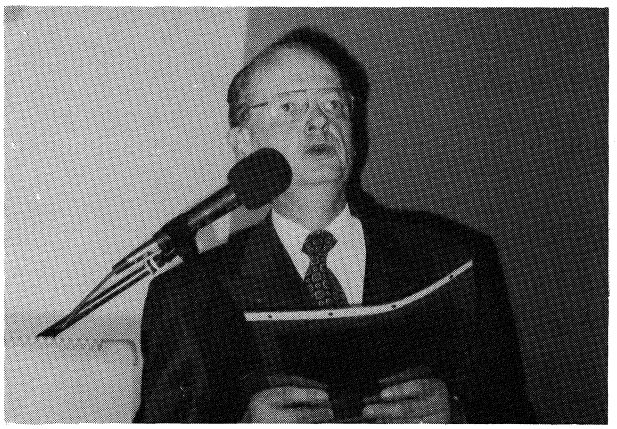
LINEAR TRAPPED ION CLOCK FOR A SMALL EXPLORER

Lute Maleki
NASA Jet Propulsion Laboratory
Passadena, CA – USA

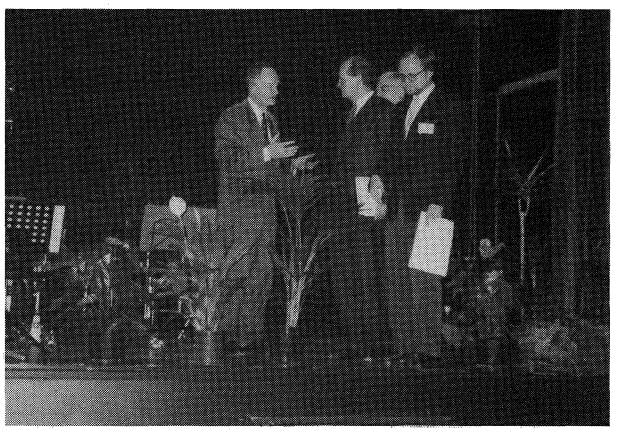
A Small Explorer mission (SMEX) is currently under study for the use of an atomic clock to test various aspects of General Relativity. Because of the constraints of time and schedule associated with the SMEX class of missions, only a Linear Ion Trap Standard (LITS) can currently meet these constraints. In this talk a short description of the proposal and the clock will be given.

Award of the «European time and frequency prize »
made by the Société Française des Microtechniques et de Chronométrie (SFMC) to

the teams of the Physikalisch-Technische Bundesanstalt (PTB)
Bundesallee 100, D-38116 Braunschweig



K. Dorenwendt



From the left: R. Besson, K. Dorenwendt, C. Audoin, A. Bauch

As the President of « Société Française des Microtechniques et de Chronométrie » so said « **Société Française de Chronométrie** », it is a great pleasure and also a great honour for me to publically introduce the winners of 1997 "**Prix Européen Temps-Fréquence**". This year the choice of the committee was not an easy one since numerous extremely good candidates have been nominated by the time and Frequency community. I would like to particularly thank in the name of our committee all those whose proposed candidates. The number of candidates this year also means that the European Award is gaining more and more recognition in our scientific community.

It is my pleasure to announce that the committee in his session of February 25th 1997 has attributed the "Prix Européen Temps-Fréquence" to the successive teams of the **Physikalisch-Technische Bundesanstalt** so said **P.T.B**. who under the main leadership of ***Pr. Gerhardt Becker, Dr. Klaus Dorenwendt, Dr. Hermann de Boer, Dr. Andreas Bauch***, have designed and operated primary cesium beam frequency standards that have shown continuously, since almost 30 years, a very high level of accuracy. They have very positively contributed to the quality of the realization of the unit of time and to the elaboration of the International Atomic Time.

These teams have proved experimentally the advantage of selecting slow cesium atoms in a thermal beam. Their results have motivated a lot of further important studies in the whole community.

It is a great pleasure for me to now call **Dr. Claude Audoin** who will help me to deliver the Award and the check to the winners and to introduce the short talk that will be given by **Dr. Klaus Dorenwendt** on a short historical review of the work at P.T.B.

Prof. Raymond J. Besson
President of SMFC

SOCIETE FRANÇAISE
DES MICROTECHNIQUES ET DE CHRONOMETRIE

Le Jury, en sa séance du 25 février 1997 a décerné
LE PRIX EUROPÉEN "TEMPS-FRÉQUENCE"

aux

Equipes du Physikalisch-Technische Bundesanstalt
(Directeurs successifs : Prof. Gerhardt BECKER, Dr Klaus DORENWENDT, Dr Hermann DE BOER, Dr Andreas BAUCH)

for outstanding contributions in designing and operating primary cesium beam frequency standards that have shown continuously, since almost 30 years, a very high level of accuracy. They have very positively contributed to the quality of the realization of the unit of time and to the elaboration of the International Atomic Time.

La remise du prix a eu lieu au **Onzième** Forum Européen "Temps-Fréquence"
à Neuchâtel le 5 mars 1997

Le Président — RJ Besson

Le Vice-Président — A. RÉMOND

MEETING REPORT

Event	11th European Frequency and Time Forum
Date	4-6 March 1997
Location	University of Neuchâtel
Participants	317
Countries	23 (for details, see Table 1)
Papers	(see Table 2)
- **Invited**	2 (opening session)
- **Oral papers**	68
- **Posters**	78
Countries	23
Sessions	17 sessions (Opening + 7 sessions led in parallel + 2 posters sessions)
Exhibitors	14 booths

a ***Tutorial session*** was held the day before the beginning of the Forum i.e March 3

Participants	49
Countries	9
Speakers	6

For the first time a ***Workshop*** was held the day after the end of the Forum i.e March 7

Participants	85
Countries	14
Papers	11

Support *:*

The organizers wish to thank :
The Republic and Canton of Neuchâtel ; The City of Neuchâtel ; The Swiss Academy of Engineering Sciences (SATW / ASST), Zürich ;
The Swiss National Science Foundation, Bern ; Le Bureau National de Métrologie, Paris - France
for their financial support.

The organizers and participants express their gratitude for the use of premises, for gifts and for any assistance provided by
The University of Neuchâtel ; L'imprimerie Krattiger, Corcelles, Neuchâtel ; PTT, Direction d'arrondissement des Télécommunications, Neuchâtel ; The Tourist Office, Neuchâtel ; The Observatory of Neuchâtel

Participants by country **TABLE 1**

Country	No of participants
Austria	5
Belgium	1
Canada	4
Czech Republic	1
Finland	7
France	104
Germany	37
Greece	1
Italy	18
Japan	8
Norway	1
Poland	7
Portugal	1
Romania	2
Russia	7
Spain	1
Sweden	1
Switzerland	43
The Netherlands	3
United Kingdom	21
United States	41
Venezuela	1
Yugoslavia	2
Total : 23 countries	**317 participants**

EFTF 1997 – contributed papers and posters **TABLE 2**

Sessions	No of participants
Opening session	2
Cesium standards	5
Optical Frequency Standards	5
Cold Cesium Atoms	4
Time and Frequency Comparisons	5
Rubidium Stadards and Flywheels	5
Time, Space, and Relativity	5
GPS Applications	5
Resonators I	5
Telecommunications	5
SAW Filters	5
Resonators II	5
Materials	5
Measurement Techniques	5
Oscillators	4
Posters Piezoelectricity	29
Posters T/F Techniques and Atomic Standards	49
Total : 14 oral sessions	**148 papers**
2 posters sessions	

1. PROGRAMME

Two fields that are becoming increasingly important for time and frequency metrology were highlighted by the invited papers of the opening session, namely space applications and relativity. **Space applications** of time and frequency are more and more perceptible to the general public, especially with the continuing development of civil uses of the Global Positioning System (GPS), to which the tutorial of March 4 and a session of the conference were entirely devoted. The best clocks of the near future will need the microgravity environment to deliver their expected performance. The possibility to launch several state-of-the art atomic clocks and time transfer equipment onboard the International Space Station was the central topic of the Workshop held on March 7 and organized jointly with the Observatoire de Paris.

With clock stabilities approaching 10^{-16} , the consequences of **Relativity** effects become of outmost importance in frequency and time comparisons and will affect many practical applications. Relativity and metrology were extensively discussed from a theoretical standpoint at the opening session and in an experimental context in one session of the conference. Progress in **atomic clocks** (cesium, rubidium, optical frequency standards and laser cooling of atoms), which was the subject of four sessions, is clearly seen to be stimulated by applications both in industry and in basic research. The field of piezoelectric devices was covered in five sessions devoted to **resonators, oscillators, surface acoustic wave filters**, and new piezoelectric **materials**.

In addition to the regular sessions of oral presentations, two poster sessions were held. The possibility of displaying the posters during the whole conference is a strong asset for this mode of presentation which is gaining ever more acceptance as it facilitates contacts and interaction between participants. In the same spirit, a "blank" session was reserved for informal discussions and was found very useful by members of various working groups to hold "special interest" meetings.

As already mentioned a one-day tutorial was held on the Global Positioning System on March 3. Special proceedings of this tutorial were edited separately. The workshop on "A new generation of space clocks" served as the first meeting of researchers planning to put the best clocks in the best conditions and, hopefully demonstrate their potential for science and technology.

2. World-wide participation

317 participants from 23 countries having a wide variety of backgrounds including national observatories, research institutes, universities and industries attended the Forum (Table 1). This diversity was reflected in the wide spectrum of topics addressed, ranging from fundamental physics and advanced research to materials, applications and measurement techniques.

3. Facilities

The Aula of the University of Neuchâtel is a perfect location for the symposium. The services, microphones, slide projection, etc. were excellent. The EFTF Conferences occupy two auditoriums as they are divided into 2 parallel sessions. The premises of the University provide enough room for the organization of the EFTF Exhibition and for the posters exhibition, although they were more numerous than other years (78 posters).

4. Exhibition

No less than 14 exhibitors from 7 countries took the opportunity to meet potential customers, reminding by their presence that their success is one important justification and motivation for the scientific results presented at the conference.

5. Conclusions

With its record number of contributions and very high attendance, the European Frequency and Time Forum has proven its attractiveness to scientists in Europe and worldwide. It is to be hoped that the next Forum, which will be held in Warsaw, Poland, will help strengthen the ties with the very active time and frequency community in eastern Europe and in the CEI.

Pierre Thomann
Scientific Chairman

List of participants EFTF - March 4-6, 1997

NAME	AFFILIATION	ADDRESS		COUNTRY
Abbott Benjamin P.	Sawtek Inc.	1818 S. HWY. 441	32703 Apopka, FL	USA
Abe Yutaka	Hokkaido University	Kita 13, Nishi 8	060 Sapporo	JAPAN
Aerni Claude	Oscilloquartz S.A.	Rue des Brévards 16 Case Postale	2002 Neuchâtel	SWITZERLAND
Alenkov Vladimir	Fomos	Buzheninova st. 16	105023 Moscow	RUSSIA
Amy-Klein Anne	Laboratoire de Physique des Lasers	Avenue J.-B. Clément	93430 Villetaneuse	FRANCE
Arnaud Roger	SICN	4, rue du Radar	74008 Annecy	FRANCE
Aubry Jean - Pierre	Oscilloquartz SA	Rue des Brévards 16	2002 Neuchâtel 2	SWITZERLAND
Audoin Claude	CNRS / LHA	Bâtiment 221, Universtié Paris -Sud	91405 Orsay	FRANCE
Azoubib Jacques	BIPM	Pavillon de Breteuil	92310 Sèvres	FRANCE
Bahoura Messaoud	LPTF / Observatoire de Paris	61, Avenue de l'Observatoire	75014 Paris	FRANCE
Baldy Michel	Tekelec TEMEX	29, Av. de la Baltique	91953 Les Ulis	FRANCE
Ballandras Sylvain	Université de Franche-Comté	32, av. de l'Observatoire	25044 Besançon Cedex	FRANCE
Barillet Roland	CNRS / LHA	Bâtiment 221, Université de Paris Sud	91405 Orsay	FRANCE
Bauch Andreas	Physikalisch-Technische Bundesanstalt	Bundesallee 100	38116 Braunschweig	GERMANY
Baumont Françoise	Observatoire de la Côte d'Azur	Av. Nicolas Copernic	06130 Grasse	FRANCE
Bava Elio	Politecnico di Milano	Piazza Leonardo da Vinci 32	20133 Milano	ITALY
Beard Ronald L.	Naval Research Laboratory	4555 Overlook Avenue, SW.	20375 5354 Washington, D.C.	USA
Bech Achim	Raltron Electronics	2315 NW 107 Avenue	33172 Miami, FL	USA
Bedrich Stefan	GeoForschungsZentrum Potsdam	Telegraphenberg A17	14473 Potsdam	GERMANY
Beeler Daniel	MSC Microsystem Consulting AG	Eichtal	8634 Hombrechtiken	SWITZERLAND
Bel Olivier	LPMO / CNRS	32 Avenue de l'Observatoire	25000 Besançon	FRANCE
Bell Donna	TFX Electronics	418 Juniper Drive	60067 Palatine, Illinois	USA
Bernier Laurent G.	Observatoire Cantonal	Rue de l'Observatoire 58	2000 Neuchâtel	SWITZERLAND
Berthoud Patrick	Observatoire cantonal de Neuchâtel	Rue de l'Observatoire	2000 Neuchâtel	SWITZERLAND
Besson Raymond J.	LCEP / ENSMM	26, ch. de l'Epitaphe	25030 Besancon Cedex	FRANCE
Bichsel René	Vibrometer SA	BP 1071	1701 Fribourg	SWITZERLAND
Bigler Emmanuel	LPMO / CNRS	32, av. de l'Observatoire	25044 Besançon Cedex	FRANCE
Bignon Olivier	C.Q.E	2, Rue Keller	10150 Pont Sainte Marie	FRANCE
Bloch Martin	Frequency Electronics Inc.	55 Charles Lindberg Blvd	11553 Mitchel Field, NY	USA
Boroditsky Roman	Valpey-Fisher Corporation	75 South Street	01748 Hopkinton, MA	USA
Boudy Caroline	C.Q.E. - Composants Quartz et Electronique	2, Rue Keller	10151 Pont-Sainte-Marie Cedex	FRANCE
Boulanger J.-Simon	National Research Council of Canada	Building, M-36 Montreal Road	K1A OR6 Ottawa, Ontario	CANADA
Bourgon Isabelle	LCEP / ENSMM	26, ch. de l'Epitaphe	25030 Besançon	FRANCE
Bourquin Roger	ENSMM / LCEP	Route de Gray - La Bouloie	25030 Besançon Cedex	FRANCE

NAME	AFFILIATION	ADDRESS		COUNTRY
Bouteville Gérard	MORS	BP 22 - La Confrérie	13610 Le Puy Sainte Reparade	FRANCE
Bouzanov Oleg A.	Fomos	Buzheninova st. 16	105023 Moscow	RUSSIA
Boy Jean-Jacques	ENSMM / LCEP	Route de Gray - La Bouloie	25030 Besançon Cedex	FRANCE
Brendel Rémi	LPMO / CNRS	32, av. de l'Observatoire	25044 Besançon Cedex	FRANCE
Briot Jean-Bernard	LPMO / CNRS	32, av. de l'Observatoire	25044 Besançon Cedex	FRANCE
Brunet Michel	CNES	18, av. Edouard Belin	31055 Toulouse Cedex	FRANCE
Brunet William	OSAFRANCE	50, Rue Marcel Dassault	92100 Boulogne	FRANCE
Busca Giovanni	Observatoire Cantonal	Rue de l'Observatoire 58	2000 Neuchâtel	SWITZERLAND
Candelier Vincent	CEPE	44, av. de la Glacière, B.P. 165	95105 Argenteuil Cedex	FRANCE
Capelle Bernard	LCEP / CNRS	4, Place Tussieu, Case 115	75252 Paris Cedex 05	FRANCE
Cash Peter	Frequency and Time Systems Inc.	34 Tozer Road	01915-5510 Beverly, MA	USA
Cérez Pierre	LHA / CNRS	Bâtiment 221, Université Paris Sud	91405 Orsay Cedex	FRANCE
Chambon Maguelonne	Bureau National de Métrologie	22, Rue Monge	75005 Paris	FRANCE
Chardonnet Christian	Laboratoire de Physique des Lasers	Avenue J.-B. Clément	93430 Villetaneuse	FRANCE
Chittick Robert	C-MAC QUARTZ CRYSTALS LTD	Edinburgh Way	CM 20 2DE Harlow, Essex	UNITED KINGDOM
Clairon André	L.P.T.F. Observatoire de Paris	61, av. de l'Observatoire	75014 Paris	FRANCE
Collet Patrick	Tekelec Temex CQE	Rue Keller / B.P. 16	10150 Pont-Sainte-Marie	FRANCE
Cordara Franco	IEN Galileo Ferraris	Strada delle Cacce 91	10135 Torino	ITALY
Costanzo Giovanni	Politecnico di Torino	Corso Duca Degli Abruzzi, 24	10129 Torino	ITALY
Coste Muriel	CNES	18, av. Edouard Belin	31055 Toulouse Cedex	FRANCE
Couteleau Laurent	LPMO / CNRS	32, av. de l'Observatoire	25044 Besançon Cedex	FRANCE
Cudia Luigi	TELECOM ITALIA	Via di Valcannuta 250	00166 Roma	ITALY
D'Albaret Bertrand	MATEL-FORDAHL	18 Rue Séjourné	94044 Créteil	FRANCE
Dalla Piazza Silvio	Asulab S.A.	Rue des Sors 3	2074 Marin	SWITZERLAND
Danaher James	3S Navigation	23141 Plaza Pointe Drive	92635 Laguna Hills, CA	USA
Davis John A.	National Physical Laboratory	Queens Road	TW11 OLW Teddington, Middlesex	UNITED KINGDOM
De Bal Marc	ALCATEL TELECOM	Berken Rodelei 75	2660 Hoboken	BELGIUM
De Jong Gerrit	NMi Van Swinden Laboratory	P.O. Box 654	2600 AR Delft	THE NETHERLANDS
De Marchi Andrea	Politecnico di Torino	Corso Duca Degli Abruzzi 24	10129 Torino	ITALY
De Young James	US Naval Observatory	3450 Massachusetts Ave. NW	20392-5420 Washington DC	USA
Defranould Philippe	Thomson-Microsonics	399, rte des Crêtes, B.P. 232	06904 Sophia Antipolis	FRANCE
Dekker Evert	SARONIX	P.O. Box 70	7000 AB Doetinchem	THE NETHERLANDS
Demidov Nikolai	KVARZ Institute of Electronic Measurement	176, Pr. Gagarina	603009 Nizhny Novgorod	RUSSIA
Desaintfuscien Michel	CNRS / LHA	Bâtiment 221, Université Paris Sud	91405 Orsay Cedex	FRANCE
Détaint Jacques	FRANCE TELECOM CNET	196, Avenue H. Ravera	92220 Bagneux	FRANCE
Deyzac François	ONERA	B.P. 72	92320 Châtillon	FRANCE

NAME	AFFILIATION	ADDRESS		COUNTRY
Dimarcq Noël	C.N.R.S. Université Paris Sud	Bâtiment 221	91405 Orsay Cedex	FRANCE
Dobrogowski Andrzej	Poznan University of Technology	Piotrowo 3A	60-965 Poznan	POLAND
Dolivo Michel	Vectron Labs, Inc.	166 Glover Ave.	06850 Norwalk, CT	USA
Domnin Yu S.	IMVP GP VNIIFTRI	Mendeleevo	141570 Moscow	RUSSIA
Dos Santos Serge	LPMO / CNRS	32, av. de l'Observatoire	25000 Besançon	FRANCE
Douglas R.J.	National Research Council of Canada	Montreal Road, Building M-36	K1A OR6 Ottawa, Ontario	CANADA
Drullinger Robert	NIST	325 Broadway	80303-3328 Boulder, CO	USA
Dudle Gregor	National Physical Laboratory		TW11 0LW Teddington, Middlesex	UNITED KINGDOM
Dulmet Bernard	ENSMM / LCEP	Route de Gray la Bouloie	25030 Besançon Cedex	FRANCE
Dumas Joël	Tekelec		Les Ulis	FRANCE
Ecabert Marcel	FSRM - Organizer EFTF 1997	Rue Jaquet-Droz 1 / CP20	2007 Neuchâtel	SWITZERLAND
Elder William	Vectron Laboratories, Inc	166 Glover Avenue	06850-5160 Norwalk, CT	USA
Engdahl Nils	Fordahl S.A.	Erlenstr. 31	2503 Bienne	SWITZERLAND
Eskelinen Pekka	Lappeenranta University of Technology	P.O Box 20	53851 Lappeenranta	FINLAND
Esteban Irène	Université Bordeaux 1 - Labo IXL	351, Cours de la Libération	33405 Talence	FRANCE
Feltham Stephen	ESA - European Space Agency ESTEC	Keplerlaan	2200 AG Noordwijk	THE NETHERLANDS
Ferrand Dominique	Hewlett-Packard France	Avenue du Canada	91947 Les Ulis Cedex	FRANCE
Ferre-Pikal Eva	NIST	325 Broadway	80303 Boulder, CO	USA
Feustel-Büechl Jörg	ESA - ESTEC	P.O. Box 299	2200 AG Noordwijk	THE NETHERLANDS
Forgacs Nandor		In der Lache	78022 Villingen-Scheweningen	GERMANY
Fossler Earl	TRAK Microwave Corporation	4726 Eisenhower Blvd.	33634-6391 Tampa, FL	USA
Fouilleul Bernard	Aura International SA	1, Place du Serpentaire, SILK 122	94513 Rungis	FRANCE
Fric Tomas	MOSER-BAER	19, Champ des Filles	1228 Plan-Les-Ouates	SWITZERLAND
Fridelance Patricia	Observatoire de la Côte d'Azur	Avenue Copernic	06130 Grasse	FRANCE
Fukuda Kyouya	Communications Research Laboratory	4-2-1, Nukui-Kitamachi, Koganei	184 Tokyo	JAPAN
Furlong Joan M.	National Physical Laboratory	Queens Road	TW11 OLW Teddington, Middlesex	UNITED KINGDOM
Galzerano Gianluca	Politecnico di Milano	Pizza L. Da Vinci, 32	20133 Milano	ITALY
Garrity Thomas J.	Cambio International	31 Dixon Street	3246 Laconia, NH	USA
Garvey Michael	Frequency and Time Systems Inc.	34 Tozer Road	01915-5510 Beverly, MA	USA
Gaspakian Shahin		In der Lache	78022 Villingen-Scheweningen	GERMANY
Gentsch Roland	Observatoire Cantonal de Neuchâtel	Rue de l'Observatoire 58	2000 Neuchâtel	SWITZERLAND
Giordano Vincent	LPMO / CNRS	32, Av. de l'Observatoire	25044 Besançon Cedex	FRANCE
Girardet Emmanuel	A.R. Electronique	Valparc - 14 rue de Franche Comté	25048 Besançon Cedex	FRANCE
Göhrig Detlef	Tele Quarz GmbH	Landstr.	74924 Neckarbischofsheim	GERMANY
Golding Mike	US Naval Research Laboratory	4555 Overlook Ave, Code 8321	20375 Washington DC	USA

NAME	AFFILIATION	ADDRESS		COUNTRY
Gonzales François	Centre National d' Etudes Spatiales	18, avenue Edouard Belin	31055 Toulouse	FRANCE
Gouni Philippe	Thomson-CSF	105 av. du général Eisenhower	31023 Toulouse	FRANCE
Graf Erhard		Ch. Chard 31	2022 Bevaix	SWITZERLAND
Green Clive	Quartzlock Instruments	Gothic, Plymouth Road	TQ9 5LH Totnes, Devon	UNITED KINGDOM
Groslambert Jacques	LPMO / CNRS	32, av. de l'Observatoire	25000 Besançon	FRANCE
Grüdler Pierre	Observatoire de la Côte d'Azur	Av. Copernic	06130 Grasse	FRANCE
Gruson Yannick	LPMO / CNRS	32, Av. de l'Observatoire	25044 Besancon Cedex	FRANCE
Guerin Pascal	Observatoire de Besançon	41 bis, av. Observatoire	25010 Besançon Cedex	FRANCE
Guillemot Philippe	CNES	18, av. Edouard Belin	31401 Toulouse Cedex	FRANCE
Guinot Bernard	Observatoire de Paris	2, Rue des Soupirs	77590 Chartrettes	FRANCE
Guyon-Gellin	LPMO / CNRS	32, Ave de l'Observatoire	25044 Besançon	FRANCE
Guyon-Gellin Jean-M.	DGA / DSP / STTC	26, Boulevard Victor	00460 Paris Armées	FRANCE
Hafner Erich	Xotex Corporation	881 Sycamore Ave.	07724 Tinton Falls, NJ	USA
Hahn Jörg	DLR	Müncher Strasse 20	82234 Wessling	GERMANY
Hamonic Yves	Tekelec Temex System	29, Av. de la Baltique	91953 Les Ulis	FRANCE
Hanson William	Piezo Crystal Company	100 K Street P.O. Box 619	17013 Carlisle, PA	USA
Hasegawa Atsushi	Communications Research Laboratory	4-2-1, Nukui-Kitamachi, Koganei	184 Tokyo	JAPAN
Hauden Daniel	CNRS / LPMO	32, av. de l'Observatoire	25044 Besançon Cedex	FRANCE
Hedley Rokos	C-MAC Quartz Crystals	Edinburgh Way	CM20 2DE Harlow, Essex	UNITED KINGDOM
Helmcke Jürgen	Physikalisch-Technische Bundesanstalt	Bundesallee 100	38116 Braunschweig	GERMANY
Henry Emmanuelle	LPMO / CNRS	32, rue de l'Observatoire	25044 Besançon cedex	FRANCE
Herman Gildas	Tekelec-Systèmes	29, av. de la Baltique ZA Courtaboeuf	91953 Les Ulis Cedex	FRANCE
Herzog Hans-Jürgen	KVG GmbH	Waibstadterstr. 2-4	74924 Neckarbischofsheim	GERMANY
Hilty Kurt	SWISS TELECOM PTT Generaldirektion	Ostermundigenstr. 91-93	3000 Bern 29	SWITZERLAND
Hisleur Christiane	LPMO / CNRS	32, Ave de l'Observatoire	25044 Besançon	FRANCE
Hodé Jean-Michel	Thomson Microsonics	399, rte des Crêtes	06904 Sophia Antipolis	FRANCE
Hodge Clifford C.	National Physical Laboratory	Queens Road	TW 11 OLW Teddington, Middlesex	UNITED KINGDOM
Howe David A.	NIST	325 Broadway	80303 Boulder, CO	USA
Hruska Carl K.	York University	4700 Keele Street Ross Bldg. Room N520	M3J 1P3 North York, Ontario	CANADA
Hübner Gerhard	Efratom Elektronik GmbH	Fichtenstr. 25	85649 Hofolding	GERMANY
Huntingford Dave	MOTOROLA AECS	Taylors road	SG5 4AY Stotfold, Hitchin, Herts	UNITED KINGDOM
Husted Charles	Raltron Electronics	2315 NW 107 Avenue	33172 Miami, FL	USA
Ikegami Takeshi	National Research Laboratory of Metrology	1- 4 -1 Umezono	Ibaraki 305 Tsukuba	JAPAN
Impiglia Paolo	Telecom Italia SpA	Via Della Vignaccia 45	00163 Roma	ITALY
Ishii Osamu	Toyo Communication Equipment Co., Ltd	1-1, Koyato 2-chome,	253-01 Samukawa-Machi , Kanagwa	JAPAN

NAME	AFFILIATION	ADDRESS		COUNTRY
Jaldehag Kenneth	Swedish National Testing and Research Institute	Box 857	50115 Boras	SWEDEN
Jamin-Changeart François	CNES	18, avenue Edouard Belin	31055 Toulouse	FRANCE
Jeanmaire Alain	Tekelec Neuchâtel Time SA	Av. du Mail 59	2000 Neuchâtel	SWITZERLAND
Jessa Mieczyslaw	Poznan University of Technology	ul. Piotrowo 3 a	60-965 Poznan	POLAND
Jordan Christine	The University of Manchester		SK11 9DL Macclesfield, Cheshire	UNITED KINGDOM
Jornod Alain	Observatoire Cantonal	Rue de l'Observatoire 58	2000 Neuchâtel	SWITZERLAND
Joyet Alain	Observatoire Cantonal de Neuchâtel	Rue de l'Observatoire 58	2000 Neuchâtel	SWITZERLAND
Jucker Hans H.	SF-FUEF Engineering Radar Systems		8600 Dübendorf	SWITZERLAND
Kajita Masatoshi	Communications Research Laboratory	4-2-1, Nukui-Kitamachi, Koganei	184 Tokyo	JAPAN
Kalliomäki Kalevi	University Oulu	P.O. Box 400	90571 Oulu	FINLAND
Kartaschoff Peter		La Pistoule 28	2036 Cormondrèche	SWITZERLAND
Kasznia Michal	Poznan University of Techology	Piotrowo 3 A	60 969 Poznan	POLAND
Kavolis Perkunas	Piezo Crystal Co.	100 K Street	17013 Carlisle, PA	USA
Kazuyuki Kita	Toyocom	12-32, Konan 2-chome, Minato-ku	108 Tokyo	Japan
Kirchner Dieter	Technische Universität Graz	Inffeldgasse 12	8010 Graz	AUSTRIA
Kiuchi Hitoshi	Communications Research Laboratory	4-2-1 Nukiu-Kita, Koganei	184 Tokyo	JAPAN
Klepczynski William	Innovative Solutions International	8500 Leesburg Pike, Suite 608	22182 Vienna, Va	USA
Klipstein William	NIST	Physics Bldg., Rm. B160	20899 Gaithersburg, MD	USA
Knight David J.E.		110 Strawberry Vale	TW1 4SH Twickenham	UNITED KINGDOM
Knight Phil	MOTOROLA AECS	Taylors road	SG5 4AY Stotfold, Hitchin, Herts	UNITED KINGDOM
Kohl Rudolf	Dornier Satellitensysteme	Postfach 80 11 69	81663 München	GERMANY
Kondratiev Serguei	Micronas Semiconductor S.A.	Ch. Chapons-des-Prés	2022 Bevaix	SWITZERLAND
Koskela Julius	Helsinki University of Technology	Rakentajanaukio 2c	02150 Espoo	FINLAND
Koutcherov Victor	Gosstandart	Leninsky prosp. 9	117049 Moscow	RUSSIA
Kramer Gunter	Physikalisch-technische Bundesanstalt	Bundesallee 100	38116 Braunschweig	GERMANY
Krasnikova Svetlana	Sawtek Inc.	1818 S. HWY. 441	32703 Apopka, FL	USA
Krempl Peter W.	AVL List GmbH	Kleiststrasse 48	8020 Graz	AUSTRIA
Krispel Ferdinand	AVL List GmbH	Kleistrasse, 48	A Graz	AUSTRIA
Kurzenknabe Glenn	Piezo Crystal Co.	P.O. Box 619	17013 Carlisle, PA	USA
Lamboley	CEPE	44, av. de la Glacière, B.P. 165	95105 Argenteuil Cedex	FRANCE
Laurent Philippe	Observatoire de Paris - LPTF	61, Av. de l'observatoire	75014 Paris	FRANCE
Laverty John	National Physical Laboratory	Queens Road	TW11 0LW Teddington, Middlesex	UNITED KINGDOM
Le Corre J.L.	CEPE	44, av. de la Glacière, B.P. 165	95105 Argenteuil Cedex	FRANCE
Lea Stephen Nicholas	National Physical Laboratory		TW11 OLW Teddington, Middlesex	UNITED KINGDOM
Lehner Christoph	DLR	Reichenbergstr. 8	82362 Weilheim	GERMANY

NAME	AFFILIATION	ADDRESS		COUNTRY
Leibfried Oliver	Tele Quarz GmbH	Landstrasse 2	74924 Neckarbischofsheim	GERMANY
Lemonde Pierre	BNM - LPTF	61 Avenue de l'Observatoire	75014 Paris	FRANCE
Leschiutta Sigfrido	Polytecnico di Torino	Corso Duca Degli Abbruzzi, 24	10129 Torino	ITALY
Levi Filippo	IEN Galileo Ferraris	str. Delle Calle 91	10135 Torino	ITALY
Lewandowski Wlodzimierz	BIPM	Pavillon de Breteuil	92312 Sèvres Cedex	FRANCE
Liacopoulos Petros	NETCOM SA	Metamorfoseos 4	153-44 Glyka Nera	GREECE
Lind Kare	Telenor Research and Development	P.O. Box 83	2007 Kjeller	NORWAY
Lipe Thanas	Tag Heuer SA	Champs-Montants 14	2074 Marin	SWITZERLAND
Lisowiec Aleksander	Tele- and Radio Research Institute	Ratuszowa 11	03-450 Warsaw	POLAND
Longet Christian		Rue Dornier 14 A	25000 Besançon Cedex	FRANCE
Luginbuhl Philippe	IMT - Université de Neuchâtel	Rue Jaquet-Droz 1	2007 Neuchâtel	SWITZERLAND
Makdissi Alaa	LPTF - Observatoire de Paris	61, ave de l'Observatoire	75014 Paris	FRANCE
Malocha Don	University of Central Florida		32816-2450 Orlando, FL	USA
Mandache Cipriana	National Inst. for Lasers, Plasma and Radiation Physics	P.0 Box MG-36	76900 Bucharest, Magurele	ROMANIA
Manetsberger Rudolf	DLR	Laberstr. 2	82362 Weilheim	GERMANY
Mannermaa Jari	University of Oulu	P.O. Box 400	90571 Oulu	FINLAND
Maroncilli Pietro	Nuova Mistral	CP n. 93	04013 Latina Scalo	ITALY
Martinovic Branislav	Institue Mihajlo Pupin	Volgina 15	11000 Beograd	YUGOSLAVIA
Mateescu Irina	Inst. of Physics and Technology of Materials	P.O. Box MG 7	76900 Bucarest Magurele	ROMANIA
Matola Marko	Kotka Institute of Technology	Paaskysentie 1	48220 Kotka	FINLAND
Matsakis Demetrios	U.S. Naval Observatory	3450 Masssachusetts Avenue, NW	20392 Washington DC	USA
Maze Thierry	DCN BREST / SDPROD / NAV		29240 Brest Naval	FRANCE
Mérigoux Henri	Université de Franche Comté	Route de Gray-La Bouloie	25030 Besançon Cedex	FRANCE
Messaoudi Tijani	LCEP / ENSMM	26, ch. de l'Epitaphe	25030 Besançon	FRANCE
Messina Filippo	Osservatorio Astronomico	Strada 54, Loc. Poggio dei Pini	09012 Capoterra	ITALY
Michel Stéphane	ENSMM / LCEP	26, ch. de l'Epitaphe	25030 Besançon Cedex	FRANCE
Mileti Gaetano	NIST	325 Broadway	80303-3328 Boulder, CO	USA
Milic Branislav	Institue Mihajlo Pupin	Volgina 15	11000 Beograd	YUGOSLAVIA
Minguy Yves	DCN BREST - SDPROD-NAV		29240 Brest Naval	FRANCE
Mitrofanov Igor	Avangard-Elionica Co. Ltd	Kondratyevsky pr. 72	195271 St. Petersburg	RUSSIA
Montandon Carine	FSRM - Organizer EFTF 1997	Case postale 20	2007 Neuchâtel	SWITZERLAND
Morison Ian	The University of Manchester	Nr. Macclesfield	SK11 9DL Macclesfield, Cheshire	UNITED KINGDOM
Morley Peter	CTS Frequency Control	400 Reimann Avenue	60548 Sandwich, IL	USA
Mourey Marc	ENSMM / LCEP	Route de Gray-La Bouloie	25030 Besançon Cedex	FRANCE
Mundschau Jonathan W.	Valpey-Fisher Corporation	75 South Street	01748 Hopkinton, MA	USA
Mureddu Leonardo	Osservatorio Astronomico	Strada 54, Loc. Poggio dei Pini	09012 Capoterra	ITALY

NAME	AFFILIATION	ADDRESS		COUNTRY
Naglowski Piotr	Institute of Electronic Materials Technology	Wolczynska 133	01-919 Warsaw	POLAND
Nau Hartmut	DLR - Deutsche Forschungsanstalt für Luft- und Raumfahrt e.V. Institut für Hochfrequenztechnik	Münchnerstr. 20	82234 Weßling	GERMANY
Neubig Bernd W.	Tele Quarz GmbH	Landstrasse 2	74924 Neckarbischofsheim	GERMANY
Ng Peggy	York University	4700 Keele Street	M3J 1P3 North York, Ontario	CANADA
Nosek Jaroslav	Technical University of Liberec	Halkova 6	46117 Liberec	CZECH REPUBLIC
Novero Carlo	IEN Galileo Ferraris	Strada delle Cacce 91	10135 Torino	ITALY
Nuotio Juha	Kotka Institute of Technology	Paaskysentie 1	48220 Kotka	FINLAND
Odiet Laurent	Fordahl S.A.	Erlenstr. 31	2503 Bienne	SWITZERLAND
Overney Frédéric	Swiss federal office of metrology	Lindenweg 50	3048 Waben	SWITZERLAND
Pabst Alexander	Semic RF Electronic GmbH	Postweg 2	82024 Taufkirchen	GERMANY
Palacio Juan	Real Observatorio de la Armada	Cecilio Pojazon S/N	11.100 San Fernando	SPAIN
Pawlitzki Alexander	TimeTech GmbH	Nobelstr. 15	70569 Stuttgart	GERMANY
Peier Ulrich R.	Peier Engineering S.A.	Endlikerstrasse 60	8400 Winterthur	SWITZERLAND
Pelletier Vincent	Tekelec TEMEX	29, Av. de la Baltique	91953 Les Ulis	FRANCE
Pentovelis Georgios	C.Q.E. - Composants Quartz et Electronique	2, rue Keller	10150 Pont-Sainte-Marie Cedex	FRANCE
Perez Jacques	SAMO		Foutenay-Le-Fleury	France
Petermann Roland	Petermann-Technik	Alte Hauptstrasse 1	82237 Wörthsee	GERMANY
Petit Gérard	BIPM	Pavillon de Breteuil	92312 Sèvres Cedex	FRANCE
Petit Pierre	C.N.R.S. Université Paris Sud	Bâtiment 221	91405 Orsay Cedex	FRANCE
Placito R.L.	Steatite Insulations Ltd.	2, The Square Broad Street	B15 1AP Birmingham	UNITED KINGDOM
Planat Michel	LPMO / CNRS	32, av. de l'Observatoire	25044 Besançon Cedex	FRANCE
Plessky Victor	Micronas Semiconductor S.A.	Ch. Chapons-des-Prés	2022 Bevaix	SWITZERLAND
Plumelle Francine	CNRS / LHA	Bâtiment 221, Université Paris-Sud	91405 Orsay Cedex	FRANCE
Poulain Pierre	CEPE	44, av. de la Glacière	95105 Argenteuil	FRANCE
Prost Léon	Office Fédéral de Métrologie	Lindenweg 50	3084 Wabern	SWITZERLAND
Raoelijaona F.	ENSMM / LCEP	26, chemin de l'Epitaphe	25030 Besançon Cedex	FRANCE
Ratier Nicolas	LPMO / CNRS	32, Av. de l'Observatoire	25044 Besançon	FRANCE
Re Vito	Frequency Electronics	55 Charles Lindberg	Mitchel Field, NY	USA
Reineck Stefan	KVG Quartz Crystal Technology GmbH	Waibstadter Str. 2-4	74924 Neckarbischofsheim	GERMANY
Ressler Hubert	Technische Universität Graz	Infeldgasse 12	8010 Graz	AUSTRIA
Riley William J.	EG & G	35, Congress Street	01970 Salem, MA	USA
Rochat Pascal	Tekelec Neuchâtel Time S.A.	Avenue du Mail 59 / CP 51	2008 Neuchâtel	SWITZERLAND
Römisch Stefania	Politecnico di Torino	Corso Duca Degli Abruzzi 24	10129 Torino	ITALY
Rose Dwane	Saunders & Associates, Inc	7440 East Karen Drive	85260 Scottsdale, Arizona	USA
Roth Hartmut	Efratom Elektronik GmbH	Fichtenstr. 25	85649 Hofolding	GERMANY
Rubiola Enrico	Politecnico di Torino	Corso Duca Degli Abruzzi 24	10129 Torino	ITALY

NAME	AFFILIATION	ADDRESS		COUNTRY
Rudolph Harald	KVG	Waibstadter Str. 2-4	74924 Neckarbischofsheim	GERMANY
Ruser Claus	Alpha Quartz	Av. Francisco de Miranda, Oeste Piso 9	1060 Caracas	VENEZUELA
Sakharov Sergey	Fomos	Buzheninova st. 16	105023 Moscow	RUSSIA
Salitz Thomas	Semic RF Electronic GmbH	Postweg 2	82024 Taufkirchen	GERMANY
Salomon Christophe	Laboratoire Kastler Brossel / ENS	24, rue L'Homond	75005 Paris	FRANCE
Salpala Anne	Kotka Institute of Technology	Paaskysentie 1	48220 Kotka	FINLAND
Samain Etienne	Observatoire de la côte d'Azur	Av. Copernic	06130 Grasse	FRANCE
Santarelli Giorgio	L.P.T.F. - Observatoire de Paris	61, av. de l'Observatoire	75014 Paris	FRANCE
Sauerland Paul	Transat Corporation	31000 Bainbridge Road	44139 Solon, Ohio	USA
Saunders Steve	Roditi Int. Corp. Ltd.	130 Regent Street	W1R 6BR London	UNITED KINGDOM
Schäfer Wolfgang	TimeTech GmbH	Nobelstr. 15	70569 Stuttgart	GERMANY
Schäfer Wolfgang	TimeTech GmbH	Nobelstr. 15	70569 Stuttgart	GERMANY
Schleinzer Gerald	AVL-List GmbH	Kleiststrasse 48	8020 Graz	AUSTRIA
Schlueter Bernard	Organizer EFTF 1997	BP 704	2001 Neuchâtel	SWITZERLAND
Schlüter Wolfgang	Institut für Angewandte Geodäsie		93444 Kötzting	GERMANY
Schmidt Reiner	Lucent Technologies Zürich		Zürich	SWITZERLAND
Schneuwly Dominik	Oscilloquartz SA	Rue des Brévards 16	2002 Neuchâtel 2	SWITZERLAND
Schweda Hartmut	Observatoire Cantonal	Rue de l'Observatoire 58	2000 Neuchâtel	SWITZERLAND
Seiler Detlef		In der Lache	78022 Villingen-Scheweningen	GERMANY
Seydel Eberhard	KVG Quartz Crystal Technology GmbH	Waibstadter Str. 2-4	74924 Neckarbischofsheim	GERMANY
Silverio Marques Fatima	IPQ - Instituto Portugues da Qualidade	Rua C à Avenida dos Três Vales	2825 Monte de Caparica	PORTUGAL
Simon Eric	LPTF	61, Avenue de l'Observatoire	75014 Paris	FRANCE
Solal Marc	Thomson Microsonics	399, rte des Crêtes	06904 Sophia Antipolis Cedex	FRANCE
Soluch Waldemar	Institute of Electronic Materials Technology	Wolczynska 133	01-919 Warsaw	POLAND
Sommer Roland	Kistler Instrumente AG	Erlachstr. 22	8408 Winterthur	SWITZERLAND
Stansfeld James	Roxbord Group PLC	124 Victoria Road, Farnborough	GU 14 7PW Hants	UNITED KINGDOM
Steele James Mc A.	National Physical Laboratory	Bldg 31, RM 25	TW11 OLW Teddington, Middlesex	UNITED KINGDOM
Steele Peter	Efratom Elektronik GmbH	Fichtenstr. 25	85649 Hofolding	GERMANY
Sthal Fabrice	ENSMM / LCEP	26, Ch. de l'Epitaphe	25030 Besançon Cedex	FRANCE
Stroppolo Luca	TELECOM ITALIA	Via di Valcannuta 250	00163 Roma	ITALY
Stührmann Holger	Tekelec Temex	Kapuzinerstrasse 9	80337 München	GERMANY
Sullivan Donald B.	NIST	325 Broadway	80303 Boulder, CO	USA
Svelto Cesare	Politecnico di Milano	P. Leonardo da Vinci 32	20130 Milano	ITALY
Szymaniec Krzysztof	LPTF / BNM - Observatoire de Paris	61, av. de l'Observatoire	75014 Paris	FRANCE
Takazawa Koji	Toyocom Europe GmbH	Bollenhöhe 5	40822 Mettmann	GERMANY
Telle Harald R.	Physikalisch-Technische Bundesanstalt	P.O. Box 3345	38023 Braunschweig	GERMANY

NAME	AFFILIATION	ADDRESS		COUNTRY
Tellier Colette R.	ENSMM / LCEP	26, Chemin de l'Epitaphe	25030 Besançon Cedex	FRANCE
Thomann Pierre	Chairman of the Scientific Committee - EFTF 1997	Rue de l'Observatoire 58	2000 Neuchâtel	SWITZERLAND
Thomas Claudine	BIPM	Pavillon de Breteuil	92312 Sèvres Cedex	FRANCE
Thorvaldsson Thor	Micronas Semiconductor S.A.	Rue Chapon-des-Prés	2022 Bevaix	SWITZERLAND
Trialoup Claude	CEPE	44, av. de la Glacière	95105 Argenteuil	FRANCE
Trümpy Kaspar	Micro Crystal	Mühlestrasse 14	2540 Grenchen	SWITZERLAND
Tuladhar Kanak K.	Oscilloquartz S.A.	Rue des Brévards 16	2002 Neuchâtel 2	SWITZERLAND
Uhrich Pierre	BNM - LPTF	61, av. de l'Observatoire	75014 Paris	FRANCE
Underhill M.J.	University of Surrey		GU2 5XH Guildford	UNITED KINGDOM
Vacheron Pascal	Micronas Semiconductor S.A.	Ch. Chapons-des-Prés	2022 Bevaix	SWITZERLAND
Valentin Constance	CNRS Laboratoire de l'horloge Atomique	Bât. 221 Université Paris-Sud	91405 Orsay	FRANCE
Vanier Jacques	IEN Galileo Ferraris	Strada Delle Cacce 91	10135 Torino	ITALY
Vedel Fernande	Université de Provence	Centre de Saint Jérôme, Case C 21	13397 Marseille Cedex 20	FRANCE
Vernotte François	Observatoire de Besançon	41 bis, av. Observatoire, B.P. 1615	25010 Besançon Cedex	FRANCE
Vessot Robert F.C.	Harvard-Smithsonian Astrophysical Observatory	60 Garden Street	02138 Cambridge, MA	USA
Wagner Karl Ch.	Siemens Matsushita Comp.	Postfach 801709	81617 München	GERMANY
Walls Fred L.	NIST	325 Broadway	80303 Boulder, CO	USA
Walter Hannes U.	ESA Microgravity Applications Promotion Office	8-10 rue Mario-Nikis	75738 Paris cedex 15	FRANCE
Walter Harald	Lucent Technologies	Thurn-und-Taxis-Strasse 10	90411 Nürnberg	GERMANY
Weidemann Werner	Efratom Time & Frequency Products	3 Parker	92618-1696 Irvine, CA	USA
Weiss Krzysztof	Tele and Radio Research Institute	Ratuszowa 11	03 450 Warsaw	POLAND
Wenzel Robert	Lucent Technologies	Thurn-und-Taxis-Strasse 10	90411 Nürnberg	GERMANY
Weyers Stefan	Physikalisch-Technische Bundesanstalt	Bundesallee 100	38116 Braunschweig	GERMANY
Whibberley P.	National Physical Laboratory		TW11 0LW Teddington, Middlesex	UNITED KINGDOM
White J.D.	US Naval Research Laboratory	4555 Overlook Ave., SW	20375-5354 Washington DC	USA
Winkler Gernot M. R.	Innovative Solutions International	8500 Leesburg Pike, Suite 608	22182 Vienna, VA	USA
Wolcoff Bernard	A. R. ELECTRONIQUE	15 rue d'Estienne d'Orves	78500 Sartrouville	FRANCE
Wolf Michael	Alcatel Telecom	Lorenzstrasse 10	70435 Stuttgart	GERMANY
Wolf Peter	BIPM	Pavillon de Breteuil	92312 Sèvres Cedex	FRANCE
Wong Sover	Oscilloquartz S.A.	Rue des Brvards 16	2002 Neuchâtel	SWITZERLAND
Wright Peter	Thomson Microsonics	399, Route des Crêtes	06904 Sophia-Antipolis	FRANCE
Zingg Walter	Microcrystal Div. of ETA S.A.	Mühlestr. 14	2540 Grenchen	SWITZERLAND
Zucco Massimo	National Physical Laboratory		TW11 0LW Teddington, Middlesex	UNITED KINGDOM

EFTF Exhibitors - March 3-7, 1997 - Neuchâtel, Switzerland

COMPANY	*ADDRESS*	*COUNTRY*
KVG GmbH	Waibstadterstr. 2-4 D-74924 Neckarbischofsheim	GERMANY
Oscilloquartz SA	Rue des Brévards 16 CH-2002 Neuchâtel 2	SWITZERLAND
Frequency Electronics Inc.	55 Charles Lindbergh Blvd. USA-11553 Mitchel Field, NY	USA
Fomos	Buzheninova st. 16 RUS-105023 Moscow	RUSSIA
Alpha Quartz	Av. Francisco de Miranda, Oeste Piso 9 VEN-1060 Caracas	VENEZUELA
Aura International SA	1, Place du Serpentaire, SILK 122 F-94513 Rungis	FRANCE
Tekelec Neuchâtel Time S.A.	Avenue du Mail 59 / CP 51 CH-2008 Neuchâtel	SWITZERLAND
Tekelec Temex System	29, Av. de la Baltique F-91953 Les Ulis	FRANCE
Tekelec Temex CQE	Rue Keller / B.P. 16 F-10150 Pont-Sainte-Marie	FRANCE
TRAK Microwave Corporation	4726 Eisenhower Blvd. USA-33634-6391 Tampa, FL	USA
Roditi Int. Corp. Ltd.	130 Regent Street GB-W1R 6BR London	UNITED KINGDOM
3S Navigation	23141 Plaza Pointe Drive USA-92635 Laguna Hills, CA	USA
MOSER-BAER	19, Champ des Filles CH-1228 Plan-Les-Ouates	SWITZERLAND
Quartzlock Instruments	Gothic, Plymouth Road GB-TQ9 5LH Totnes, Devon	UNITED KINGDOM
TFX Electronics	418 Juniper Drive USA-60067 Palatine, Illinois	USA
MOTOROLA AECS	Taylors road GB-SG5 4AY Stotfold, Hitchin, Herts	UNITED KINGDOM
TimeTech GmbH	Nobelstr. 15 D-70569 Stuttgart	GERMANY
Lucent Technologies	Thurn-und-Taxis-Strasse 10 D-90411 Nürnberg	GERMANY

INDEX OF THE AUTHORS