Forum Européen
Fréquence et temps

European Frequency
and Time Forum

Besançon, Neuchâtel

19è FORUM EUROPEEN TEMPS-FREQUENCE

19th EUROPEAN FREQUENCY AND TIME FORUM

ACTES
PROCEEDINGS

Besançon, France
March 21st-24th, 2005

Besançon
Quatre saisons Orchestrent votre temps

Organisé par / Organized by :
Société Chronométrique de France, dite Société Française des Microtechniques et de Chronométrie (S.F.M.C.)

With it's strategic location in a curve of the Doubs river (the "Boucle"), at the crossing of two major european traderoutes (southwestern Europe - Germany, northwestern Europe - Italy through Switzerland), Besançon, at all times in history, has always been associated to the entire European community. The Celtic fortified town of Vesontio became with the Roman Empire the capital city of the Provincia Maxima Sequanorum, later known as the Free County of Burgundy during the Middle-Ages (approximately today's Franche-Comté). Since early medieval times, it is the siege of a large archbishopric that contributed to the growth of the city through the centuries. It's close links with the Holy Roman Empire (it hosted an Imperial Council in 1157 and the Emperor Friedrich Barbarossa married the heiress of the County) have made it a flourishing city : the Emperor Rudolf of Habsburg eventually made it a Free City of the Empire in 1290, and was then politically separated from the County of Burgundy.

During the 16th century, after a period of relative peace and economic growth, it provided the Empire with great diplomats such as Nicolas Perrenot de Granvelle and his son Antoine, confidents and ministers of Charles V and Philipp II, kings of Spain and Holy Roman Emperors, heirs of the title of Count of Burgundy. After several military attempts and memorable sieges, France annexed the city, along with Franche-Comté, by the Treaty of Nijmegen in 1678. Since then, it is again the capital city of Franche-Comté and hosts the University, created in Dole by the duke of Burgundy Philipp the Good in 1423.

Close to the Swiss Confederation, Besançon and the entire region became a very important center for the French clockmaking industry as soon as the end of the 18th century, with the reknowned "Horloges Comtoises". The watchmaking industry became in the middle of the 20th century an important employment sector in Besançon and Franche-Comté, until the closing of LIP in the 70's, the biggest watchmaking company of Besançon. After a successfull conversion to icro-mechanics and micro-technics, it is now a leading pole for research in Time, Frequency and Optics, with many well-known laboratories and schools : ENSMM (National School for Mechanics and Micro-technics), LPMO (Laboratory of Physics and Metrology of Oscillators), LOPMD (Laboratory of Optics P.-M. Duffieux), LCEP (Laboratory of Electronic Chronometry and Piezoelectricity), the Laboratory Time-Frequency (hosted by the Astronomical Observatory of Besançon, which was created in 1882 to support the growing local watchmaking industry). The old acquaintance of Besançon with Clocks and Time was recently officialised with the opening of the Musée du Temps (Museum of Time), located in the 16th century palace of the Granvelle family.

Today, Besançon is still considered as the French capital for clocks and watches, and many companies related to the watchmaking industry are still present in Besançon and the region. Hometown of Victor Hugo and the Lumière brothers, the city is also reknowned for it's life quality, with 590 acres (2400 ha.) of parks and forests for a total superficy of 1600 acres (6500 ha.) and 125,000 inhabitants. It's historical monuments (Roman arena, Roman triumphal arch, Granvelle palace, Quais Vauban, the Citadelle, the Astronomical Clock,...), it's international events (international music festival, Foire Comtoise,...) and it's various museums (museum of Fine Arts and Archaeology, museum of Time,...) have all contributed to offer Besançon the label of "Ville d'Art & d'Histoire" (City of Art and History).

REMERCIEMENTS

La Société Chronométrique de France, dite Société Française des Microtechniques et de Chronométrie (S.F.M.C.) qui organise le Forum Européen Temps-Fréquence, souhaiterait remercier pour leur aide et leur soutien :

- La Ville de Besançon
- Le Conseil Général du Doubs
- Le Conseil Régional de Franche-Comté

ACKNOWLEDGEMENTS

The Société Chronométrique de France, said Société Française des Microtechniques et de Chronométrie (S.F.M.C.) which organizes the European Frequency and Time Forum, wishes to thank for their help and support :

- The City of Besançon
- The Conseil Général du Doubs
- The Conseil Régional de Franche-Comté

Local Organising Committee

Raymond Besson, Femto-ST/LCEP, France (Forum Chairman)
Roger Bourquin, Femto-ST/LCEP, France (Forum Technical Chairman)
Isabelle Coutrot, DRRT, France
Joël Petetin, Observatoire de Besançon, France
Sonia Clairemidi, SFMC, France
Michel Vincent, SFMC, France

Executive Committee

Claude Audoin, France
Raymond Besson, Laboratoire de Chronométrie Electronique Piézoélectricité, France
Marcel Ecabert, FSRM, Switzerland
Lute Maleki, Jet Propulsion Laboratory, USA
Fritz Riehle, Physikalisch-Technische Bundesanstalt, Germany
Bernard Schlueter, FSRM, Switzerland
Vincent Giordano, Femto-ST/LPMO, France
Pierre Thomann, Observatoire Cantonal de Neuchâtel, Switzerland
Michael Underhill, University of Surrey, UK

Scientific Committee

Roger Bourquin, Femto-ST/LCEP, France
Jean-Pierre Aubry, Oscilloquartz SA, Switzerland
Claude Audoin, France
Andreas Bauch, Physikalisch-Technische Bundesanstalt, Germany
Elio Bava, Politecnico di Milano, Italy
Laurent-Guy Bernier, METAS, Switzerland
Emmanuel Bigler, Femto-ST/LCEP, France
Jean-Simon Boulanger, National Research Council of Canada, Canada
Michel Brunet, Centre National d'Etudes Spatiales, France
Giovanni Busca, Switzerland
André Clairon, BNM - SYRTE, France
Gerrit De Jong, NMi Van Swinden Laboratory, The Netherlands
Andrea De Marchi, Politecnico di Torino, Italy
François Deyzac, ONERA, France
Noel Dimarcq, BNM - SYRTE, France
Christopher Ekstrom, US Naval Observatory, USA
Pekka Eskelinen, Helsinki University of Technology, Finland
Jeremy Everard, University of York, UK
Steve Feltham, European Space Agency ESTEC THE, The Netherlands
Michael Garvey, Symmetricom TRC, USA
Patrick Gill, National Physical Laboratory, UK
Vincent Giordano, Femto-ST/LPMO, France
Aldo Godone, IEN Galileo Ferraris, Italy

Michel Granveaud, BNM - SYRTE, France
Yuri Gulyaev, Russian Academy of Sciences, Russia
Jörg Hahn, European Space Agency ESTEC THE, The Netherlands
Daniel Hauden, LPMO - CNRS, France
Kurt Hilty, METAS - Office Fédéral de Métrologie et d'Accréditation, Switzerland
Carl Hruska, York University, Canada
Kalevi Kalliomaki, University of Oulu, Finland
Masami Kihara, NTT Network Innovation Laboratories, Japan
Dieter Kirchner, Technische Universität Graz, Austria
Peter Krempl, AVL List GmbH, Austria
Lute Maleki, Jet Propulsion Laboratory, USA
Demetrios Matsakis, US Naval Observatroy, USA
Bernd Neubig, Germany
Rodriguez Juan Palacio, Real Instituto y Observatorio de la Armada, Spain
Christophe Salomon, Ecole Nationale Supérieure, France
Giorgio Santerelli, BNM - SYRTE, France
Jörn Stenger, Physikalische-Technische Bundesanstalt, Germany
Patrizia Tavella, Istituto Elettrotecnico Nacionale Galileo Ferraris, Italy
Pierre Thomann, Observatoire Cantonal de Neuchâtel, Switzerland
Dmitry Tsarapkin, MPEI(TU), Russia
Pierre Uhrich, BNM - SYRTE, France
Adolph Ulianov, Kvarz Institute of Electronic Measurements, Russia
Michael Underhill, University of Surrey, UK
Rainer Unverdross, Unverdross Technik, Germany
Jacques Vanier, Université de Montréal, Canada
Walter Zingg, Rolex SA, Switzerland

The Young Scientist Award was awarded this year to **M. John KITCHING** (NIST, U.S.A.), for outstanding theoretical and experimental work on various sources of noise and technical limitations, showing that excellent performance can be achieved in mm-sized physic packages for atomic clocks.

The European Time and Frequency Award was this year split between three outstanding scientists :

M. Hidetoshi KATORI (University of Tokyo, Japan), for a brilliant break-through in the field of optical frequency standards ;

M. Peter KREMPL (AVL List GmbH, Austria), for his pioneering research on the physical and electro-acoustical properties of Gallium Phosphate, and the development of industrial growth processes for large crystals including the epitaxial growth of $GaPO_4$ on quartz, leading to state of the art resonators and sensors ;

M. George MANSFELD (Academy of Sciences, Russia), for fundamental contributions in understanding physical and quantum acoustics,acoustoelectronics in piezoelectric semiconductors and acoustomagnetic effects in ferrites. For applications to frequency control and high Q microwave resonators using ultralow loss materials.

"Opening ceremony, with a few words from M. Raymond Besson, Chairman of EFTF 2005, M. Jean-Louis Fousseret, Mayor of Besançon, and M. François Vernotte, Director of the Observatory of Besançon."

Play Stop

"Closing ceremony, with a few words from M. Fritz Riehle, Chairman of EFTF 2006 in Braunschweig (Germany), M. Pierre Thomann, Chairman of EFTF 2007 in Geneva (Switzerland), M. Roger Bourquin, technical chairman of EFTF 2005, and M. Raymond Besson, Chairman of EFTF 2005."

Play Stop

19th European Frequency and Time Forum, Besançon, France

Opening Session

SYNTHETIC APERTURE RADAR

European Frequency and Time Forum 2005
Besançon, France
21-24 march 2005

S.Watts
Thales UK
Aerospace Division
Manor Royal, Crawley, West Sussex

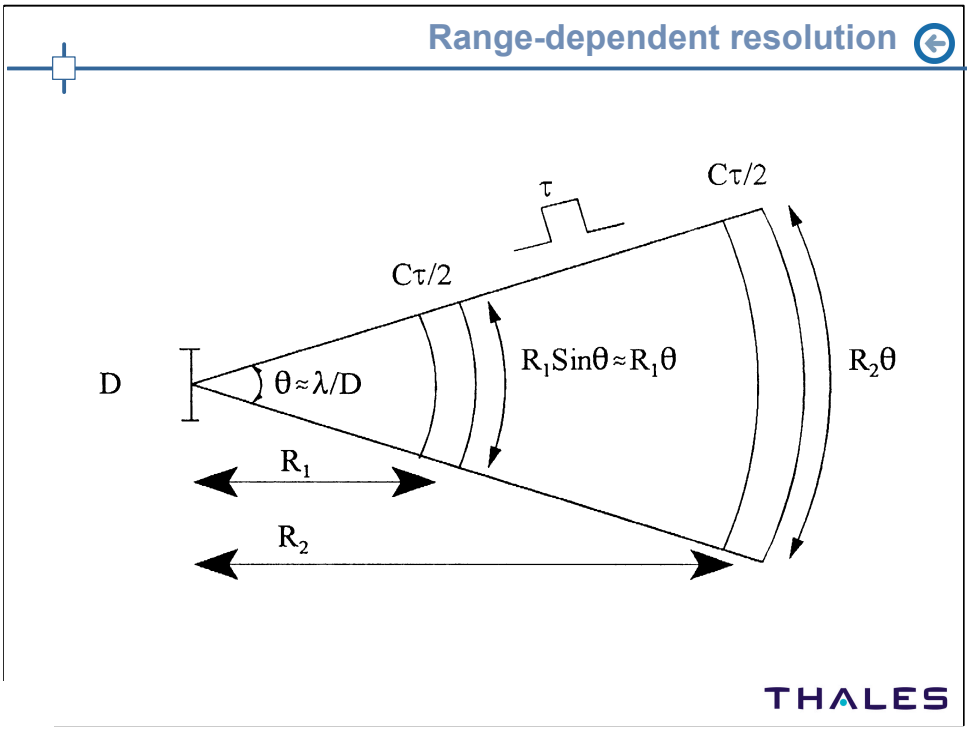
THALES

INTRODUCTION

This presentation provides a basic introduction to the principles of synthetic aperture radar. These basic principles are then used to illustrate some of the specific issues relating to frequency and time, associated with achieving good focused imagery in practical conditions.

SAR is a technique of radar imaging of the land, sea or individual targets, achieving very high spatial resolution.

The use of synthetic aperture (SAR) techniques is becoming increasingly widespread in a number of applications ranging from satellite remote sensing of land and sea, through to target imaging from airborne radars for military purposes.

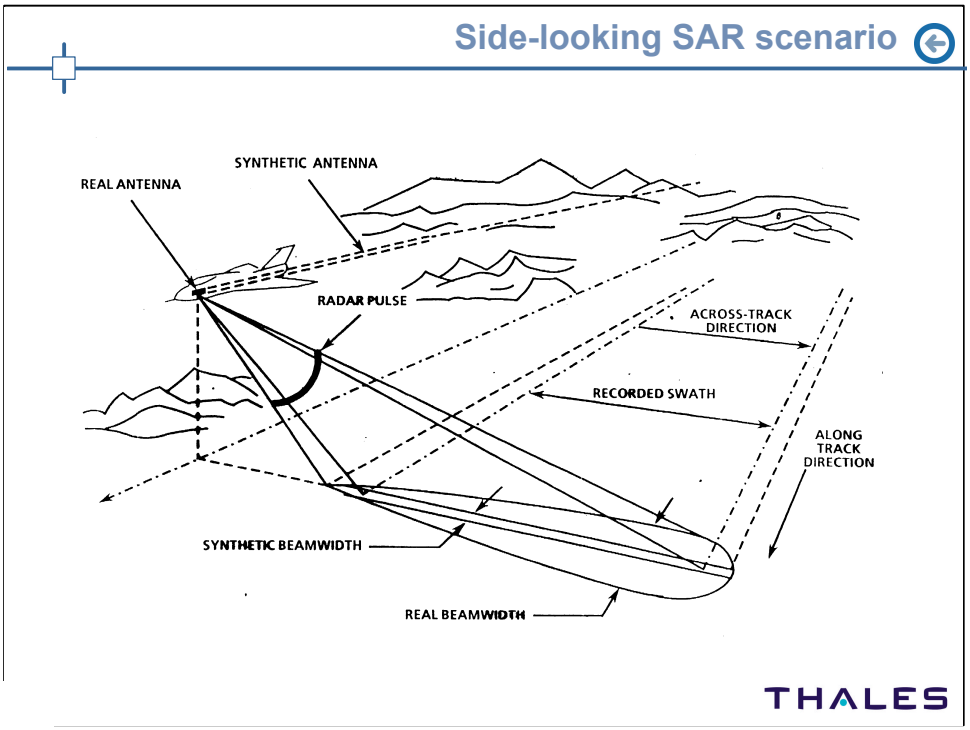


Synthetic Aperture Radar

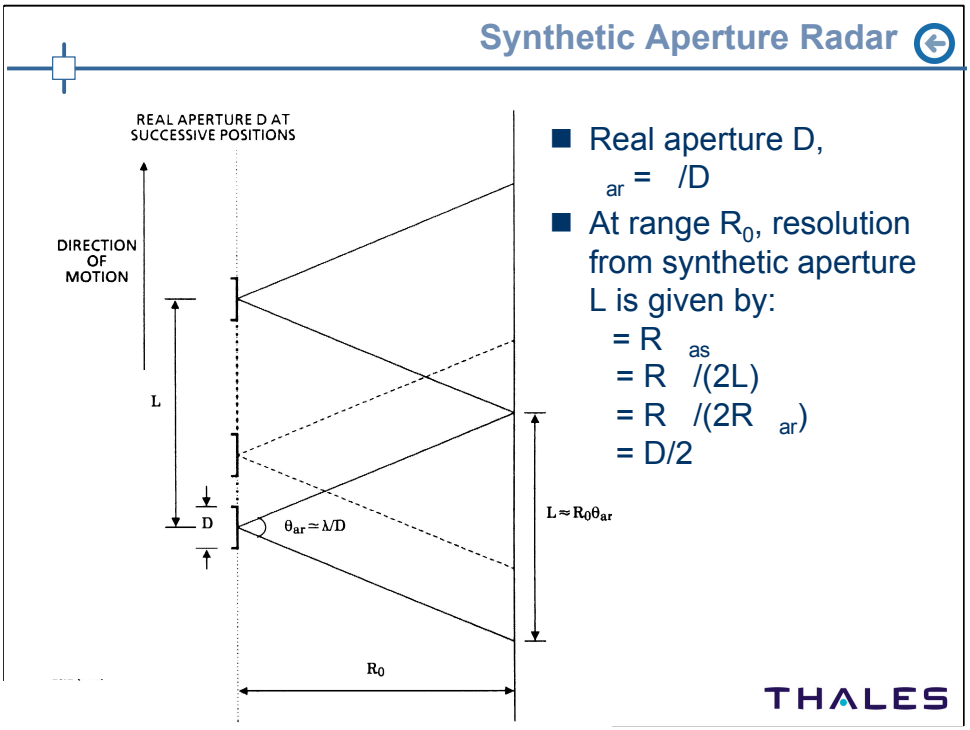
Consider a real antenna aperture D, with pulse length .

At long ranges it is clear that the spatial resolution of a radar may be very poor. A typical real aperture size of 1m and a wavelength of 3 cm would provide a resolution of no better than 3 km at a range of 100 km, whilst a realistic military requirement may be to resolve individual vehicles spaced only a few metres apart.

The range resolution achievable with a typical side-looking radar will be, say, 3m but a comparable azimuth resolution at 100 km range would require a real aperture of about 1 km which is not realistic. Indeed we shall see that range resolutions as low as 10 cm are possible and we would like cross-range resolution comparable to that.



By using a coherent radar it is possible to achieve considerably improved azimuth resolution by utilising the Doppler or phase relationship of successive radar returns. This is the basis for synthetic aperture radar. The basic principle relies on the translation or rotation of the target relative to the radar. For terrain mapping the range and hence phase of the stationary target (e.g. a point on the ground) is caused to change by the forward motion of the side-looking airborne radar, as shown above. The phase history from a sequence of radar returns is used to provide increased azimuth resolution. One interpretation of the process is the synthesis of an aperture of sufficient dimension to give the required resolution. The synthetic aperture effectively comprises a phased array.



The basic principles of SAR are shown here. A target at range R_o can be illuminated by a synthetic aperture length L, where L R_o ar. The effective beamwidth which can be achieved by a synthesised aperture of this sort is given by:

$$_{as} = k_s / (2L) \quad k_s / (2 R_0 \ _{ar})$$

where k_s is the illumination taper factor of the synthesised aperture, and the factor 2 in the expression for $_{as}$ arises due to the two-way path utilised in the beam forming. Now:

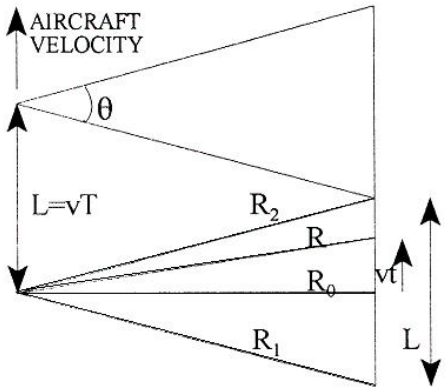
$$_{ar} = k_r / D$$

so that it can be seen that the best possible azimuth resolution which can be achieved with a real aperture size D is in fact given by:

$$R_0 \ _{as} \quad \frac{k_s}{k_r} \frac{D}{2}$$

The best possible azimuth resolution obtainable from a side-looking synthetic aperture is independent of range and has a value of approximately half the real aperture dimension.

Phase change over the synthetic aperture



$R - R_0 \quad v^2 t^2/(2 R_0)$

Two-way phase shift:

$= 2 \quad v^2 t^2 / (\quad R_0)$

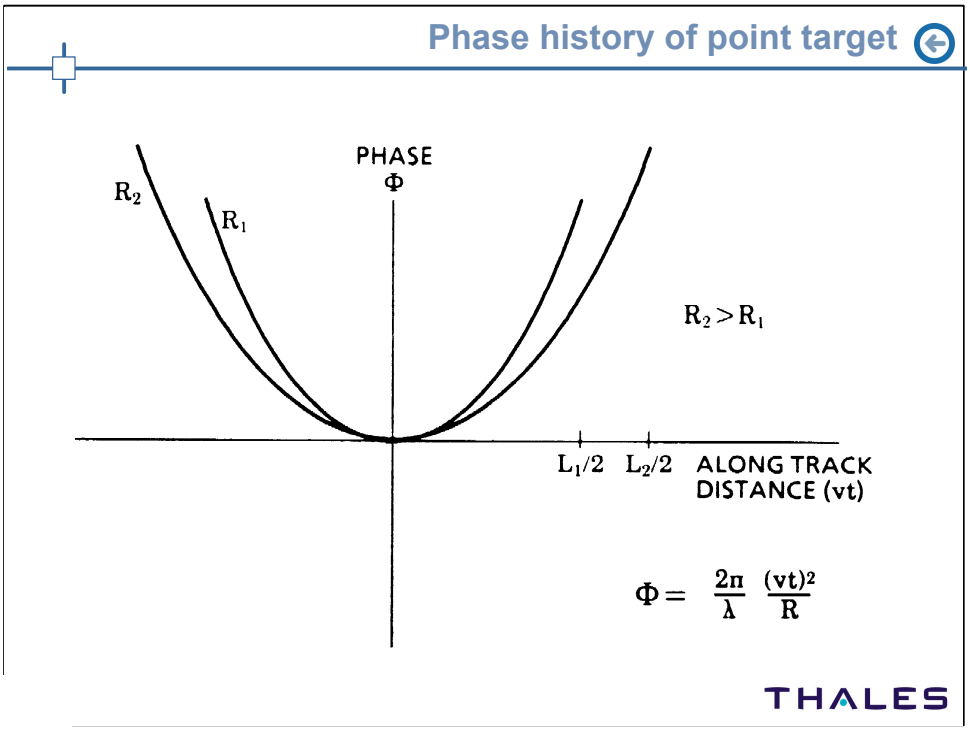
Doppler shift:

$(1/2 \ )d \ /dt = 2 v^2 t / (\quad R_0)$

THALES

An important point to note is that in general the returns from a target will be in the near field of the synthetic array, which will therefore need to be focused. In practice the radar stores the returns received at successive positions and then processes the stored returns at each range R by appropriately amplitude and phase weighting and then summing them to achieve the focused image. As with a real array, the element positions must be known to within a fraction of a wavelength so that the expected phase history of the target returns can be accurately predicted. To achieve this, the aircraft's position must be very accurately measured as a function of time and the phase and timing of the radar returns changed to correct for any deviation from a nominal expected flight trajectory. The instrumentation accuracy required for this is extremely demanding and will normally set a practical limit to the achievable resolution which is poorer than that predicted by the simple theory outlined above. Unpredictable motion of the target on the ground may also degrade the final resolution which can be achieved.

The variation of range, R, to a target within the synthetic aperture can be expressed as a change in relative phase of the radar signal. If the aircraft is travelling at a constant velocity the rate of change of phase gives a linear frequency sweep over the the synthetic aperture.



It can be seen that for the ideal aircraft motion and a side looking antenna, the range R from a point target will vary over the synthetic aperture so that:

$$R = (R_0^2 + (vt)^2)^{1/2}$$

with v : aircraft along track velocity (assumed constant)

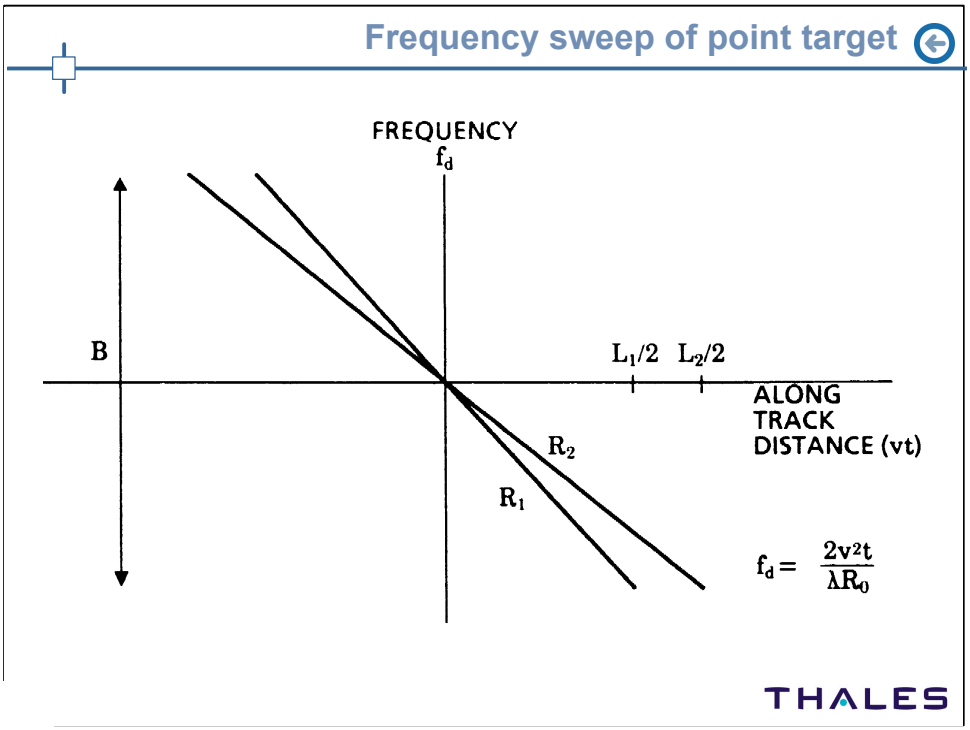
and t : time

Now R $R_0 + \frac{1}{2R_0}(vt)^2$ for R_0^3 $(vt)^4$

The phase shifts required for focused processing of a series of returns from target are therefore given by:

$$= \frac{4}{\lambda}(R - R_0) = \frac{2}{\lambda}\frac{(vt)^2}{R_0}$$

Note that the phase history is a function of range to the target, R.



A quadratic phase shift over the synthetic aperture can also be expressed as a linear rate of change of Doppler shift f_d given by:

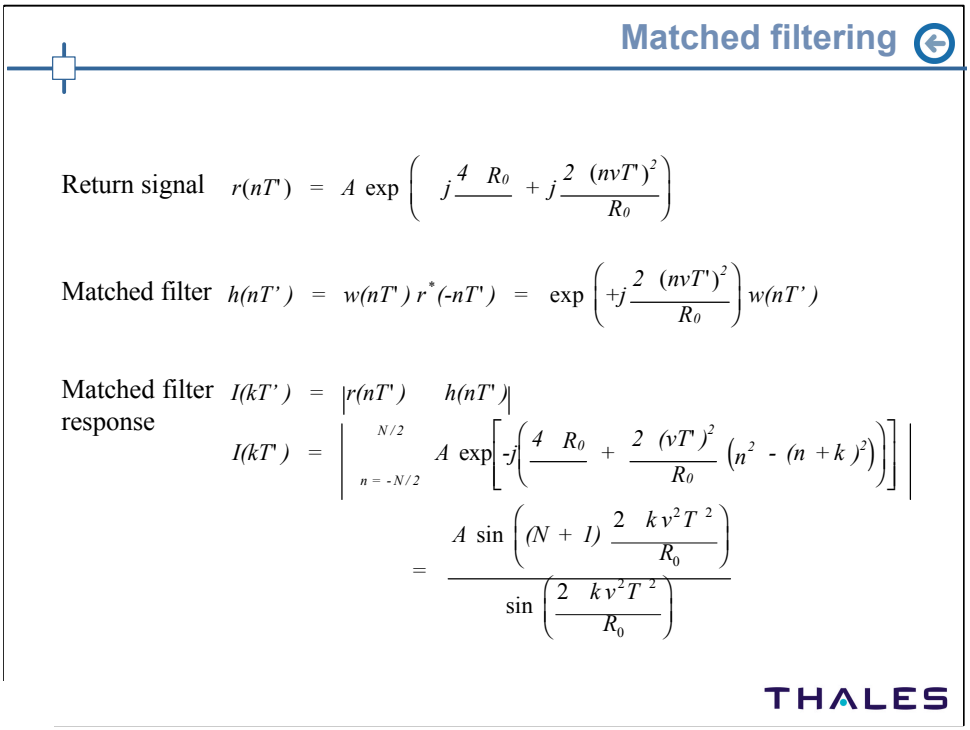
$$f_d = \frac{1}{2}\frac{d}{dt} = \frac{2\ v^2\ t}{R_0}$$

The basic expression for SAR resolution can be derived from a number of different viewpoints, in addition to that of a synthetic antenna array. For example, the linear frequency sweep shown above can be interpreted in a similar way to a linear chirp pulse compression waveform. The waveform time duration is given by T=L/v and its bandwidth (total frequency excursion) is given by

$$B = \frac{4v^2}{R}\left(\frac{T}{2}\right)$$

When this waveform is processed by a matched filter (i.e. the synthetic array elements are focused) the compressed "pulse length" will be given by 1/B, corresponding to a spatial resolution, $_{as}$, given by

$$_{as} = \frac{v}{B} = \frac{2\ R_0 v}{4v^2 T} = \frac{R_o}{2L} = \frac{D}{2}$$



For a coherent radar the phase of the return signals is measured by reference to the phase of the transmitted signal. The return signal can then be represented by

$$r(nT') = A \exp\left(j\frac{4}{} \frac{R_0}{} + j\frac{2}{R_0}(nvT')^2 \right)$$

The constant phase term $4\pi R_0/$ can be ignored and the SAR processing required to provide a focused image can be regarded as a matched filtering operation with a filter impulse response of the form:

$$h(nT') = w(nT')\, r^*(-nT') = \exp\left(+j\frac{2}{R_0}(nvT')^2\right) w(nT')$$

In practice some amplitude weighting may also be applied to the filter response to control the azimuth sidelobes and this is shown as w(nT) in equation (13).

The required SAR image is then the modulus of the matched filter output given by

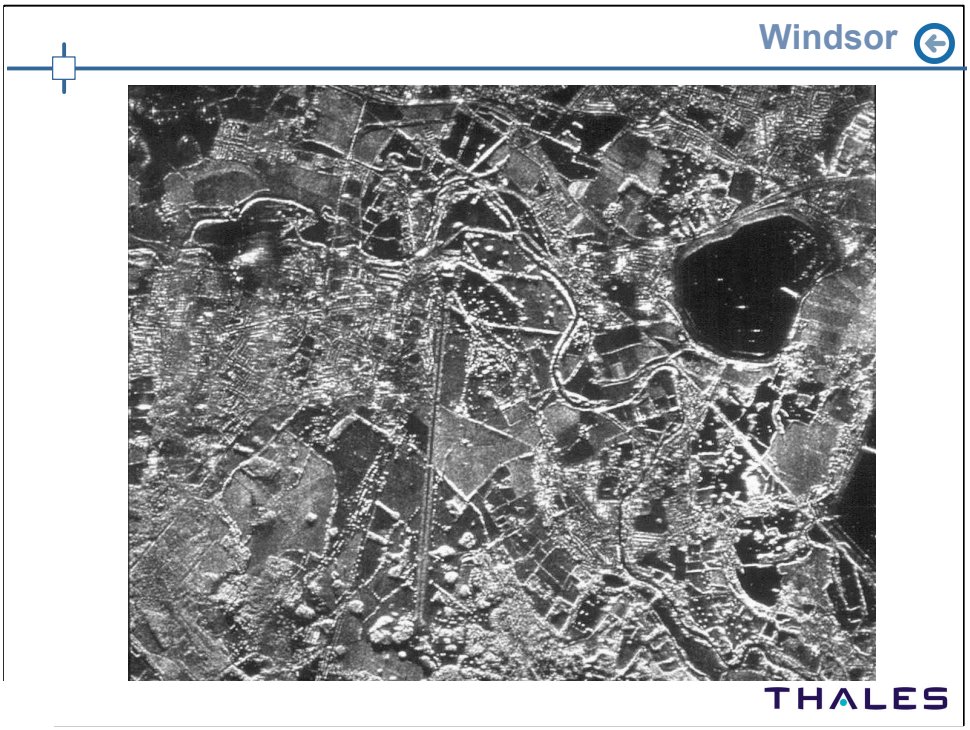
$$I(kT') = |r(nT') \quad h(nT')|$$

For the simple case of w(nT) = 1, it can be seen that

$$I(kT') = \left| \sum_{n=-N/2}^{N/2} A \exp\left[-j\left(\frac{4}{} \frac{R_0}{} + \frac{2}{R_0}(vT')^2 \left(n^2 - (n+k)^2\right)\right)\right] \right|$$

$$= \frac{A \sin\left((N+1)\frac{2}{R_0} kv^2T^2\right)}{\sin\left(\frac{2}{R_0} kv^2T^2\right)}$$

This matched filter operation can also be described as a correlation operation with the received waveform being correlated with r*(nT).



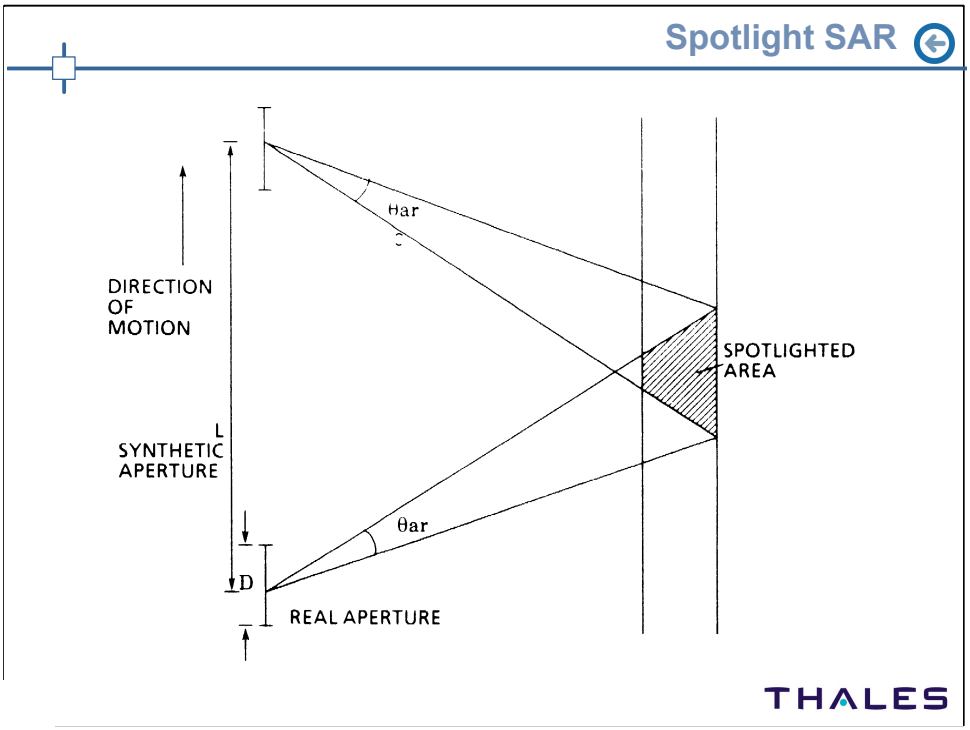
Achieving high resolution imagery is much more complicated than this simple introduction might suggest. We must cope with range walk, range drift and also remember that the quadratic phase history is only an approximation.

Nevertheless, good imagery can be achieved.

The image above shows Windsor Castle and Windsor Great Park. It was recorded by the X-band experimental SAR fitted in a Canberra.



This is an image of the countryside near Seaford on the South coast of England. Note the built-up area to the right and the woods to the left of the image. The different crops in the fields can be seen as different textures. Note also the shadows cast by tall objects – the radar was illuminating the scene from the direction of the top of the image.



Spotlight SAR mode

So far we have considered a side-looking radar, where the length of the synthetic aperture is constrained by the real aperture beamwidth. If we can steer the beam to illuminate a patch of ground over a wider angle, we can synthesise a larger synthetic aperture and achieve a resolution <D/2. We trade the ability to achieve continuous mapping to get higher resolution.

SAR example

Radar wavelength		3 cm	1.8 cm
Max range R_m		100km	30 km
Range resolution	$_r$	3m	10 cm
Aircraft velocity	v	250 m/s	30 m/s
Real antenna aperture	D	2m	0.5 m
Real antenna beamwidth	$_{ar}$	1	2
Synthetic azimuth resolution	$_{as}$	3m	10 cm
Radar PRF	$c/2R \quad f_p > 2kv_{ar}/$	1 kHz	1 kHz
Max aperture Synthesis Time	$T = R_m /(2_{as}v)$	2s	90 s
Doppler bandwidth	$B = 2v^2T/R$	83 Hz	300 Hz

THALES

These are some basic parameters for two hypothetical examples of SAR systems. The first is a long range radar operating in X-band (= 3 cm) from a fast moving platform and the second is a very high resolution short-range radar operating in J-band (= 1.8 cm) from a much slower moving platform.

The resolution desired by the J-band radar is < D/2, so spotlight mode must be used. Note the very long aperture synthesis time required to achieve this performance.

Tolerable residual motion errors

- Radial motion error r_e results in phase error

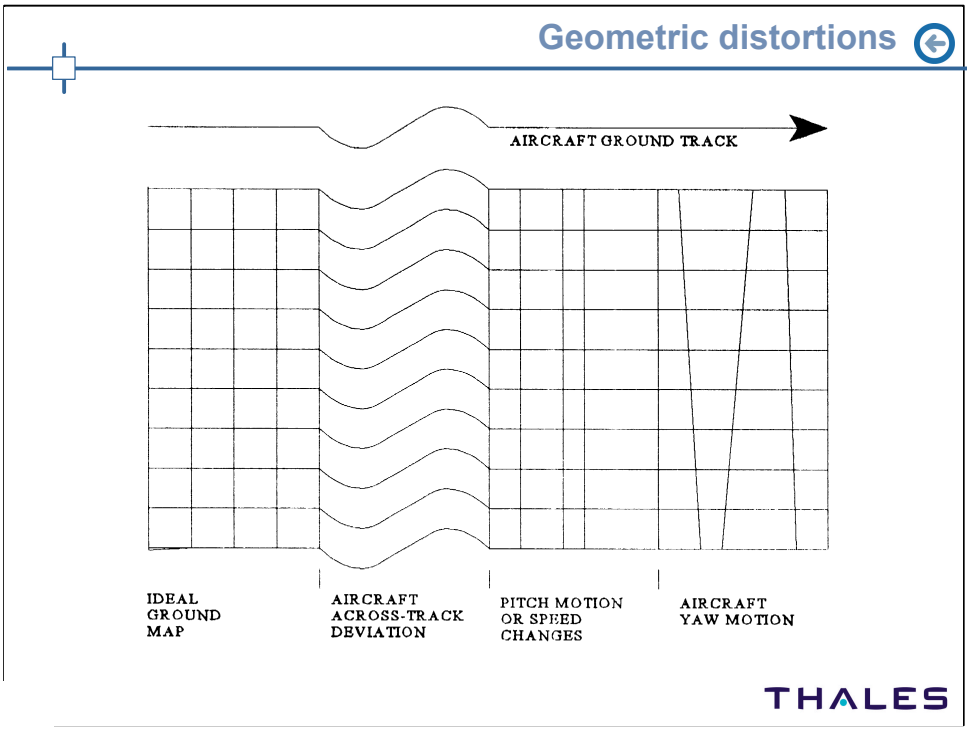
$$_e = 4 \ r_e/$$

- Uncertainty in slowly varying, non-linear motions along LOS should be $r_e <$ /16
- Uncorrected motion – timing and phase errors
- Geometric errors; defocusing; geolocation errors
  - Along track velocity error – defocused image
  - Across-track velocity error – rotated image
  - Across-track acceleration – defocused image

THALES

The derivation of ideal SAR performance assumes radar samples taken at precisely spaced intervals (< /2) along a perfectly linear track.

It is not possible in practise to achieve this precision from an airborne platform and the radar returns must be compensated to correct for deviations from the ideal motion and sampling intervals.



SAR ERRORS

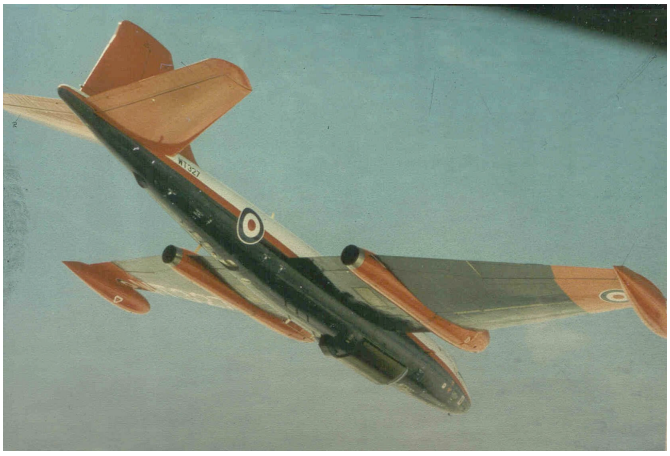
In all the basic analyses presented so far, the assumption has been made that the radar antenna is travelling in a straight line and with constant velocity and height. In practice an airborne platform will be subject to considerable deviations from the ideal for all these parameters. The correction of a wide range of motion errors is central to the successful operation of a SAR radar and an understanding of the effects of uncompensated errors is essential for the predicting of performance for a practical system. The following sections briefly outline the main sources of error which must be addressed in a practical system.

Beam Pointing Errors

For a side-looking SAR the real beam must be maintained with a constant beam pointing direction relative to the aircraft track. Typically the SAR processor is arranged to process returns centred around the zero doppler and the antenna must point at 90 degrees to the track. In practice the aircraft may yaw resulting in movement of the antenna beam. This movement may result in failure to illuminate the ground correctly with the returns now having an offset centre doppler frequency. If the zero doppler region falls outside the mainbeam, the SAR processor will not provide an output and characteristic “yaw bands” will be seen on the processed map.

The ideal solution is to stabilise the antenna in yaw and pitch (pitch errors have a similar effect if the beam is substantially depressed from the horizontal) by the use of inertial motion sensors and servo controls. For correction in one axis only, clutter locking techniques which centre the mainbeam doppler on zero frequency can also be used. An alternative solution to mechanical stabilization is to sense the mean doppler signal of the returns and to apply an appropriate pulse to pulse phase correction. This will allow the SAR processor to focus the returns but the map will be distorted as the yaw angle changes due to the changing direction of look relative to the track.

Canberra SAR



X-band experimental SAR system built by THORN EMI Electronics (now Thales) for RSRE (now QinetiQ)

THALES

Motion Errors

Along Track Velocity Errors

If the along track velocity is incorrect, the phase correction applied by the SAR processor will be mismatched. The mismatch will be similar to that observed due to the incorrect range being used (see section on depth of focus). The principal result of an along track velocity error will be a defocusing of the SAR image.

Allowing a /2 phase error at the extreme of the synthetic aperture gives a maximum along track constant velocity error v having a value given by

$$v \qquad \frac{R_0}{2v\,T^2}$$

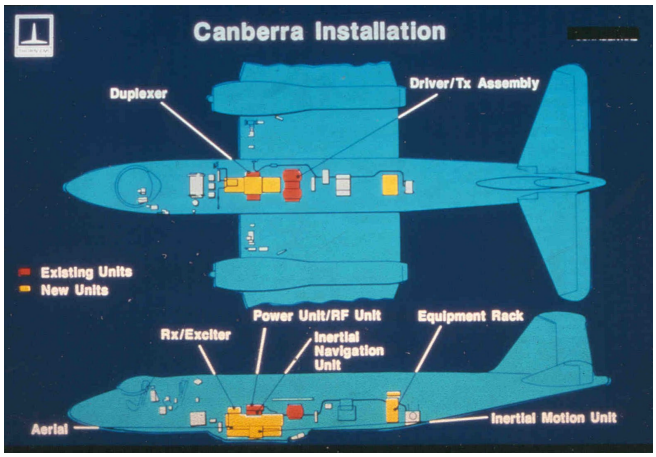
In addition to defocusing the image, an uncorrected along track velocity error will also result in a scaling error in the resulting map since it is assumed that the radar pulse intervals correspond to a certain spatial distance travelled over the ground. It is common practice to servo the radar prf to the aircraft velocity to minimise this effect.

Across-Track Motion Errors

An airborne platform will generally experience across-track acceleration as well as the rotations discussed in the section on beam pointing errors.

A constant across-track velocity will simply provide a doppler offset to the returns and can be interpreted as a change in aircraft track with a constant yaw error on the beam pointing direction. Very low frequency accelerations can therefore be likened to a variable beam pointing error with similar consequences for the processing. The velocity error must be corrected by appropriate phase shifting of the returns and any range walk must be accommodated by tracking the radar range gates.

Canberra installation



Radar installation, including INS and IMU for motion compensation

THALES

This figure shows the radar installation in the Canberra bomb bay.

The antenna was installed on a 3-axis stabilised scanner mounting. This allows the antenna pointing direction to be maintained independent of the aircraft rotation (pitch, roll and yaw); the antenna could also be controlled in tilt.

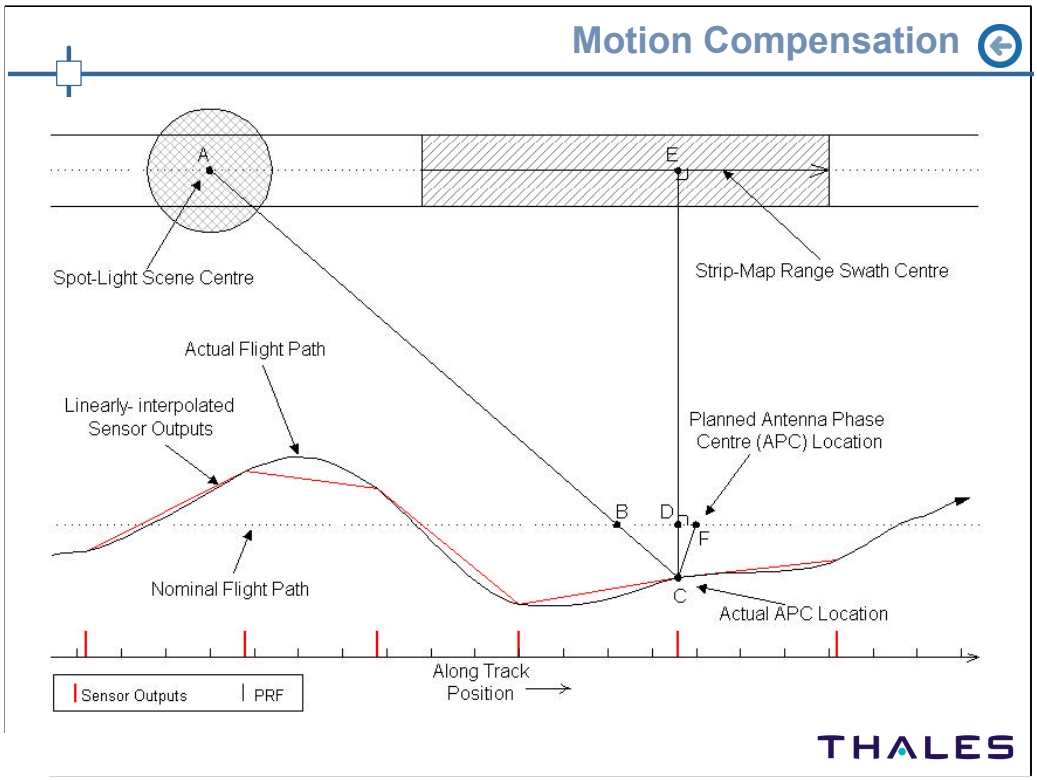
Fixed to the antenna is an IMU – three axis gyros and accelerometers.

Inboard, near the radar, is a GPS aided INS system

Navigation sensors for motion sensing

- IMU, INS, GPS
- IMU on antenna – as close as possible to antenna phase centre
- INS/GPS – on antenna for large systems or on aircraft as close as possible to antenna scanner

THALES



During the time taken to synthesise an aperture, the radar must measure the deviation of the antenna phase centre from a pre-defined trajectory (usually a straight line) to accuracies of a fraction of the radar wavelength of 3 cm.

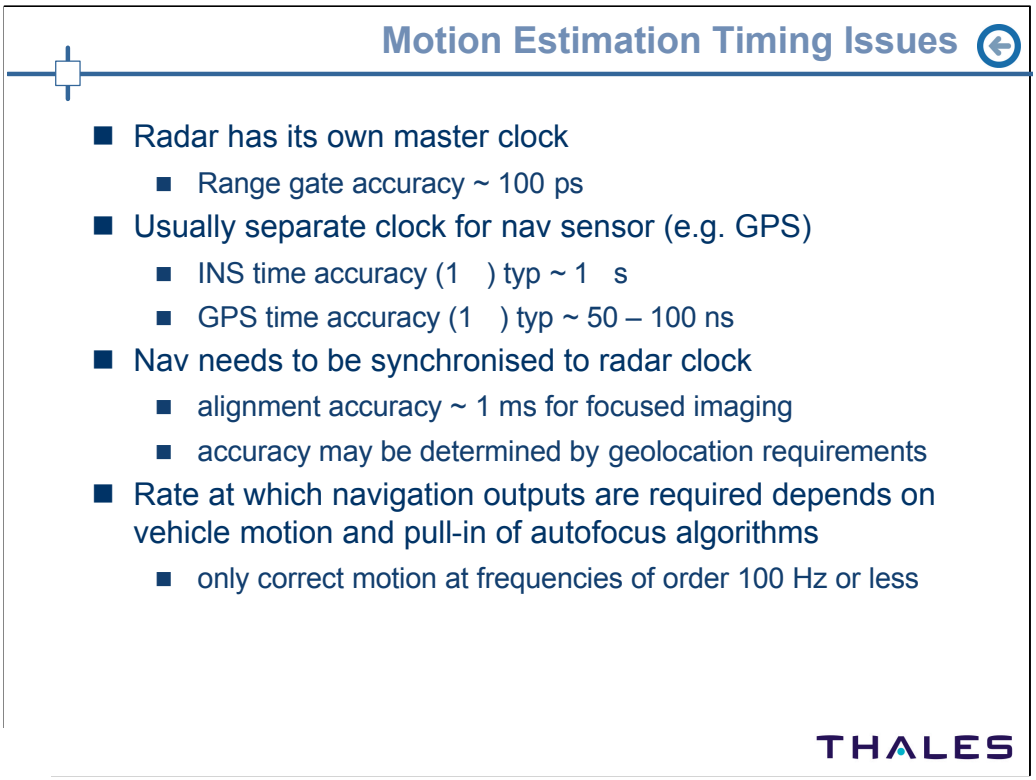
This information is then used to correct the timing and phase of the radar pulse returns to make them appear as if they had come from equally spaced points along a straight line.

The axis along the bottom shows the timing of the radar pulses (PRF = pulse repetition frequency) and the outputs from the inertial sensor. The phase and timing corrections to be applied depend on the type of SAR processing (e.g. swath SAR or spotlight SAR).

Motion Compensation

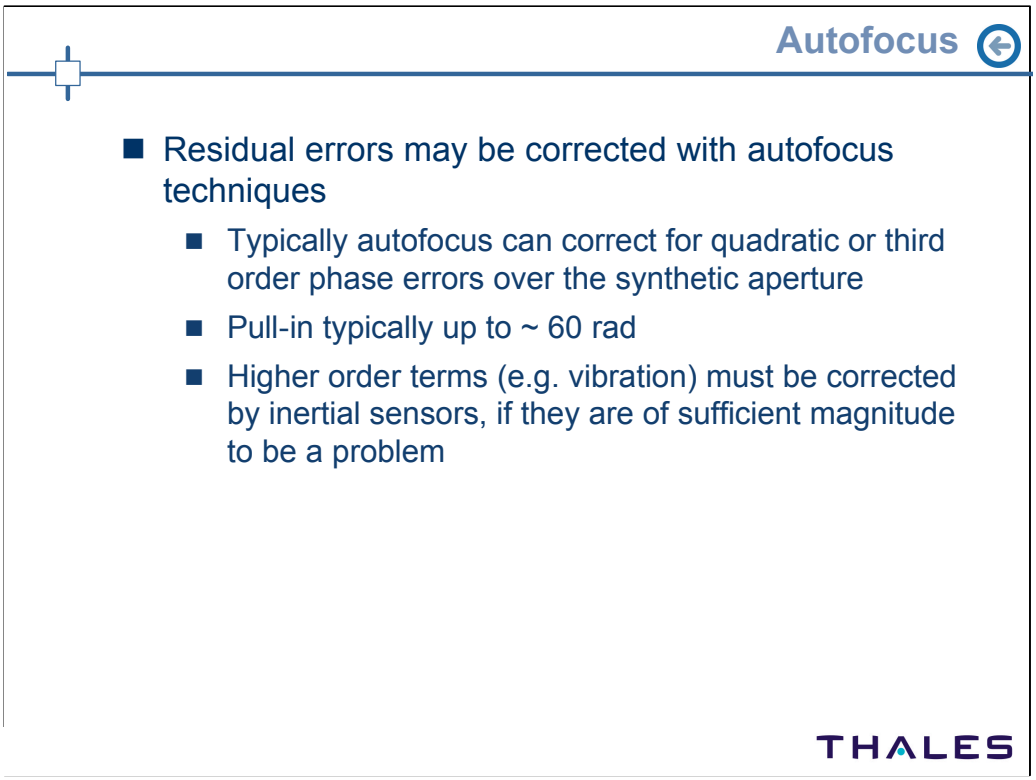
- Antenna phase centre motion estimation
- Coarse down-range alignment
- Fine down-range alignment
- Cross-range phase compensation

THALES

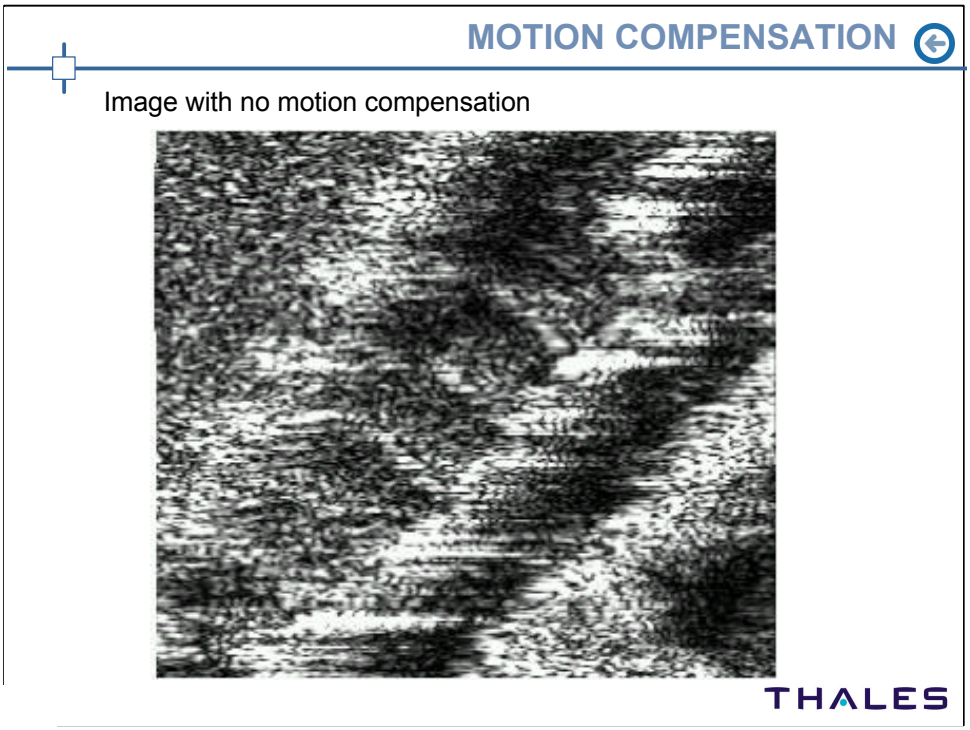


Most importantly, the sampled outputs of the inertial sensors must be accurately time-stamped by the same clock used to time the radar pulses. The outputs from the inertial sensors, often supported by GPS measurements for long-term stability, are processed to provide motion estimates, which in turn are used to correct the timing and phase of the radar pulse returns. This process can be very complex, requiring accurate transfer of measurements through different coordinate systems and between the inertial sensors and antenna phase centre. The correction of the radar returns often requires extensive interpolation between sampled pulse returns to reproduce the required samples in space and time over the synthetic aperture. Different corrections may be required for returns from different parts of the scene being imaged by the radar. In practice, inertial sensors alone cannot usually measure the full motion with sufficient accuracy, especially at the lowest frequencies of the motion spectrum. The typical drift rate of high quality unaided INS systems might result in a CEP of 10m after a flight of 7 min; over aperture time of 60 sec drift is likely to exceed the specification for complete motion correction.

Residual low-frequency errors must be corrected using autofocus techniques, as part of the radar image signal processing.



The inertial sensor must provide sufficiently accurate correction for the autofocus to function and must correct for higher frequency motion.

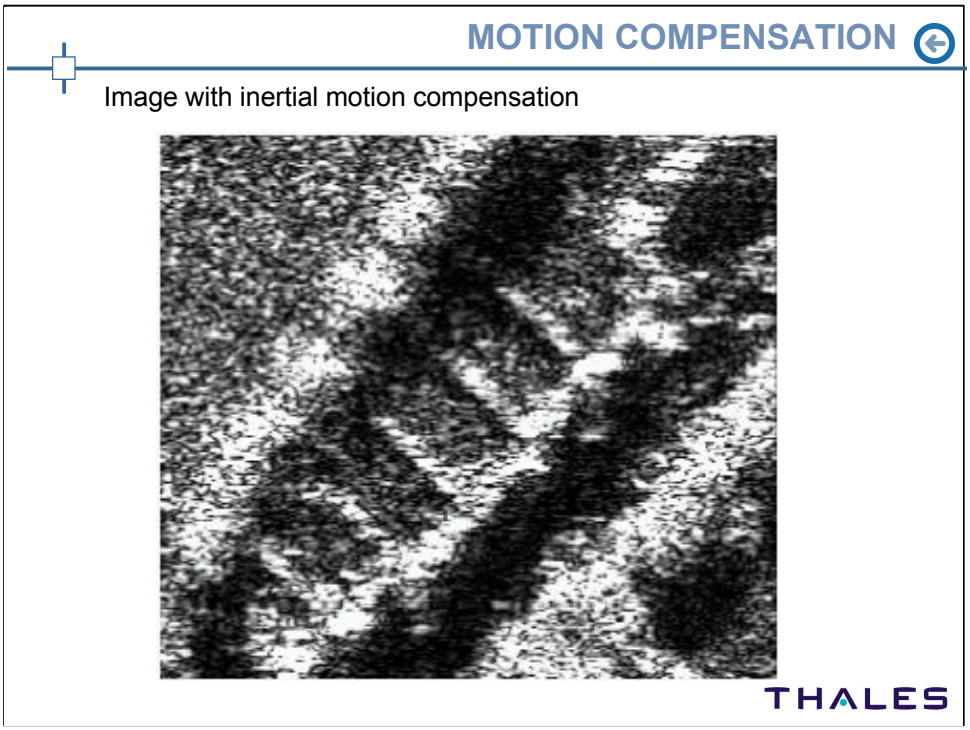


This is an image of houses with gardens, divided by fences. No motion compensation has been applied. Clearly the aircraft was flying a reasonably straight and level, as some imaging was achieved. However, the image is not well focused.

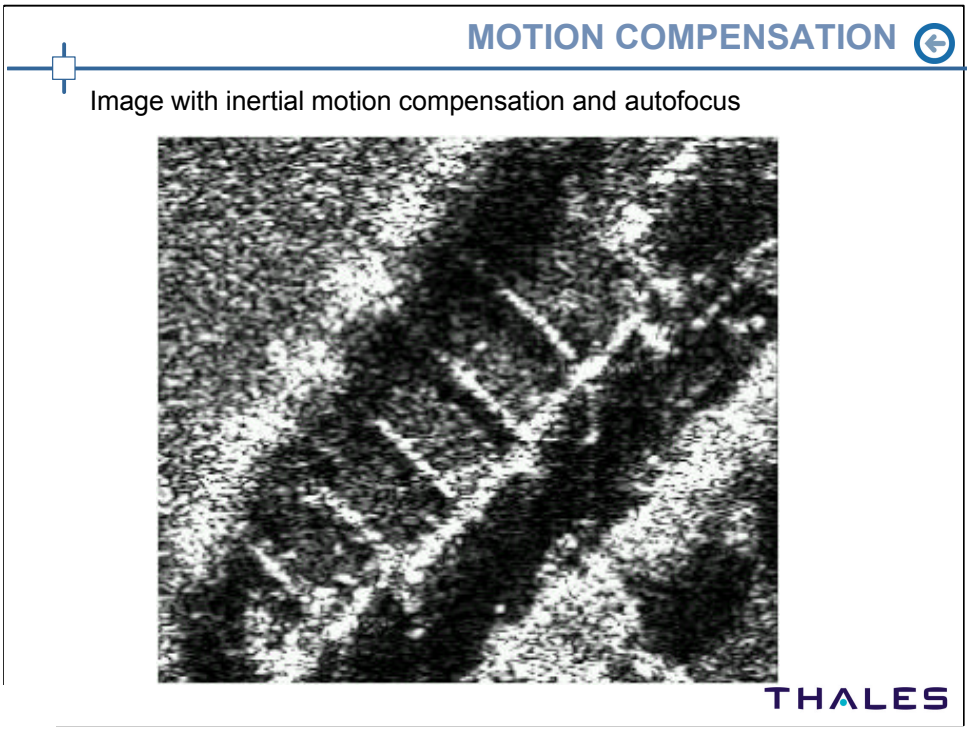
The quality of the final SAR image will be limited by a number of factors. These include amongst others, image speckle, range and azimuth sidelobes in the image response function, defocused images and geometric distortions. A single-look SAR image suffers from a problem known as “speckle” generally observed with coherent imaging systems, such as, for example, laser holography. The returns from a uniform background appear with random fluctuations from one resolution cell to another. This speckle is observed because the returns from a SAR map cell are comprised of the backscatter from a large number of randomly distributed scatterers within the cell. The returns from individual scatterers added with random phase, resulting in a Gaussian distribution of amplitude in the radar receiver at i.f.

A highly speckled image is characteristic of a “single-look” SAR image. The speckle fluctuations can be reduced by summing the envelopes of the map cell returns (i.e. non-coherent integration) for a number of independent “looks” at the map. Independent looks can be achieved in a number of ways. These include forming maps at different radar frequencies and by using azimuth multilooks). Speckle reduction can also be obtained by collapsing in range and summing adjacent cells non-coherently.

The overall image quality can be quantitatively characterized by a number of measures such as the “integrated sidelobe ratio” and the “small area contrast ratio”. An important part of image interpretation of SAR maps is the ability to detect changes between successive maps and to overlay them on geographic maps. For these purposes it may be important to eliminate the geometric distortions peculiar to radar imaging as well as providing a well focused image with low speckle and sidelobes. Many of the geometric distortions encountered are similar to those found in real aperture radars, whilst others will be introduced by errors in the SAR processing.

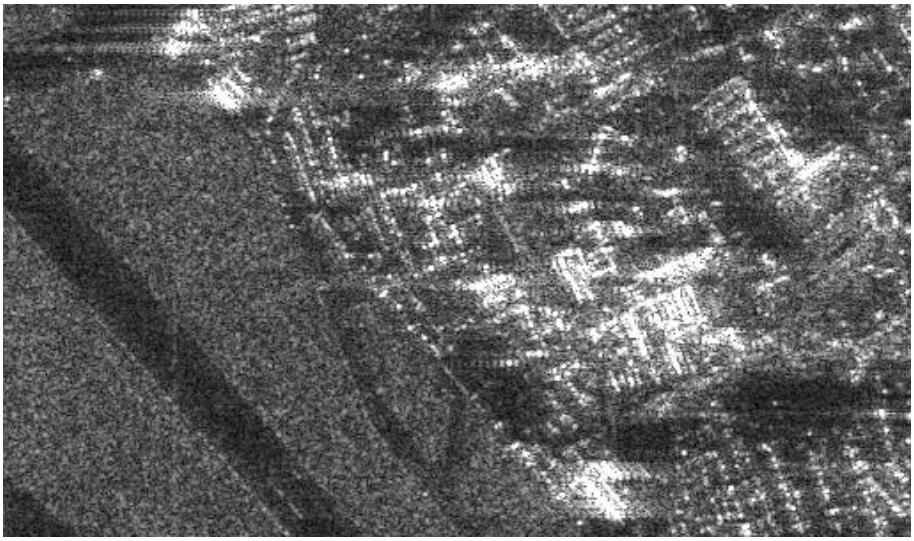


The same image with motion compensation applied using inertial sensor information.



Finally, the image with autofocus processing as well as inertial motion compensation.

SAR image, no range-gate interpolation



THALES

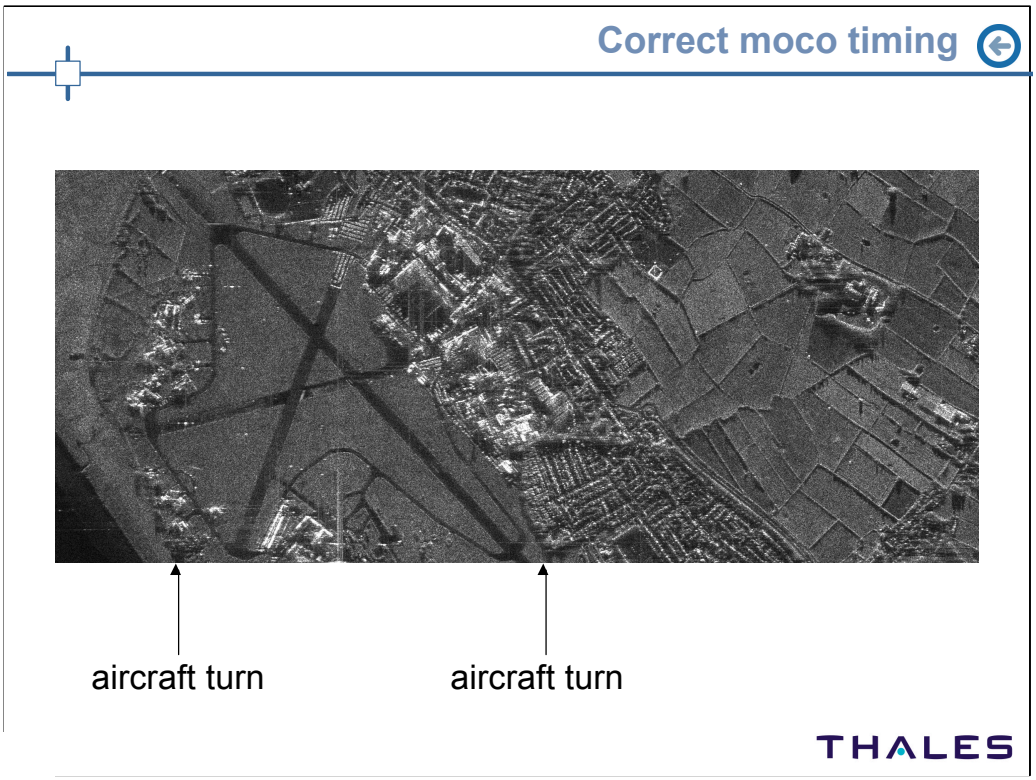
With high resolution imagery we must accurately correct the timing of the range samples, as well as the phase. Here is an example of an image where samples have only been corrected to the nearest range sample.

SAR Image, fully range-aligned

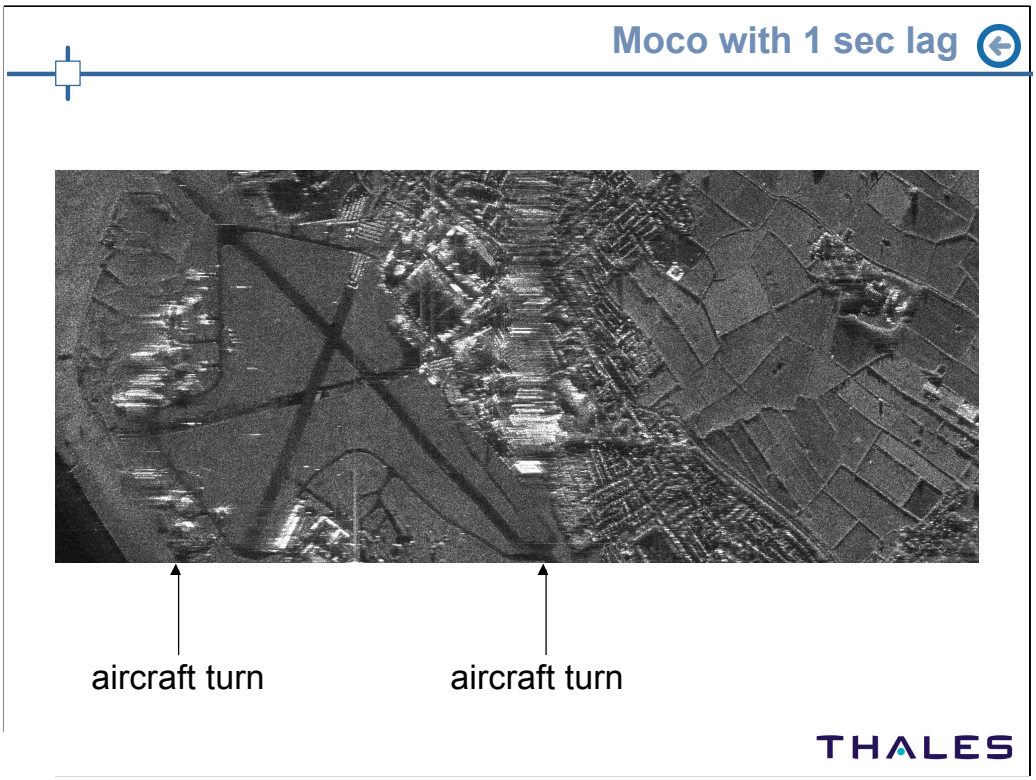


THALES

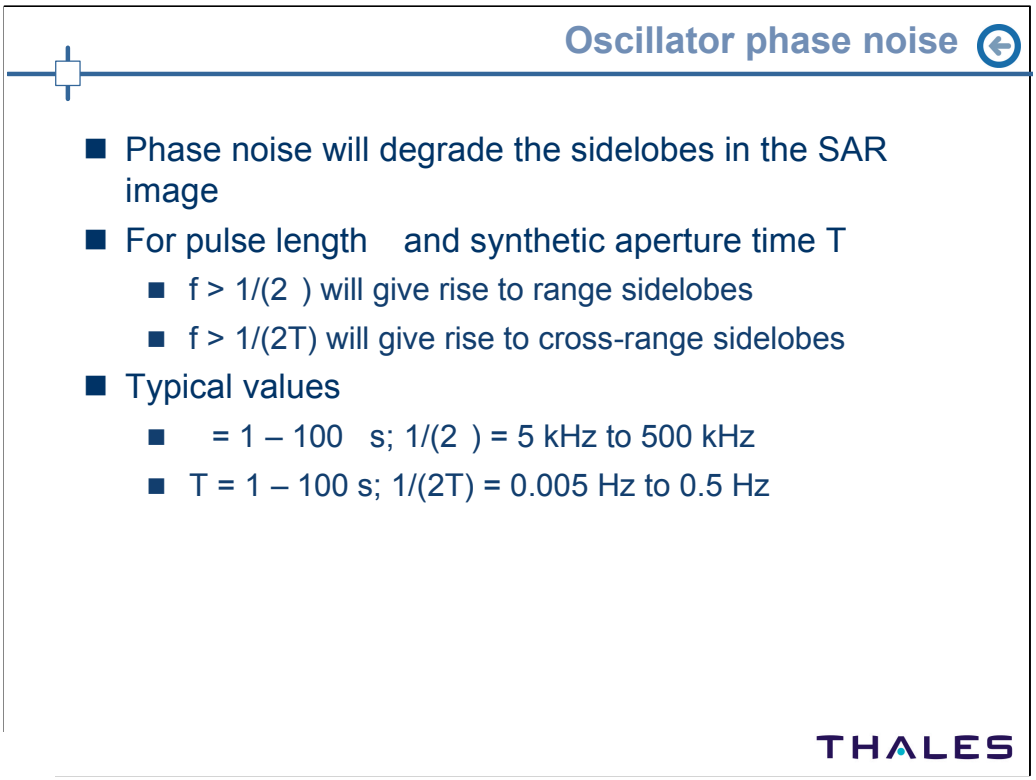
Here the range samples have been fully interpolated to give accurate range alignment.



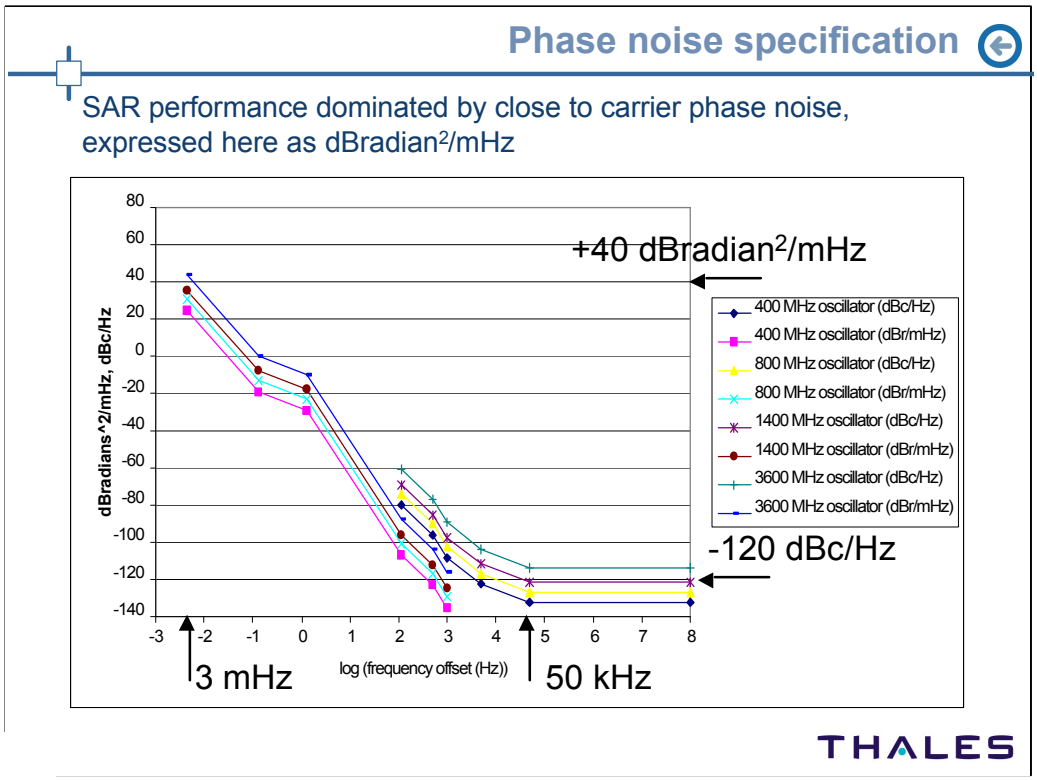
Along track motion compensation timing does not have to be highly accurate but gross errors can cause significant problems. Here is a well-focused image, which includes some turns, with linear motion between the turns.



Here the motion compensation data has a 1 sec time lag – now the corrections derived from the data are incorrect in the turns, giving rise to blurring and geometric distortions.

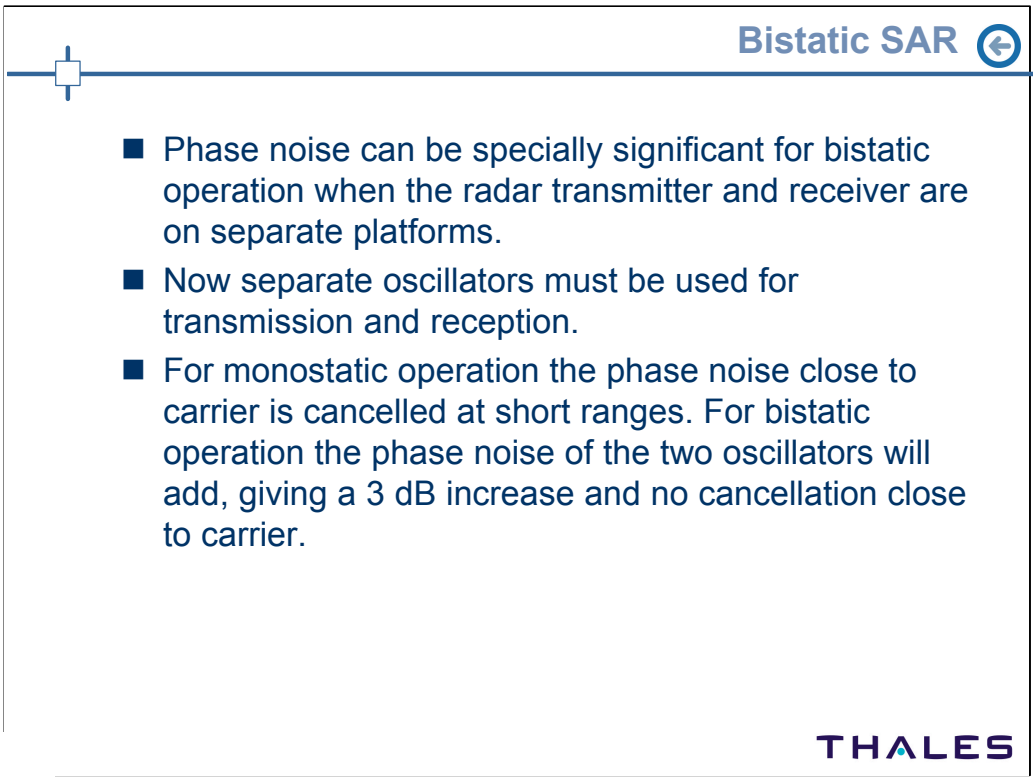


An addition source of error or distortion on the image is phase noise from the radar local oscillator used to generate and detect the radar signals.

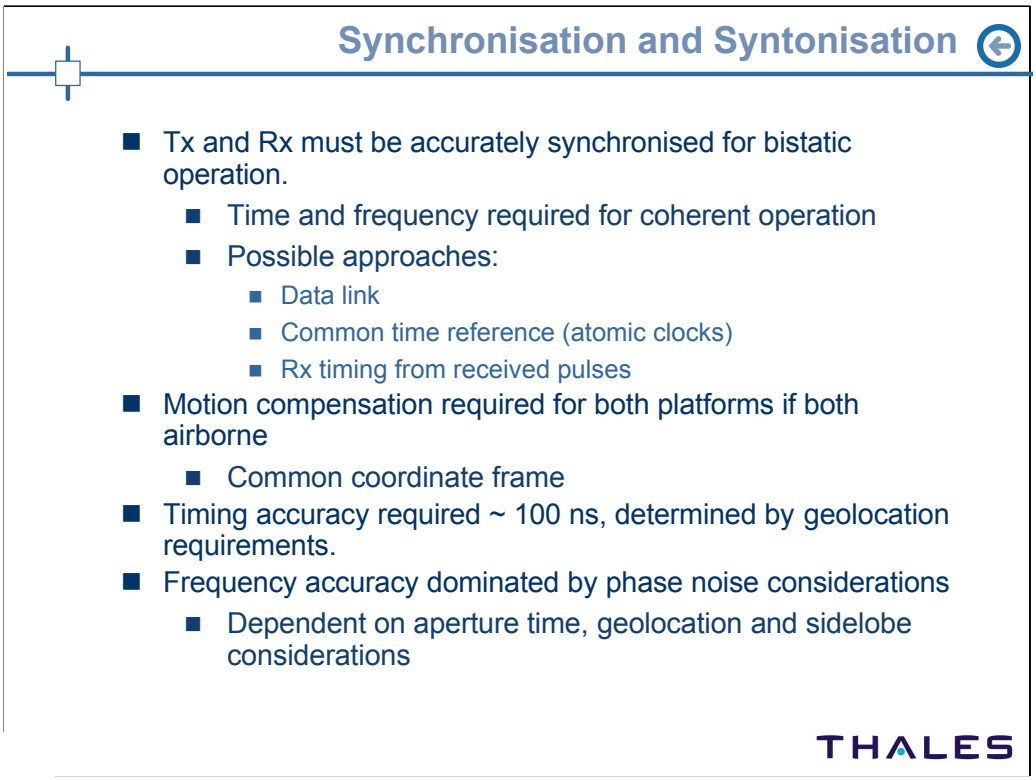


This is a typical specification for the phase noise power spectral density for a series of oscillators in a radar (the higher frequency signals are assumed to have been derived from the 400 MHz oscillator). For normal Doppler radar, the aperture time T is small and oscillator phase noise away from the carrier dominates.

For SAR it noise very close to carrier that is relevant. Here the modulation index for the phase noise is very large and the phase noise has been specified as dB radian^2/mHz. Higher frequency noise is specified as dBc/Hz.



A particular application when phase noise may be an issue is bistatic radar operation. This also has special timing problems.



Bistatic radar must be synchronised and matched in frequency.

SAR Imagery

- There are many examples of SAR imagery to be found on the internet.
- Interesting sources include:
  - www.sandia.gov/radar/
  - www.lynxsar.com/images.html
  - http://www2.jpl.nasa.gov

THALES

Conclusions

- The processing of radar returns to produce well-focused SAR imagery requires complex signal processing, including accurate motion compensation.
- Local oscillator noise requirements are not excessively demanding, although SAR systems require better close-to-carrier noise performance than other coherent radar processes.
- Future work on airborne systems is aimed at both very long range systems on fast air platforms and high resolution short-range systems for unmanned air vehicles (UAV) employed in battlefield surveillance.
- For UAV applications, the challenge is to achieve high performance with very light weight and low volume. The challenge is to understand what minimum quality of inertial motion sensors and local oscillators is sufficient for these applications.

THALES

Session 4C / Optical Frequency Standards

Optical Lattice Clock

Masao Takamoto[1], Feng-Lei Hong[3], Ryoichi Higashi[1] and Hidetoshi Katori[1,2]
[1]Department of Applied Physics, Graduate School of Engineering, The University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[2]PRESTO, Japan Science and Technology Agency, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[3]National Institute of Advanced Industrial Science and Technology (AIST)
Tsukuba Central 3, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8563, Japan
E-mail: katori@amo.t.u-tokyo.ac.jp

ABSTRACT

Precision spectroscopy was performed for the $5s^2\ {}^1S_0$ (F=9/2)-$5s5p\ {}^3P_0$ (F=9/2) clock transition of ^{87}Sr atoms trapped in a one-dimensional optical lattice. The transition frequency for the Sr lattice clock was measured to be 429,228,004,229,952(15) Hz by an optical frequency comb referenced to the SI second.

1. INTRODUCTION

Since the demonstration of the optical frequency comb technique [1,2] that established a coherent link between optical and radio frequencies, optical clocks have gained increasing interest as future atomic clocks. Optical clocks promise better stability and accuracy than Cs microwave clocks because of their higher carrier frequencies of up to 1000 THz. To date, optical clocks with single trapped ions [3,4,5] and ultracold neutral atoms in free fall [6,7] have shown a record high performance that is approaching that of the best Cs fountain clocks [8].

A single ion trapped in an RF quadrupole field is an ideal system for realizing optical clocks, as the trap prepares a quantum absorber completely at rest in free space for an extended time and its electric field vanishes at the center of the trap [9]. Employing this scheme, quantum projection noise (QPN)-limited spectroscopy has been performed [10] with a projected accuracy of 10^{-18}, whereas the stability of the single-ion based optical clocks is severely limited by the QPN and a long averaging time is required to meet its ultimate accuracy. One may think of increasing the number of quantum absorbers N as employed in optical clocks using neutral atoms. In this case, however, the atom-laser interaction time sets upper bounds for the effective transition line Q-factor [11], which is two orders of magnitude smaller than that for ion clocks [10], since an atom cloud in free space expands with finite velocity and is strongly accelerated by the gravity during the measurement. Furthermore, the residual Doppler shifts as a result of the imperfect wavefronts of the probe beams and atom-atom collisions during the measurement may affect its ultimate accuracy [6].

In this paper, we demonstrate an "optical lattice clock" [12] that utilizes millions of neutral atoms prepared in an optical lattice. The sub-wavelength localization of atoms in each lattice site suppresses the first-order Doppler shift as well as the collisional frequency shifts; therefore the scheme simulates millions of single-ion clocks operated simultaneously. In striking contrast with conventional approaches toward optical clocks, this optical lattice clock scheme interrogates atoms while they are strongly perturbed by an external field. We have shown that this perturbation can be cancelled out to a precision level below 10^{-17} [13] by designing the light shift trap so as to adjust the dipole polarizabilities for the electronic states used as a clock transition. By applying the scheme to the $5s^2\ {}^1S_0$ (F=9/2)-$5s5p\ {}^3P_0$ (F=9/2) clock transition of ^{87}Sr atoms [14], our recent experiment resolved the clock transition with a linewidth of 27 Hz, which is one order of magnitude narrower than that observed for atomic clocks in free fall [6,7], and the clock stability was estimated to be comparable to that of single-ion clocks [10].

2. LIGHT SHIFT CANCELLATION

In general, such a lattice-trapping field significantly modifies the internal states of atoms by so-called light shifts; hence this system was not seriously considered for atomic clocks until the light shift cancellation technique was demonstrated [14]. The transition frequency ν of atoms exposed to a lattice trapping laser with intensity I is given by the sum of the unperturbed transition frequency ν_0 and the light shift ν_{ac},

$$\nu(\lambda_L,\mathbf{e}) = \nu_0 + \nu_{ac} = \nu_0 - \frac{\Delta\alpha(\lambda_L,\mathbf{e})}{2\varepsilon_0 ch} I + O(I^2). \quad (1)$$

Here ε_0, c, and h are the permittivity of free space, the speed of light, and the Planck constant, respectively, and $\Delta\alpha(\lambda_L,\mathbf{e}) = \alpha_u(\lambda_L,\mathbf{e}) - \alpha_l(\lambda_L,\mathbf{e})$ is the difference between the ac polarizabilities of the upper and lower states, which depends on the lattice laser wavelength λ_L and its polarization vector $\mathbf{e}$. By adjusting these polarizabilities to satisfy $\Delta\alpha(\lambda_L,\mathbf{e})=0$, the observed atomic transition frequency ν will be equal to ν_0 independent of the lattice laser intensity I as long as the higher-order corrections

$O(I^2)$ are negligible. In the search for a clock transition in which the light shift cancellation condition $\Delta\alpha(\lambda_L,\mathbf{e})=0$ can be mainly determined by a lattice laser wavelength λ_L, it is preferable to use the $J=0$ state, which exhibits a scalar light shift and does not depend on the light polarization. In this experiment, we adopted the $5s^2\ ^1S_0$ (F=9/2) → $5s5p\ ^3P_0$ (F=9/2) forbidden transition of ^{87}Sr with a nuclear spin of 9/2 as the clock transition (Fig.1) [12], in which hyperfine mixing provides a finite lifetime of Γ_0^{-1}=150 s.

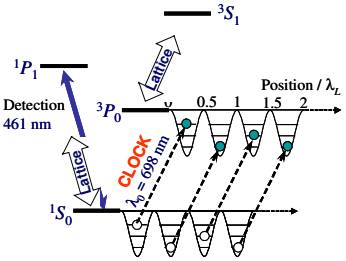


Figure 1: Strontium energy levels. The $^1S_0$ and $^3P_0$ states are coupled to the respective upper spin states by an off-resonant laser to produce an optical lattice with equal energy shifts in the clock transition. Atoms are excited on the electronic-vibrational transitions in the lattice potentials.

Figure 2 shows the calculated differential light shift ν_{ac} for the $^1S_0$ - $^3P_0$ clock transition of strontium as a function of the lattice laser wavelength λ_L, where we assumed the lattice laser intensity of I =10 kW/cm^2 (solid line) and I =5 kW/cm^2 (dotted line). By tuning the lattice laser to $\lambda_L \approx$ 800 nm, the $^1S_0$ and the $^3P_0$ states have the same amount of light shift; therefore the light shift ν_{ac} for the clock transition vanishes. At around this *magic* wavelength, the light shift can be controlled with $|\nu_{ac}| <$ 1 mHz by defining the lattice laser frequency only within 1 MHz. Other systematic effects, such as the vector and tensor light shift arising from the hyperfine splitting, are summarized in the right column of Table 1 and in ref. [13]. In these calculations, we assumed the typical lattice laser intensity of I =10 kW/cm^2.

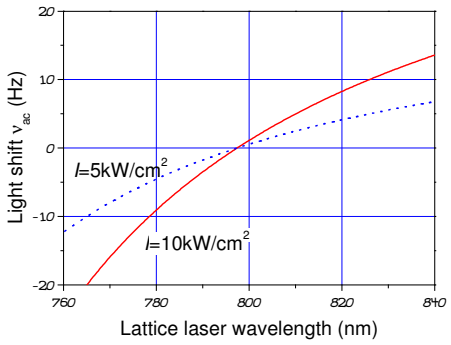


Figure 2: Light shift ν_{ac} on the clock transition as a function of the lattice laser wavelength for a lattice laser intensity of I =10 kW/cm^2 (solid line) and I =5 kW/cm^2 (dotted line).

3. EXPERIMENT

3.1 A clock laser

A laser diode (Hitachi, HL6726MG, λ_{typ}=685 nm, P_0=30 mW), which is loaded with an external cavity and tuned to λ_0=698 nm, is used as a clock laser. The laser is stabilized to a reference cavity made of ULE (ultra low expansion) glass with the finesse of 30,000. The temperature of the ULE grass was stabilized at around T = 3.0 °C, at which the material approached a zero expansion coefficient and the clock laser showed a frequency drift rate of 0.13Hz/s. The cavity is put inside a vacuum chamber (at a pressure of 7×10^{-8} Torr), which is placed on a passive vibrational isolation table (nano-K model 150BM-1). The whole cavity system is set in a wooden box, the inside of which is lined with urethane foam for acoustic absorbance. A beat note between the two lasers, which were locked to different reference cavities constructed as described above, indicated laser linewidth of 10 Hz. We presume that this linewidth is mainly introduced by an acoustic noise that entered the reference cavity because of an imperfect acoustic isolation of a noisy laboratory environment. Further reduction of the laser linewidth is under way.

3.2 Laser cooled ^{87}Sr in an optical lattice

The experimental configuration for the lattice clock is shown in Fig. 3a [15]. ^{87}Sr atoms were laser-cooled and trapped on the $^1S_0$-$^3P_1$ transition. Roughly 10^4 atoms with a temperature of ≈ 2 μK were loaded into a 20-μK-deep 1D optical lattice that was formed by a standing wave of a Ti-sapphire laser and provided atoms with subwavelength confinement along its axial direction. The clock laser was sent into the same fiber that delivered the Ti-sapphire laser to set its wave-vector parallel to the axis of the 1D lattice potential so as to satisfy the Lamb-Dicke condition. Both lasers were passed through a polarizer with an extinction ratio of 10^{-5} to guarantee linear polarization. We note that for this π polarized lattice laser, the frequency spread due to the tensor as well as the vector light shift among Zeeman sublevels m is less than 1 mHz [13].

3.3 Determination of the magic wavelength and spectroscopy

We irradiated the clock laser for 10 - 40 ms to excite atoms trapped in the optical lattice to the $^3P_0$ state. The excitation of the clock transition was then monitored by the laser-induced fluorescence on the $^1S_0$ - $^1P_1$ cyclic transition. The magic wavelength for the lattice laser was determined as follows. We measured the lattice laser intensity (I)-dependent light shift ν_{ac} at several wavelengths λ_L near 813.5 nm [15] as shown in Fig. 4a.

The slope of the fitted lines, which gives the differential light shift $\delta\nu_{ac}(\lambda_L,I)/\delta I$ and is proportional to the polarizability difference $\Delta\alpha(\lambda_L)$ (see Eq. (1)), is plotted as a function of the lattice laser wavelength λ_L in Fig. 4b. The *magic* wavelength, where $\Delta\alpha(\lambda_L) = 0$ holds, is therefore determined as 813.420(7) nm. We note that this uncertainty introduces a scalar-light-shift uncertainty of 4 Hz (see Table 1).

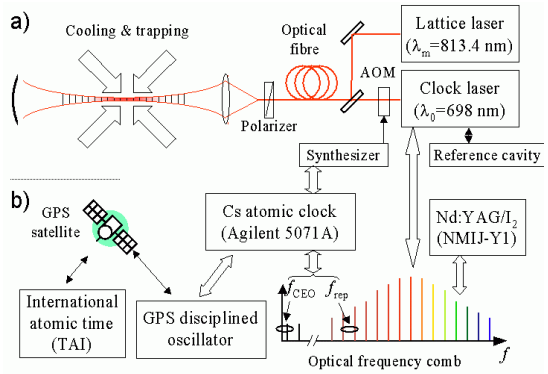


Figure 3: The schematics of the experiment. a) Ultracold Sr atoms were trapped in a one-dimensional optical lattice formed by a standing wave of a Ti-sapphire laser tuned to a magic wavelength. The clock laser was introduced along the fast axis of the lattice potential to guarantee the Lamb-Dicke confinement. b) The frequency of the clock laser was measured by an optical frequency comb referenced to the SI second.

Figure 5 shows the spectrum of the clock transition measured at the magic wavelength; each empty circle was measured with an interrogation time of 40 ms and a cycle time of 1 s, which is required for cooling, trapping, and spectroscopy. We observed a nearly Fourier-limited linewidth of 27 Hz, which is the narrowest linewidth yet reported for neutral-atom-based optical clocks [6,7,11]. With 20 data points that took ≈20 s to measure, the Lorentzian fitting of the spectrum determined the centre frequency with an uncertainty of 2 Hz, corresponding to an Allan deviation [16],

$$\sigma_y(\tau) \approx \frac{2\text{Hz}}{429\text{THz}} \frac{1}{\sqrt{\tau/20}} = 2\times10^{-14}/\sqrt{\tau}.$$

With the present measurement, the signal-to-noise ratio is limited by the shot-to-shot atom number fluctuation (≈10%) loaded into the optical lattice. The normalization of the excitation rate [6], in addition to the reduction of the clock laser linewidth [17], will improve the clock stability towards the projection noise limit for 10^6 atoms [6,8] of $\sigma_y(\tau) \approx 10^{-18}/\sqrt{\tau}$ [13] with τ measured in second.

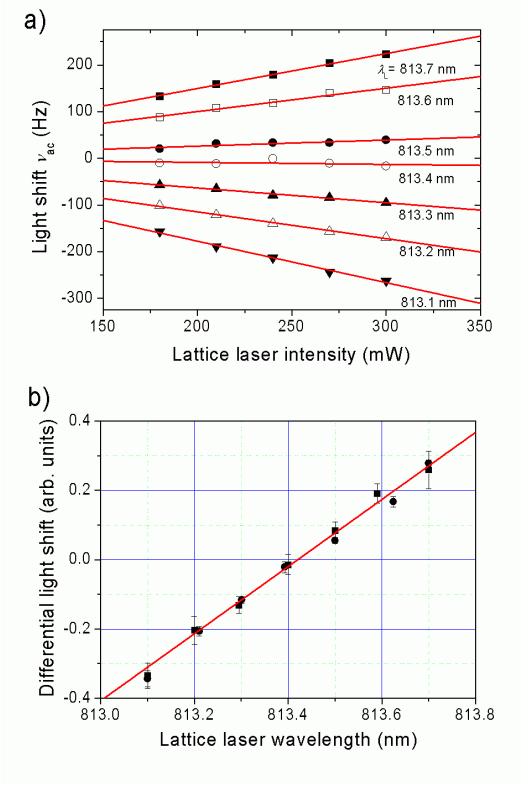


Figure 4 Determination of the magic wavelength. a) The light shift ν_{ac} of the clock transition was measured as a function of the lattice laser intensity I at several wavelengths λ_L near 813.5 nm. The differential light shift $\delta\nu_{ac}/\delta I$ was determined by the slope of the fitted lines. b) The differential light shift, which is proportional to the dipole polarizability difference $\Delta\alpha(\lambda_L)$, is plotted as a function of the lattice laser wavelength λ_L. The magic wavelength, where $\Delta\alpha(\lambda_L)$=0, was determined to be λ_L=813.420(7) nm by a linear fit to the data points.

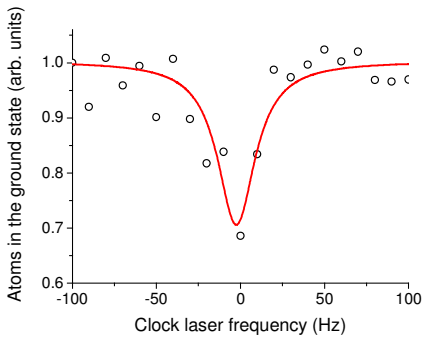


Figure 5: Spectrum of the clock transition at 429 THz. Nearly Fourier transform-limited linewidth of 27 Hz was obtained for the interrogation time of 40 ms for each empty circle. The uncertainty of the center frequency was 2 Hz given by the Lorentzian fitting of the measured spectrum.

4. ABSOLUTE FREQUENCY MEASUREMENT

The schematic for the absolute frequency measurement is shown in Fig. 3b. The frequency f_c of the clock laser prestabilized to a stable high-finesse reference cavity was frequency-scanned with an acousto-optic modulator (AOM) to find the atomic resonance frequency ν_0 every 20 s by observing a spectrum as shown in Fig. 5. Since the frequency setting of the AOM $\delta(t) = f_c(t) - \nu_0$ measured the temporal drift of the clock laser, the atomic resonance frequency ν_0 could be determined by simultaneously measuring the clock-laser frequency $f_c(t)$. We observed a beat frequency $f_b = |f_c - f_n|$ between the clock laser and the n-th comb component at 698 nm and used it to servo-control the f_{rep} of the comb. Consequently, the clock-laser frequency $f_c(t)$ was converted to f_{rep}, which was then mixed down to about 7 MHz and recorded by a frequency counter with a gate time of 20 s. The clock-laser frequency $f_c(t)$ was calculated by using the formula $f_c = n \times f_{rep} + f_{CEO} \pm f_b$ with an integer $n \approx 5.4\times10^5$. The ambiguous sign was removed by observing the sign of the variation in the beat frequency f_b, while f_{rep} was varied. We thus obtained two series of data sets; the offset frequency $\{\delta(t)\}$ and the corresponding clock-laser frequency $\{f_c(t)\}$, both of which were determined every t_m = 20 s. The atomic resonance frequency was calculated as $\nu_0(k) = f_c(kt_m) - \delta(kt_m)$ for $k = 1, 2, ..., k_{max}$, where $k_{max} = T_{max} / t_m$ with T_{max} typically 10^4 s. Each filled square in Fig. 6 represents the mean value of $\{\nu_0(k)\}$, whose standard deviation of $\sigma \approx$ 620 Hz was in agreement with the Cs clock's instability at τ = 20 s. The standard deviation of the mean $\sigma / \sqrt{k_{max}}$ gave error bars of about 28 Hz.

In an alternative experimental scheme, a transportable iodine-stabilized Nd:YAG laser (NMIJ-Y1) [18] was used as a flywheel in the frequency measurement. The instability of the Nd:YAG laser $\sigma_y(\tau)$ is 2×10^{-13} for τ=1 s. This value improves after τ=100 s towards 3×10^{-14}, which is smaller than that of the Cs clock until τ=10^5 s. Furthermore, the $\sigma_y(\tau)$ of the Nd:YAG laser is smaller than that of the clock laser for τ > 30 s, which makes it possible to observe the slow drift of the clock laser. Because of the excellent short-term stability of the Nd:YAG laser, we also performed a frequency measurement by phase-locking the comb to the Nd:YAG laser. The clock-laser frequency $f_c(t)$ was measured by simultaneously introducing a second beat measurement between the clock laser and the comb component at 698 nm. The measurement results obtained with this scheme coincide with those obtained with the first measurement scheme and are indicated by filled circles in Fig. 6.

The frequency of the Cs clock (Agilent 5071A) was monitored by the GPS time during the period of the frequency measurement (over one month). The timing difference between the 1 pps (pulse per second) signals generated by the Cs clock and a GPS-disciplined oscillator was recorded continuously using a time interval counter. The frequency offset between the Cs clock and the GPS time was calculated from the recorded timing difference and found to be $-1.04(8) \times 10^{-13}$. The relationship between the GPS time and the TAI is published on the web page of the Bureau International des Poids et Mesures (BIPM) [19]. The agreement between the GPS time and the TAI was about one part in 10^{15} during the period of the frequency measurement.

Figure 6 summarizes the absolute frequency of the Sr lattice clock referenced to the Cs clock. In the figure, each filled square/circle represents a frequency measurement of typically 10^4 s to reduce the fractional instability of the Cs clock down to 6.5×10^{-14} corresponding to an uncertainty of 28 Hz. With a total averaging time of $\tau \approx 9.4\times10^4$ s over 9 days, we reached an uncertainty of 9 Hz. The sources of the uncertainties and the corrections are listed in Table 1 to give a final value of 429,228,004,229,952(15) Hz [27], which is consistent with a measurement reported by Courtillot *et al.* [20] within their measurement uncertainty of 20 kHz.

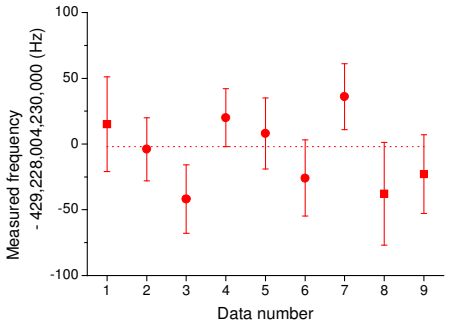


Figure 6: Absolute frequency measurement of the $^1S_0$ - $^3P_0$ transition of Sr atoms in an optical lattice. The resonance frequency was measured using an optical frequency comb generator referenced to a commercial Cs clock. Each filled square/circle corresponds to a frequency measurement over 10^4 s that gave a typical uncertainty of 28 Hz. The 9 days of measurement determined the clock transition to be 429,228,004,229,952(15) Hz, including the systematic correction of –45.7 Hz.

5. DISCUSSION AND UNCERTAINTIES

Table 1 gives the uncertainty budgets for this experiment as well as those expected for future experiments. In the following we report present uncertainties and some technical improvements that could be made in the future.

The first-order Doppler shift may occur due to the relative motion between the clock laser and the lattice potential, which is often introduced by the vibration of the optical fibre or air turbulence where the clock laser travels. Assuming a clock laser linewidth of 10 Hz to be

this Doppler width, the frequency uncertainty may decrease down to 30 mHz after averaging 10^5 measurements. By applying fibre phase-noise cancellation by referring to a reflection from a mirror surface that forms the standing wave for the lattice potential, this Doppler shift could be actively stabilized to less than 1 mHz [21] even for an averaging time of 1 s.

The first-order Zeeman shift for Δm=0 transitions is m ×106 Hz/G. In this experiment the residual magnetic field of 22.5 mG and random population distribution among m substates could be responsible for the Zeeman shift or broadening up to 10 Hz. The reduction of the magnetic field down to the μG level would be feasible by applying a magnetic field shielding.

The collisional frequency shift, which was measured by varying the atom density by a factor of 3, was less than a measurement uncertainty. However, the shift may be present in this 1D lattice configuration, since tens of atoms were typically trapped in a single lattice potential with an atom density of $\approx 10^{12}$ cm^{-3}. This collision shift would ultimately be removed by spin-polarizing the ^{87}Sr atoms to activate the Fermi suppression of collisions [22] or by applying a 3D lattice [12, 13] with less than unity occupation. In the latter case, the frequency shift due to the resonant dipole-dipole interaction energy between atoms separately confined in the 3D lattice potential with an interatomic distance of $\lambda_L/2 \sim 400$ nm is given in the Table 1.

The Blackbody shift scales as $\nu_B \approx 3\times10^{-10}\times(T/\mathrm{K})^4$ Hz, which would introduce a 3 mHz uncertainty for the temperature imhomogeneity of $\Delta T = 0.1$ K at $T = 293$ K.

The vector light shift can be problematic in 3D optical lattices. A phase stable optical lattice could be applied, where the 3D laser intersection is formed by a single linearly-polarized standing-wave with folded mirrors [23]. In this configuration, since the amplitude and the phase of the trapping light fields are correlated in each lattice site, high polarization stability is expected.

The forth-order light shift given in Table 1 is estimated without including the resonant contribution [13]. The magic wavelength that we determined in this experiment can be used to evaluate the resonant contribution of a fourth-order light shift. The $5s7p\ ^1P_1$ state connected to the clock upper state ($5s5p\ ^3P_0$) via the $5s6s\ ^3S_1$ state with 2 photons at $\lambda = 813.359$ nm may play a decisive role. However, as indicated in Fig. 4, its influence, which can manifest itself as the light shift quadratic to the lattice laser intensity or the dispersive behavior of the light shift near the two-photon resonance, is not visible within the present experimental precision. It will be accurately estimated once a dipole moment on the $5s7p\ ^1P_1$ - $5s6s\ ^3S_1$ transition is available.

Effect	Correction (Hz)	Uncertainty (Hz) Achieved	Attainable
1st order Doppler	0	3×10^{-2}	$<10^{-3}$
2nd order Doppler	0	2×10^{-6}	$<2\times10^{-6}$
Recoil shift	0	0	0
1st order Zeeman	0	10	10^{-3}
Collision shift	0.6	2.4	10^{-5}
Blackbody shift	2.4	0.1	3×10^{-3}
Probe laser light shift	0.1	0.01	10^{-3}
Scalar light shift	-3.8	4	10^{-3}
Vector light shift	0	10^{-3}	10^{-3}
Tensor light shift	0	10^{-3}	10^{-3}
4th order light shift	0	10^{-3}	10^{-3}
Cs clock offset	-45	3	
Frequency measurement	0	9	
Systematic total	-45.7		
Total uncertainty δν		15	4×10^{-3}

Table 1: Uncertainty budget for the optical lattice clock with strontium.

6. SUMMARY

In summary, we have demonstrated a novel atomic clock utilizing neutral atom ensembles confined in a light-shift free optical lattice, which would surpass conventional clock schemes in stability and accuracy [13]. This lattice clock scheme is applicable to other divalent atoms that have a hyperfine-mixing induced J=0→J=0 transition between long-lived states, such as ^{43}Ca [24], ^{171}Yb [25], and ^{199}Hg, in addition to the ^{87}Sr used here. A frequency comparison of lattice clocks operated with different atomic species will allow us to study the time variation of the fine structure constant [26]. The very high potential stability of this scheme would enable us to measure the fractional frequency difference at the 10^{-18} level, which corresponds to the gravitational red shift for a 1 cm height difference, in only one second. These features will have an immense impact on future metrology that will range from fundamental physics to such aspects of engineering as the remote sensing of natural resources.

ACKNOWLEDGEMENTS

The authors are grateful to M. Yasuda, Y. Fukuyama and J. Jiang for their assistance with the experiments and A. Onae and S. Ohshima for their participation in helpful discussions. This work received support from the Strategic Information and Communications R&D Promotion Programme (SCOPE) of the Ministry of Internal Affairs and Communications of Japan.

REFERENCES

[1] Th. Udem, J. Reichert, R. Holzwarth, and T. W. Hänsch, Absolute optical frequency measurement of the

cesium D_1 line with a mode-locked laser, *Phys. Rev. Lett.* **82,** 3568-3571 (1999).
[2] D. J. Jones *et al.* Carrier-envelope phase control of femtosecond mode-locked lasers and direct optical frequency synthesis, *Science* **288,** 635-639 (2000).
[3] S. A. Diddams *et al.* An optical clock based on a single trapped $^{199}Hg^{+}$ ion, *Science* **293,** 825-828 (2001).
[4] E. Peik *et al.* Limit on the present temporal variation of the fine structure constant, *Phys. Rev. Lett.* **93,** 170801 (2004).
[5] H. S. Margolis *et al.* Hertz-level measurement of the optical clock frequency in a single $^{88}Sr^{+}$ ion, *Science* **306,** 1355-1358 (2004).
[6] G. Wilpers *et al.* Optical clock with ultracold neutral atoms, *Phys. Rev. Lett.* **89,** 230801 (2002).
[7] C. W. Oates, E. A. Curtis, and L. Hollberg, Improved short-term stability of optical frequency standards: approaching 1Hz in 1s with the Ca standard at 657 nm, *Opt. Lett.* **25,** 1603-1605 (2000).
[8] F. Pereira Dos Santos *et al.* Controlling the cold collision shift in high precision atomic interferometry, *Phys. Rev. Lett.* **89,** 233004 (2002).
[9] H. G. Dehmelt, Mono-ion oscillator as potential ultimate laser frequency standard, *IEEE Trans. Instrum. Meas.* **31,** 83-87 (1982).
[10] R. J. Rafac *et al.* Sub-dekahertz Ultraviolet Spectroscopy of $^{199}Hg^{+}$, *Phys. Rev. Lett.* **85,** 2462-2465 (2000).
[11] F. Ruschewitz *et al.* Sub-kilohertz optical spectroscopy with a time domain atom interferometer, *Phys. Rev. Lett.* **80,** 3173-3176 (1998).
[12] H. Katori, Spectroscopy of strontium atoms in the Lamb-Dicke confinement, in *Proceedings of the 6th Symposium on Frequency Standards and Metrology.* (ed. Gill, P.) 323-330 (World Scientific, Singapore, 2002).
[13] H. Katori, M. Takamoto, V. G. Pal'chikov, and V. D. Ovsiannikov, Ultrastable optical clock with neutral atoms in an engineered light shift trap, *Phys. Rev. Lett.* **91,** 173005 (2003).
[14] H. Katori, T. Ido, and M. Kuwata-Gonokami, Optimal design of dipole potentials for efficient loading of Sr atoms. *J. Phys. Soc. Jpn.* **68,** 2479-2482 (1999).
[15] M. Takamoto, and H. Katori, Spectroscopy of the $^1S_0 - {}^3P_0$ clock transition of ^{87}Sr in an optical lattice, *Phys. Rev. Lett.* **91,** 223001 (2003).
[16] D. W. Allan, Statistics of atomic frequency standards, *Proc. IEEE.* **54,** 221-230 (1966).
[17] B. C. Young, F. C. Cruz, W. M. Itano, and J. C. Bergquist, Visible lasers with subhertz linewidths, *Phys. Rev. Lett.* **82,** 3799-3802 (1999).
[18] F. -L. Hong *et al.* Frequency comparison of $^{127}I_2$-stabilized Nd:YAG lasers, *IEEE Trans. Instrum. Meas.* **48,** 532-536 (1999).
[19] Bureau International des Poids et Mesures (BIPM), Circular T, No. 200 http://www1.bipm.org/en/scientific/tai/time_ftp.html (August 2004).
[20] I. Courtillot *et al.* Clock transition for a future optical frequency standard with trapped atoms, *Phys. Rev. A* **68,** 030501 (2003).
[21] L.-S. Ma, P. Jungner, J. Ye and J. L. Hall, Delivering the same optical frequency at two places: accurate cancellation of phase noise introduced by an optical fiber or other time-varying path, *Opt. Lett.* **19,** 1777-1779 (1994).
[22] S. Gupta *et al.* Radio-frequency spectroscopy of ultracold fermions, *Science* **300,** 1723-1726 (2003).
[23] A. Rauschenbeutel, H. Schadwinkel, V. Gomer and D. Meschede, Standing light fields for cold atoms with intrinsically stable and variable time phases, *Opt. Commun.* **148,** 45-48 (1998).
[24] C. Degenhardt, H. Stoehr, U. Sterr, F. Riehle and C. Lisdat, Wavelength-dependent ac Stark shift of the $^1S_0$ -$^3P_1$ transition at 657nm in Ca, *Phys. Rev. A* **70,** 023414 (2004).
[25] S. G. Porsev, A. Derevianko and E. N. Fortson, Possibility of an optical clock using the $6^1S_0 \rightarrow 6^3P_0^o$ transition in $^{171,173}Yb$ atoms held in an optical lattice, *Phys. Rev. A* **69,** 021403 (2004).
[26] E. J. Angstmann, V. A. Dzuba, and V. V. Flambaum, Relativistic effects in two valence-electron atoms and ions and the search for variation of the fine-structure constant, *Phys. Rev. A* **70,** 014102 (2004).
[27] M. Takamoto, F.-L. Hong, R. Higashi, and H. Katori, An optical lattice clock, *Nature* **435**, 321-324 (2005).

STABILITY AND ACCURACY OF $^{171}Yb^{+}$ SINGLE-ION OPTICAL FREQUENCY STANDARDS

T. Schneider, Chr. Tamm, E. Peik
Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany
e-mail: ekkehard.peik@ptb.de

ABSTRACT

Two single-ion optical frequency standards are compared in order to investigate their stability and systematic frequency shifts. Without externally applied perturbations, a mean frequency difference of 0.26(42) Hz is observed, corresponding to a relative agreement of $3.8(6.1)\times10^{-16}$.

1. INTRODUCTION

It has long been pointed out that frequency standards based on narrow optical transitions in laser-cooled ions stored in radio-frequency traps have the potential to reach instabilities and systematic uncertainties beyond the capabilities of the best microwave standards. Recently, an absolute frequency measurement on $^{88}Sr^{+}$ has reached an uncertainty not much larger than those of a cesium fountain clock [1] and shifts of optical frequencies have been investigated at the level of a few hertz [2,3]. Studies of the systematic frequency shifts of an optical frequency standard by measurements of its frequency relative to established microwave standards are inherently limited by the instabilities and uncertainties of the latter. These limitations can be overcome by comparing the optical frequency standard with a second one as we demonstrate in this paper by comparing two single-ion frequency standards based on a reference transition in $^{171}Yb^{+}$ [4].

2. EXPERIMENT AND RESULTS

We compare the narrow absorption resonances from the $6s\ ^2S_{1/2} \rightarrow 5d\ ^2D_{3/2}$ electric quadrupole transition at 688 THz in two laser-cooled $^{171}Yb^{+}$ ions that are stored in separate Paul traps. Independent frequency servo systems are used to lock two probe frequencies that are derived from a frequency stabilized diode laser to the resonance line centers of the ions [5]. Fourier-limited resonance linewidths in the range of 10 Hz to 30 Hz are realized. The difference between the probe frequencies is recorded once per second. For times that are longer than the servo time constant of about 10 s, this allows to investigate the stability of the frequency standards and to observe line shifts that can be induced by varying operation parameters. A relative instability of the optical frequency difference of $\sigma_y(1000\ s)=5*10^{-16}$ has been obtained with a reference transition linewidth of 30 Hz, so that shifts in the sub-hertz range can be resolved within a few minutes.

The digital servo systems that stabilize the probe frequencies use second-order integrating algorithms in order to minimize servo errors due to clock laser frequency drift. For averaging times $\tau>100$ s, $\sigma_y(\tau)$ follows the expected $\tau^{-1/2}$ behavior. The observed instability is independent of the temporal overlap of the probe pulses applied to the two ions and is in good agreement with numerical Monte Carlo simulations of the servo action for the case that the fluctuations of the atomic resonance signals are dominated by quantum projection noise.

The most important systematic shifts in the $^{171}Yb^{+}$ frequency standard are caused by the interaction of the ion with electric fields: the quadrupole shift due to the interaction with a field gradient and the quadratic Stark shift [6]. Field gradients can be produced by a static part of the quadrupole trap potential and also by patch charges in the vicinity of the trap. The quadratic Stark shift becomes important if the ion is not localized at the saddle point of the rf quadrupole potential of the Paul trap.

By applying a static voltage to the ring electrode of the trap, a field gradient of known magnitude and orientation was applied and the quadrupole moment of the $D_{3/2}$ state was determined from the measured quadrupole shift with respect to the second trap. The experimental result is $2.08(11)\ ea_0^2$, where e is the elementary charge and a_0 the Bohr radius. It is in good agreement with recent ab initio calculations [3]. From a measurement of the quadratic Stark shift produced by the trap field when one ion was displaced from the trap center, we also determined the differences of the scalar and tensor polarizabilities of the two levels connected by the reference transition [7].

Without externally applied perturbations, a mean frequency difference of 0.26(42) Hz is observed between the two trapped ions. The average was taken over measurements performed for different orientations of the magnetic field and the

uncertainty is the average statistical uncertainty of the individual measurements. Within the statistical uncertainty the data yields no evidence for frequency shifts that depend on the orientation of the magnetic field, like the quadrupole shift or the tensorial quadratic Stark shift.

This result corresponds to a relative agreement of $3.8(6.1)\times10^{-16}$, comparable to the best agreement found in a comparison of cesium fountain clocks [8]. Lowering the statistical uncertainties, it seems therefore possible to test the performance of optical frequency standards at an accuracy level exceeding that of the best cesium clocks in the near future.

REFERENCES

[1] H. S. Margolis et al., Science **306**, 1355 (2004).

[2] G. P. Barwood, H. S. Margolis, G. Huang, P. Gill, H. A. Klein, Phys. Rev. Lett. **93**, 133001 (2004).

[3] W. H. Oskay, W. M. Itano, J. C. Bergquist, Phys. Rev. Lett. (in print).

[4] E. Peik et al., Phys. Rev. Lett. **93**, 170801 (2004).

[5] Chr. Tamm, T. Schneider, and E. Peik, in: *Laser Spectroscopy XVI*, Eds.: P. Hannaford, A. Sidorov, H. Bachor, and K. Baldwin (World Scientific, Singapore, 2004).
(see also: arXiv physics/0402120)

[6] W. M. Itano, J. Res. Natl. Inst. Stds. Techn. **105**, 829 (2000).

[7] T. Schneider, E. Peik, Chr. Tamm, to be published.

[8] S. Bize et al., C. R. Physique **5**, 829 (2004)

AN OPTICAL CLOCK WITH NEUTRAL ATOMS CONFINED IN A SHALLOW TRAP

Pierre Lemonde[†] and Peter Wolf[†*]
[†]SYRTE, Observatoire de Paris
61, Avenue de l'observatoire, 75014, Paris, France
[*] BIPM, Pavillon de Breteuil, 92312 Sèvres Cedex, France

Abstract

We study the trap depth requirement for the realization of an optical clock using atoms confined in a lattice. We show that site-to-site tunneling leads to a residual sensitivity to the atom dynamics hence requiring large depths ($50-100\,E_r$ for Sr) to avoid any frequency shift or line broadening of the atomic transition at the $10^{-17}-10^{-18}$ level. In a second step we propose to circumvent the problem by the use of an accelerated lattice. Acceleration lifts the degeneracy between adjacents potential wells which strongly inhibits tunneling. We show that using the Earth's gravity, much shallower traps (down to $5\,E_r$ for Sr) can be used for the same accuracy goal.

1 Introduction

We study the effect of the atom motion on the performances of an optical lattice clock [1]. In ref. [1], it is implicitly assumed that each microtrap of the lattice can be treated separately as a quadratic potential in which case the situation is very similar to the case of trapped ions [2]. With an accuracy goal in the $10^{-17}-10^{-18}$ range in mind (corresponding to the mHz level in the optical domain), we shall see later on, that this is correct at very high trap depths only. The natural energy unit for the trap dynamics is the recoil energy associated to the absorption or emission of a photon of the lattice laser, $E_r = \frac{\hbar^2 k_L^2}{2m_a}$ with k_L the wave vector of the lattice laser and m_a the atomic mass. For Sr and for the above accuracy goal the trap depth U_0 corresponding to the independent trap limit is typically $U_0 = 100\,E_r$, which corresponds to a peak laser intensity of $25\,\mathrm{kW/cm^2}$.

For a number of reasons however, one would like to work with traps as shallow as possible. First, the residual shift by the trapping light of the clock transition is smaller and smaller at a decreasing trap depth. The first order perturbation is intrinsically cancelled by tuning to the magic wavelength except for a small eventual tensorial effect. Higher order terms may be much more problematic depending on possible coincidences between two photon resonances and the magic wavelength [1, 3]. The associated shift scales as U_0^2. The shift would then be minimized by a reduction of U_0 and its evaluation would be greatly improved if one can vary this parameter over a broader range. Second, for some of the possible atoms, such as Hg for which the magic wavelength is about 340 nm, two-photon ionization can occur which may limit the achievable resonance width and lead to a frequency shift. Finally, technical aspects like the required laser power at the magic wavelength can be greatly relaxed if one can use shallow traps, especially if the magic wavelength is in a region of the spectrum where no readily available high power laser exists. For Hg for instance, a trap depth of $100\,E_r$ would necessitate a peak intensity of $500\,\mathrm{kW/cm^2}$ at 340 nm.

When considering shallow traps one cannot neglect tunneling of the atoms from one site of the lattice to another. In section 3 we investigate the ultimate performance of the clock taking this effect into account. We show that depending on the initial state of the atoms in the lattice, one faces a broadening and/or a shift of the atomic transition. For Sr, this requires U_0 of about $50-100\,E_r$ to ensure a fractional accuracy better than 10^{-17}. We show in section 4 that a much lower U_0 can be used provided the tunneling process is made non-resonant by lifting the degeneracy between adjacent sites. This can be done by adding a constant acceleration to the lattice [4, 5]. More specifically, we study the case where this acceleration is simply the Earth's gravity. In this configuration, trap depths down to $U_0 \sim 5\,E_r$ can be sufficient for the above accuracy goal.

2 Confined atoms coupled to a light field

In this section we describe the theoretical frame used to investigate the residual effects of the motion of atoms in an external potential. The internal atomic structure is approximated by a two-level system $|g\rangle$ and $|e\rangle$ with energy difference $\hbar\omega_{eg}$. The internal Hamiltonian is:

$$\hat{H}_i = \hbar\omega_{eg}|e\rangle\langle e|. \qquad (1)$$

We introduce the coupling between $|e\rangle$ and $|g\rangle$ by a laser of frequency ω and wavevector k_s propagating along the

x direction:

$$\hat{H}_s = \hbar\Omega\cos(\omega t - k_s\hat{x})|e\rangle\langle g| + h.c., \quad (2)$$

with Ω the Rabi frequency.

In the following we consider potentials induced by lasers tuned at the magic wavelength and/or by gravity. The external potential $\hat{H}_{ext}$ is then identical for both $|g\rangle$ and $|e\rangle$ with eigenstates $|m\rangle$ obeying $\hat{H}_{ext}|m\rangle = \hbar\omega_m|m\rangle$ (Note that $|m\rangle$ can be a continuous variable in which case the discrete sums in the following are replaced by integrals). If we restrict ourselves to experiments much shorter than the lifetime of state $|e\rangle$ (for ^{87}Sr, the lifetime of the lowest $^3P_0$ state is 100 s) spontaneous emission can be neglected and the evolution of the general atomic state

$$|\psi_{at}\rangle = \sum_m a_m^g\, e^{-i\omega_m t}\,|m,g\rangle + a_m^e\, e^{-i(\omega_{eg}+\omega_m)t}\,|m,e\rangle \quad (3)$$

is driven by $i\hbar\frac{\partial}{\partial t}|\psi_{at}\rangle = (\hat{H}_{ext} + \hat{H}_i + \hat{H}_s)|\psi_{at}\rangle$ leading to the following set of coupled equations

$$\begin{aligned} i\,\dot{a}_m^g &= \sum_{m'} \frac{\Omega^*}{2} e^{i\Delta_{m',m}t}\langle m|e^{-ik_s\hat{x}}|m'\rangle a_{m'}^e \quad (4)\\ i\,\dot{a}_m^e &= \sum_{m'} \frac{\Omega}{2} e^{-i\Delta_{m,m'}t}\langle m|e^{ik_s\hat{x}}|m'\rangle a_{m'}^g. \end{aligned}$$

To derive eq. (4) we have made the rotating wave approximation (assuming $\omega - \omega_{eg} << \omega_{eg}$) and defined $\Delta_{m',m} = \omega - \omega_{eg} + \omega_m - \omega_{m'}$.

3 Periodic potential

3.1 Eigenstates and coupling by the probe laser

We now consider the case of atoms trapped in an optical lattice. As is clear from eq. (4), only the motion of the atoms along the probe laser propagation axis plays a role in the problem and we restrict the analysis to 1D. We assume that the lattice is formed by a standing wave leading to the following external hamiltonian:

$$\hat{H}_{ext}^I = \frac{\hbar^2\hat{\kappa}^2}{2m_a} + \frac{U_0}{2}(1 - \cos(2k_l\hat{x})), \quad (5)$$

where k_l is the wave vector of the trap laser. The eigenstates $|n,q\rangle$ and eigenenergies $\hbar\omega_{n,q}$ of $\hat{H}_{ext}^I$ are derived from the Bloch theorem [6]. They are labelled by two quantum numbers: the band index n and the quasi-momentum q. Furthermore they are periodic functions of q with period $2k_l$ and the usual convention is to restrict oneself to the first Brillouin zone $q \in]-k_l, k_l]$.

Following a procedure given in Ref. [5] a numerical solution to this eigenvalue problem can be easily found in the momentum representation. The atomic plane wave with

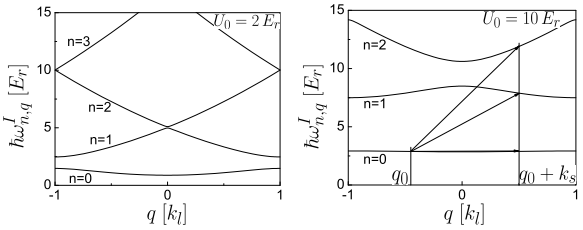


Figure 1: Band structure for two different lattice depth: $U_0 = 2\,E_r$ (left) and $U_0 = 10\,E_r$ (right). Each state $|n, q_0\rangle$ is coupled to all the states $|n', q_0 + k_s\rangle$ by the probe laser.

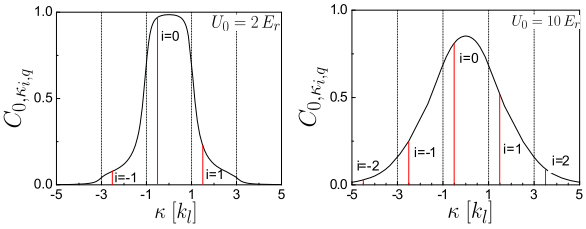


Figure 2: $C_{0,\kappa_{i,q}}$ for two different lattice depth: $U_0 = 2\,E_r$ (left) and $U_0 = 10\,E_r$ (right). The bold vertical lines illustrate the case $q = -k_l/2$. The dotted lines delimit the Brillouin zones. For a state $|n = 0, q = ak_l\rangle$ with $a \in]-1, 1]$ the solid envelope gives the contribution of the plane waves $|\kappa_{i,ak_l} = ak_l + 2ik_l\rangle$.

wave vector κ obeys

$$\hat{H}_{ext}^I|\kappa\rangle = \left(\frac{\hbar^2\kappa^2}{2m_a} + \frac{U_0}{2}\right)|\kappa\rangle - \frac{U_0}{4}(|\kappa + 2k_l\rangle + |\kappa - 2k_l\rangle). \quad (6)$$

For each value of q, the problem then reduces to the diagonalization of a real tridiagonal matrix giving the eigenenergies and eigenvectors as a linear superposition of plane waves:

$$\begin{aligned} \hat{H}_{ext}^I|n,q\rangle &= \hbar\omega_{n,q}^I|n,q\rangle \\ |n,q\rangle &= \sum_{i=-\infty}^{\infty} C_{n,\kappa_{i,q}}|\kappa_{i,q}\rangle, \quad (7) \end{aligned}$$

with $\kappa_{i,q} = q + 2ik_l$. For each value of q one obtains a discrete set of coefficients $C_{n,\kappa_{i,q}}$, which are real and normalized such that $\sum_i C^2_{n,\kappa_{i,q}} = 1$. In figures 1 and 2 are shown $\hbar\omega^I_{n,q}$ and $C_{0,\kappa_{i,q}}$ for various values of U_0. Except when explicitly stated, all numerical values throughout the paper are given for ^{87}Sr at a lattice laser wavelength 813 nm which corresponds to the magic wavelength reported in Ref. [7]. In frequency units E_r then corresponds to 3.58 kHz. In figure 3 is shown the width of the lowest energy band as a function of U_0 in units of E_r and in frequency units.

Substituting $\langle m| \to \langle n,q|$ and $|m'\rangle \to |n',q'\rangle$ in eq. (4),

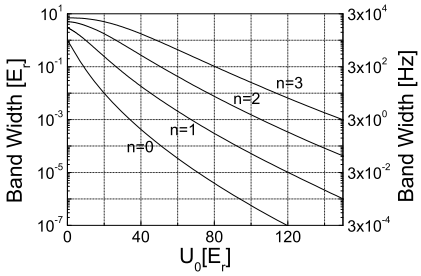


Figure 3: Lowest four bands widths as a function of the lattice depth U_0 in units of E_r (left scale) and in frequency units (right scale).

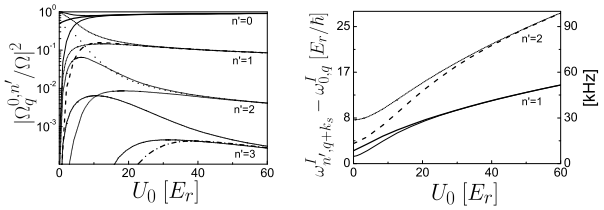


Figure 4: Left: Relative strength of the transitions to different bands ($n = 0 \rightarrow n'$) for an atom prepared in state $|n = 0, q = -k_l\rangle$ (bold lines), $|n = 0, q = -k_l/2\rangle$ and $|n = 0, q = k_l/2\rangle$ (thin lines). Right: detuning of the first two sidebands for an atom prepared in state $|n = 0, q = -k_l\rangle$ (bold lines) and $|n = 0, q = 0\rangle$ (thin lines) in units of E_r (left scale) and in frequency units (right scale).

the action of the probe laser is described by the coupled equations

$$i\,\dot{a}^g_{n,q} = \sum_{n'} \frac{\Omega_q^{n',n^*}}{2} e^{i\Delta_q^{n',n}t} a^e_{n',q+k_s} \qquad (8)$$

$$i\,\dot{a}^e_{n,q+k_s} = \sum_{n'} \frac{\Omega_q^{n,n'}}{2} e^{-i\Delta_q^{n,n'}t} a^g_{n',q} \quad ,$$

with $\Omega_q^{n,n'} = \Omega \sum_i C_{n',\kappa_{i,q}} C_{n,\kappa_{i,q+k_s}}$ and $\Delta_q^{n,n'} = \omega - \omega_{eg} + \omega^I_{n',q} - \omega^I_{n,q+k_s}$. As expected from the structure of the Bloch vectors in (7), the translation in momentum space $e^{ik_s\hat{x}}$ due to the probe laser leads to the coupling of a given state $|n, q\rangle$ to the whole set $|n', q + k_s\rangle$ (see figure 1) with a coupling strength $\Omega_q^{n',n}$ and a shift with respect to the atomic resonance $\omega^I_{n',q+k_s} - \omega^I_{n,q}$. Both quantities depend on n, n' and q and to go further we have to make assumptions on the initial state of the atoms in the lattice.

3.2 Discussion

We first consider the case where the initial state is a pure $|n, q\rangle$ state. The strengths of the resonances $\Omega_q^{n,n'}$ are shown in figure 4 for the case $n = 0$ and various values of q.

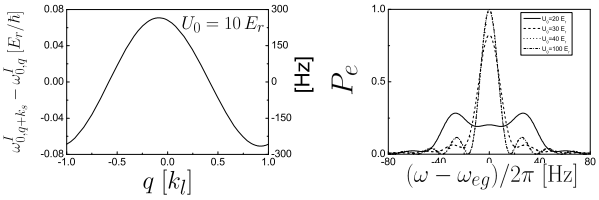


Figure 5: Left: shift of the "carrier" resonance in the first band for a lattice depth $U_0 = 10\,E_r$. Left scale: in units of E_r. Right scale: in frequency units. Right: Expected resonances in the case where the first band is uniformely populated for $\Omega = 10\,\text{Hz}$ and $U_0 = 20\,E_r$, $30\,E_r$, $40\,E_r$, and $100\,E_r$. The duration of the interaction is such that the transition probability at resonance is maximized.

At a growing lattice depth $\Omega_q^{n,n'}$ become independent on q and the strength of all "sidebands" ($n' - n \neq 0$) asymptotically decreases as $U_0^{-|n'-n|/4}$ for the benefit of the "carrier" ($n' = n$). On the other hand the frequency separation of the resonances rapidly increases with U_0 (Fig. 4). For U_0 as low as $5\,E_r$, this separation is of the order of 10 kHz. For narrow resonances (which are required for an accurate clock) they can be treated separately and the effect of the sidebands on the carrier is negligible. If for example one reaches a carrier width of 10 Hz, the sideband pulling is of the order of 10^{-5} Hz. On the other hand the "carrier" frequency is shifted from the atomic frequency by $\omega^I_{n,q+k_s} - \omega^I_{n,q}$ due to the band structure. This shift is of the order of the width of the nth band (Fig. 5 and 3). It can be seen as a residual Doppler and recoil effect for atoms trapped in a lattice and is a consequence of the complete delocalisation of the eigenstates of the system over the lattice. The "carrier" shift is plotted in figure 5 for the case $n = 0$ and $U_0 = 10\,E_r$. For this shift to be as small as 5 mHz over the whole lowest band, which corresponds in fractional units to 10^{-17} for Sr atoms probed on the $^1S_0 - ^3P_0$ transition, the lattice depth should be about $100\,E_r$ (Fig. 3).

Another extreme situation is the case where one band is uniformly populated. In this case the "carrier" shift averaged over q cancels and one can hope to operate the clock at a much lower U_0 than in the previous case. The problem is then the ultimate linewidth that can be achieved in the system which is of the order of the width of the band and is reminiscent of Doppler broadening. This is illustrated in figure 5 for which we have computed the expected "carrier" resonances in the case where the lowest band is uniformly populated, by numerically solving equations (4). This was done for a Rabi frequency $\Omega = 10\,\text{Hz}$ and an interaction duration which is adjusted for each trapping depth so as to maximize the transition probability at zero detuning. All resonances plotted in figure 5 are not shifted to within the numerical accuracy (less than 10^{-5} Hz). However, at decreasing U_0 the contrast of the resonance starts to drop for $U_0 < 40\,E_r$ and the

resonance broadens progressively, becoming unusable for precise spectroscopy when the width of the energy band reaches the Rabi frequency. To get more physical insight into this phenomenon, let's consider the particular example of this uniform band population where one well of the lattice is initially populated. This corresponds to a given relative phase of the Bloch states such that the interference of the Bloch vectors is destructive everywhere except in one well of the lattice. The time scale for the evolution of this relative phase is the inverse of the width of the populated energy band which then corresponds to the tunneling time towards delocalization. The broadening and loss of contrast shown in figure 5 can be seen as the Doppler effect associated with this tunneling motion.

In the experimental setup described in [7] about 90 % of the atoms are in the lowest band and can be selected by an adequate sequence of laser pulses. The residual population of excited bands can then be made negligible ($< 10^{-3}$). On the other hand, knowing and controlling with accuracy the population of the various $|q\rangle$ states in the ground band is a difficult task. The actual initial distribution of atomic states will lie somewhere between the two extremes studied above: a pure state and a uniform superposition of all states in the bottom band. If we assume that the population of the $|q\rangle$ states can be controlled so that the frequency shift averages to within one tenth (resp. 1%) of the band width, then a fractional accuracy goal of 10^{-17} implies $U_0 = 70\,E_r$ (resp. $50\,E_r$) or more. Therefore, operation of an optical lattice clock requires relatively deep wells and correspondingly high laser power, which, in turn, is likely to lead to other difficulties as described in the introduction. Fortunately, the requirement of deep wells can be significantly relaxed by adding a constant acceleration to the lattice, as described in the next section.

4 Periodic potential in an accelerated frame

4.1 Wannier-Stark states and coupling by the probe laser

The shift and broadening encountered in the previous section are both due to site-to-site tunneling and to the corresponding complete delocalization of the eigenstates of the lattice. As is well-known from solid-state physics, one way to localize the atoms is to add a linear component to the Hamiltonian [4]: adjacent wells are then shifted in energy, which strongly inhibits tunneling. In this section we study the case where the lattice and probe laser are oriented vertically so that gravity plays the role of this linear component. The external hamiltonian is then:

$$\hat{H}^{II}_{ext} = \frac{\hbar^2\hat{\kappa}^2}{2m_a} + \frac{U_0}{2}(1-\cos(2k_l\hat{x})) + m_a g\hat{x}, \qquad (9)$$

with g the acceleration of the Earth's gravity. This hamiltonian supports no true bound states, as an atom initially confined in one well of the lattice will end up in the continuum due to tunneling under the influence of gravity. This effect is known as Landau-Zener tunneling [4, 5]. Its timescale however increases exponentially with the depth of the lattice and for the cases considered here is orders of magnitude longer than the duration of the experiment. In the case of Sr and for U_0 as low as $5\,E_r$, the lifetime of the ground state of each well is about 10^{10} s! The coupling due to gravity between the ground and excited bands can therefore be neglected here. In the frame of this approximation the problem of finding the "eigenstates" of $\hat{H}^{II}_{ext}$ reduces to its diagonalization in a sub-space restricted to the ground band [8, 9] (we drop the band index in the following to keep notations as simple as possible). We are looking for solutions to the eigenvalue equation, of the form:

$$\begin{aligned} \hat{H}^{II}_{ext}|W_m\rangle &= \hbar\omega^{II}_m|W_m\rangle \qquad (10)\\ |W_m\rangle &= \int_{-k_l}^{k_l} dq\, b_m(q)|q\rangle\,. \end{aligned}$$

In eq. (10) the $|q\rangle$ are the Bloch eigenstates of $\hat{H}^{I}_{ext}$ (c.f. section 3) for the bottom energy band ($n = 0$), m is a new quantum number, and the $b_m(q)$ are periodic: $b_m(q + 2ik_l) = b_m(q)$. After some algebra, eq. (10) reduce to the differential equation

$$\hbar(\omega^I_q - \omega^{II}_m)b_m(q) + im_a g\partial_q b_m(q) = 0 \qquad (11)$$

where ω^I_q is the eigenfrequency of the Bloch state $|n = 0, q\rangle$ of section 3. Note that equations (10) and (11) only hold in the limit where Landau-Zener tunneling between energy bands is negligible [10, 5]. Here the periodicity of $b_m(q)$ and a normalization condition lead to a simple solution of the form

$$\begin{aligned} \omega^{II}_m &= \omega^{II}_0 + m\Delta_g \qquad (12)\\ b_m(q) &= \frac{1}{\sqrt{2k_l}} e^{-\frac{i\hbar}{m_a g}(q\omega^{II}_m - \gamma_q)} \end{aligned}$$

with $\omega^{II}_0 = \frac{1}{2k_l}\int_{-k_l}^{k_l} dq\,\omega^I_q$, $\hbar\Delta_g = m_a g\lambda_l/2$, $\omega^I_q = \partial_q\gamma_q$ and $\gamma_0 = 0$. The $|W_m\rangle$ states are usually called Wannier-Stark states and their wave functions are plotted in figure 6 for various trap depths. In the position representation $|W_m\rangle$ exhibits a main peak in the mth well of the lattice and small revivals in adjacent wells. These revivals decrease exponentially at increasing lattice depth. At $U_0 = 10\,E_r$ the first revival is already 10^2 times smaller than the main peak. Conversely, in the momentum representation, the distribution gets broader with increasing U_0. The phase shift between b_m and b_{m-1} in (12), $b_m(q) = e^{-i\pi q/k_l}b_{m-1}(q)$, corresponds to a translational symmetry of the Wannier-Stark states in the position representation $\langle x + \lambda_l/2|W_m\rangle = \langle x|W_{m-1}\rangle$.

Substituting $\langle m| \to \langle W_m|$ and $|m'\rangle \to |W_{m'}\rangle$ in eq. (4) shows that the effect of the probe laser is to couple the

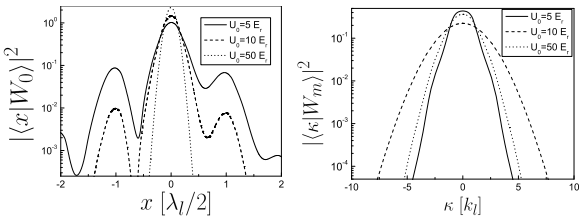


Figure 6: Wannier-Stark states in position (left) and momentum (right) representation for $U_0 = 5\,E_r$, $U_0 = 10\,E_r$ and $U_0 = 50\,E_r$. Numerically we first compute the momentum representation $\langle\kappa|W_m\rangle = b_m(\kappa)C_{0,\kappa}$ and then obtain the position representation by Fourier tranformation.

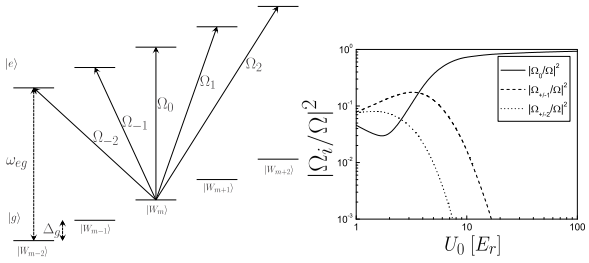


Figure 7: Left: Wannier-Stark ladder of states and coupling between states by the probe laser. Right: Relative strength of the carrier $|\Omega_0/\Omega|^2$ and of the first four sidebands $|\Omega_{\pm1}/\Omega|^2$ and $|\Omega_{\pm2}/\Omega|^2$ as a function of the lattice depth U_0.

Wannier-Stark states to their neighbours with the coupling strengths

$$\langle W_m|e^{ik_s\hat{x}}|W_{m'}\rangle = \int_{-\infty}^{\infty} d\kappa\, b_m^*(\kappa + k_s)b_{m'}(\kappa)C_{0,\kappa}C_{0,\kappa+k_s}, \tag{13}$$

obtained from direct substitution of (10) and (7) [1].

Using the translational symmetry of the Wannier-Stark states it is easy to show that

$$\langle W_m|e^{ik_s\hat{x}}|W_{m'}\rangle = e^{i\pi m k_s/k_l}\langle W_0|e^{ik_s\hat{x}}|W_{m'-m}\rangle. \tag{14}$$

From that property, equation (13), and using $b_m(\kappa) = b_m^*(-\kappa)$ (note that $\gamma_q = \gamma_{-q}$) one can then show that

$$\langle W_m|e^{ik_s\hat{x}}|W_{m+j}\rangle = e^{i\pi j k_s/k_l}\langle W_m|e^{ik_s\hat{x}}|W_{m-j}\rangle, \tag{15}$$

which is a useful result when studying the symmetry of coupling to neighbouring states (see next section).

The differential equations (4), governing the evolution of the different states under coupling to the probe laser are then

$$\begin{aligned} i\,\dot{a}_m^g &= \sum_{m'} \frac{\Omega_{m-m'}^*}{2} e^{-i\pi m' \frac{k_s}{k_l}} e^{i\Delta_{m-m'}t} a_{m'}^e \\ i\,\dot{a}_m^e &= \sum_{m'} \frac{\Omega_{m'-m}}{2} e^{i\pi m \frac{k_s}{k_l}} e^{-i\Delta_{m'-m}t} a_{m'}^g, \end{aligned} \tag{16}$$

in which we have used (14) and defined $\Omega_m = \Omega\langle W_0|e^{ik_s\hat{x}}|W_m\rangle$ and $\Delta_m = \omega - \omega_{eg} + m\Delta_g$.

4.2 Discussion

We now study the case where the initial state of the atom is a pure Wannier-Stark state. According to eq. (16) excitation by the laser will lead to a set of resonances separated by Δ_g (see Fig. 7). In the case of Sr, $\Delta_g/2\pi = 866\,\text{Hz}$ and for the narrow resonances required for high performance clock operation, they are easily resolved. The resonances obtained by first numerically integrating (13) and then numerically solving (16) are plotted

[1] For similar reasons as in section 3 one can neglect the coupling to excited bands in the system for narrow enough resonances.

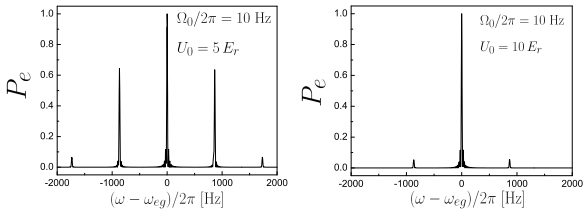


Figure 8: Computed resonances when the initial state is a pure Wannier-Stark state. Left: $U_0 = 5\,E_r$, right: $U_0 = 10\,E_r$. Both resonances are plotted for an effective Rabi frequency of the carrier $\frac{\Omega_0}{2\pi} = 10\,\text{Hz}$ and for an interaction time of 50 ms.

in figure 8 for the cases $U_0 = 5\,E_r$ and $U_0 = 10\,E_r$. They exhibit remarkable properties. First the "carrier" (which corresponds to the transition $|W_m\rangle \to |W_m\rangle$) has a frequency which exactly matches the atomic frequency ω_{eg}. It doesn't suffer from any broadening or contrast limitation which would be due to the atomic dynamics. Second, the sidebands ($|W_m\rangle \to |W_{m\pm i}\rangle$) have a coupling strength which very rapidly decreases as U_0 increases (see fig. 7). In addition they are fully symmetric with respect to the carrier which results from eq. (15), and hence lead to no line pulling of the carrier. This absence of shift and broadening remains true even for very shallow traps down to a depth of a few E_r, the ultimate limitation being the lifetime of the Wannier-Stark states. This situation is in striking contrast with the results of section 3 in the absence of gravity.

The system is more complex if the initial state of the atom is a coherent superposition of neighbouring wells. In this case off-resonant excitation of the sidebands will interfere with the carrier excitation. This interference leads to a modification of the carrier transition probability which is of the order of Ω_1/Δ_g (for the first, and most significant, sideband). For an interaction close to a π pulse, an order of magnitude of the corresponding carrier pulling is then $\Omega_1\Omega_0/\Delta_g$ which can be significant. For $U_0 = 10\,E_r$ and $\Omega_0/2\pi = 10\,\text{Hz}$ the shift is about $2 \times 10^{-2}\,\text{Hz}$, *i.e.* several

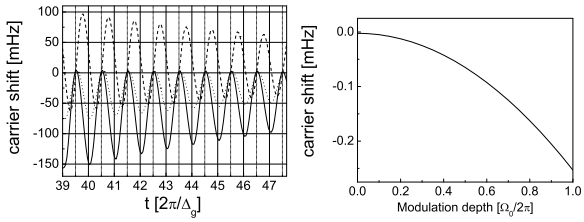


Figure 9: Left: Frequency shift of the carrier as a function of the interaction duration in the case where the initial state of the atom is a coherent superposition of neighbouring Wannier-Stark states. Solid line: $a_n^g(t=0) = a_{n+1}^g(t=0)$ for all n. Dashed line: $a_n^g(t=0) = a_{n+1}^g(t=0)e^{i\pi/2}$ for all n. Dotted line: $a_{-1}^g(t=0) = a_0^g(t=0)$ and $a_n^g(t=0) = 0$ for $n \neq -1, 0$. The shift is defined as the equilibrium point of a frequency servo loop using a square frequency modulation of optimal depth and computed for $U_0 = 5\,E_r$ and a carrier Rabi frequency $\Omega_0/2\pi = 10\,\mathrm{Hz}$. The interaction duration corresponding to a π pulse is $t\Delta_g/2\pi = 43.3$. Right: Frequency shift of the carrier as a function of the square modulation depth. The calculation has been performed for $U_0 = 5\,E_r$, $\Omega_0/2\pi = 10\,\mathrm{Hz}$ and an interaction time of $t\Delta_g/2\pi = 43.5$. The initial atomic state is the one corresponding to the dotted line on the left graph.

times 10^{-17} of the clock transition frequency. This shift is a priori all the more problematic as the initial atomic state is difficult to know and control accurately.

We have numerically solved eq. (16) for various initial atomic states, lattice depths and interaction parameters. The results are illustrated in figure 9 for the case $U_0 = 5\,E_r$. A clear signature of the effect can be identified from its dependence on the interaction duration: the frequency shift oscillates with a frequency $\Delta_g/2\pi$ resulting from the Δ_g term in $\Delta_{m-m'}$ in (16). This provides a powerful method for investigating site-to-site coherences in the lattice. More interestingly from a clock point of view, the shift becomes negligible for all interaction durations t such that $t = (n+1/2)2\pi/\Delta_g$. Since Δ_g is extremely well known (potentially at the 10^{-9} level) this condition can be accurately met. Note that choosing such a value of the interaction duration does not significantly affect the contrast, as the two relevant timescales have different orders of magnitude in the narrow resonance limit ($\Omega^{-1} >> \Delta_g^{-1}$), and therefore a range of values of t such that $t = (n+1/2)2\pi/\Delta_g$ correspond to almost optimal contrast (e.g. all such values of t in figure 9). A more detailed study shows that the level of cancellation depends on the depth of the modulation used to determine the frequency shift (see caption of Fig. 9) which results from a slight distortion of the carrier resonance. This effect is shown in figure 9, which clearly indicates that the shift can be controlled to below 1 mHz even for a very shallow lattice depth down to $U_0 = 5\,E_r$.

5 Conclusion

We studied the trap depth requirement for the operation of an optical lattice clock with a projected fractional accuracy in the $10^{-17} - 10^{-18}$ range. We have shown that using a purely periodic potential necessitates a lattice of depth $50 - 100\,E_r$ limited by tunneling between adjacent sites of the lattice. A possible way to vastly reduce this depth and the associated frequency shift or required laser power is to use gravity to lift the degeneracy between the potential wells which strongly inhibits tunneling. Trap depths down to $5 - 10\,E_r$ are then sufficient to cancel the effects of the atoms dynamics at the desired level.

Acknowledgements

We thank Sébastien Bize and André Clairon for fruitful and stimulating discussions. SYRTE is Unité Associée au CNRS (UMR 8630) and acknowledges support from Laboratoire National de Métrologie et d'Essai (LNE).

References

[1] H. Katori, M. Takamoto, V. G. Pal'chikov, and V. D. Ovsiannikov, Phys. Rev. Lett. **91**, 173005 (2003).

[2] D. Leibfried, R. Blatt, C. Monroe, and D. Wineland, Rev. Mod. Phys. **75**, 281 (2003).

[3] S. G. Porsev and A. Derevianko, Phys. Rev. A **69**, 042506 (2004).

[4] G. Nenciu, Rev. Mod. Phys. **63**, 91 (1991).

[5] M. Glück, A. Kolovsky, H. Korsch, and N. Moiseyev, Euro. Phys. J. D **4**, 239 (1998).

[6] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders, Philadelphia, 1976).

[7] M. Takamoto and H. Katori, Phys. Rev. Lett. **91**, 223001 (2003).

[8] G. H. Wannier, Phys. Rev. **117**, 432 (1960).

[9] J. Callaway, Phys. Rev. **130**, 549 (1963).

[10] J. Avron, Ann. Phys. **143**, 33 (1982).

Session 4A / Cold atom clocks

Frequency comparison between two atomic fountain clocks at the 10^{-16} level

C. Vian, P. Rosenbusch, H. Marion, S. Bize, L. Cacciapuoti,* M. Abgrall,† D. Chambon, G. Santarelli, Ph. Laurent, and A. Clairon
SYRTE, Observatoire de Paris, 61 Avenue de l'Observatoire, 75014 Paris, France.

A. Luiten and M. Tobar
The University of Western Australia, School of Physics, 35 Stirling Hwy, Crawley, Western Australia.

C. Salomon
Laboratoire Kastler Brossel, ENS, 24 rue Lhomond, 75005 Paris, France.

Keywords — atomic fountain, frequency comparison, primary standard, cold collisions.

Abstract — At SYRTE, we have performed a direct local comparison between two atomic fountains, FO1 and FO2 using as the interrogation oscillator a cryogenic sapphire resonator oscillator weakly phase-locked to a hydrogen maser . The best short term stabilities measured against this interrogation oscillator are 2.9 and 1.6×10^{-14} at 1 s, respectively. When accounting for the real time measurement of the cold collision shift as well as correction for all the other systematic shifts, we measure combined stability of 5.0×10^{-14}, going down to 2.2×10^{-16} at 50000 s. The measured systematic frequency difference is $(4 \pm 2.2) \times 10^{-16}$. It is fully consistent with the clock accuracies of 7.5×10^{-16} for FO1 and 6.5×10^{-16} for FO2.

INTRODUCTION

During the last years, several atomic fountains have been built around the world [1–4] and a few distant comparisons have been performed using satellite links such as the Global Positioning System (GPS) and Two Way Satellite Time and Frequency Transfer (TW-STFT) [5], [6]. These remote measurements allow comparisons at the 10^{-15} level over long averaging time ($\sim$ week). Instead, comparisons between atomic fountains at the same location are free from the additional link noise and unknown biases. Such comparisons are only limited by the fountain clock stability allowing for frequency resolution one to two orders of magnitude lower for the same measurement time. In this paper, we describe such a local comparison between SYRTE fountains, FO1 and FO2 [7] using a cryogenic oscillator (CSO) [8, 9] to generate an ultra-stable interrogation signal. After describing the experimental setup, we discuss the clock accuracies. We finally report the direct frequency comparison between the two fountains. For the first time, a frequency resolution in the low 10^{-16} between two primary standards is achieved.

EXPERIMENTAL SETUP

Fountain set-ups

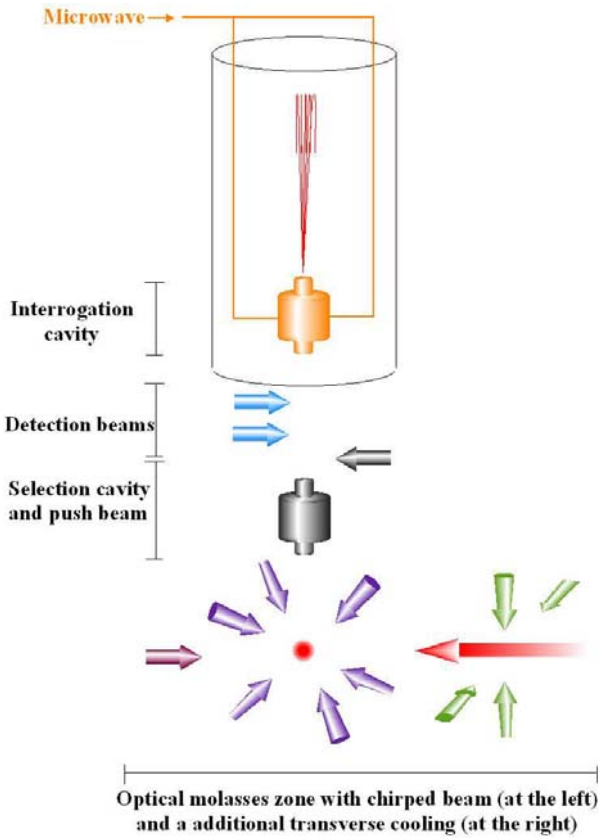


FIG. 1: Schematic of an atomic fountain : first, the atomic beam is slowed down by a combination of a transverse cooling and a chirped laser beam. An optical molasses is loaded at the intersection of six laser beams. The cloud is launched upward through a selection cavity and a push beam. It then reaches the interrogation cavity, where the atoms interact twice with the microwave. The transition probability is measured by a time of flight method, using two resonant laser beams.

FO1 and FO2 fountain clocks, operate with lin $\perp$ lin optical molasses loaded from a laser slowed atomic beam (FIG. 1). With this setup, 3×10^8 atoms can be loaded in $\sim$ 400 ms for FO1. FO2 has an additional

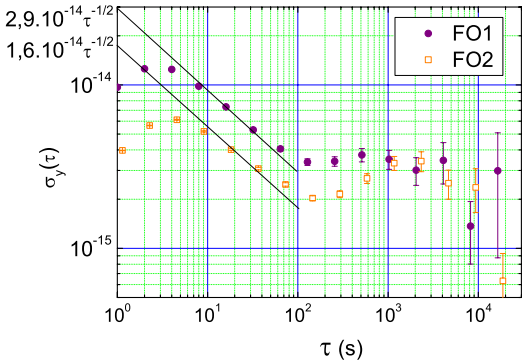


FIG. 2: Fractional frequency instability of FO1 and FO2, measured against the CSO as a function of the averaging time, τ. These measurements are realized with a high number of detected atoms ($\sim 5 \times 10^6$)

transverse cooling of the atomic beam which increases the loading rate to 10^9 atoms for the same duration. Atoms are launched upwards at 4 m.s^{-1} using the moving molasses technique, and cooled down to $\sim$ 1 μK. For regular clock operation, atoms are selected in the $|F = 3, m_F = 0\rangle$ lower clock state. The TE_{011} mode of a copper cylindrical cavity is used to probe the hyperfine transition in a Ramsey interrogation scheme. This cavity can be fed through two coupling irises oppositely located on the cavity diameter. The detection zone is made of two standing waves of resonant light, allowing for the measurement of the transition probability with a signal-to-noise ratio of $\sim$ 5 000.

Frequency stability

Figure 2 shows the fractional frequency instability of the two clocks measured against the interrogation oscillator, composed of a cryogenic sapphire oscillator weakly phase-locked to a reference hydrogen maser. For FO1 and FO2, the short term frequency instabilities are 2.9 and 1.6×10^{-14} at 1 s, respectively. To date these are the best reported short term frequency stability for primary standards. For averaging time above 100 s, the frequency instability is dominated by the frequency fluctuations of the interrogation oscillator.

Reaching these short term frequency stabilities is possible thanks to the low phase noise of the interrogation signal [10] and the high signal-to-noise ratio (5 000 for $\sim 10^7$ detected atoms). Achieving such stabilities is a crucial step for studying and evaluating systematic shifts at a level of 10^{-16}.

ACCURACY

In our continuing effort to improve primary standards, recent work focused on controlling the most limiting systematic shifts in the accuracy budget. In particular, the cold collision shift is known to be a key issue for Cs fountains. Furthermore, the stability of the magnetic field in the interrogation region has been significantly improved. These efforts allow to have an overall uncertainty of $\pm 7.5 \times 10^{-16}$ for FO1 and $\pm 6.5 \times 10^{-16}$ for FO2 (see TABLE I).

TABLE I: Systematic frequency shifts of FO1 and FO2. All values are relative to 10^{16}.

	FO1	FO2
Quadratic Zeeman effect	1199.7 ± 4.5	1927.3 ± 0.3
Blackbody radiation	−162.8 ± 2.5	−168.2 ± 2.5
Collisions and cavity pulling	−197.9 ± 2.4	−357.5 ± 2.0
Microwave spectral purity & leakage	< 3.3	< 4.3
First order Doppler effect	< 3.0	< 3.0
Ramsey & Rabi pulling	< 1.0	< 1.0
Microwave recoil	< 1.4	< 1.4
Second order Doppler effect	< 0.1	< 0.1
Background collisions	< 1.0	< 1.0
Total uncertainty	±7.5	±6.5

The following paragraphs describe the means of controlling these systematic frequency shifts.

Cold collision frequency shift

Our laboratory has proposed and demonstrated a new method [11], based on rapid adiabatic population transfer (AP), in order to prepare atomic samples with a well determined and reproducible density ratio. The initial selection of the $|\mathrm{F} = 3, m_\mathrm{F} = 0\rangle$ clock state is performed by sweeping both the selection microwave frequency across the atomic resonance and the amplitude. Two selection configurations are used. 1- A complete adiabatic population transfer from $|F = 4, m_F = 0\rangle$ to $|F = 3, m_F = 0\rangle$. 2- An interrupted adiabatic population transfer, where the microwave interaction is interrupted when the detuning crosses zero. The two different selection configurations lead to a selection efficiency of 100% and 50% of the $|4, 0\rangle$ atoms respectively, with an accuracy of $\sim 10^{-3}$ regardless of experimental imperfections, such as the field distribution in the selection cavity and microwave power fluctuations. The selection of either 100% or 50% of the atoms is performed without changing neither the atomic velocity distribution nor the position distribution. Consequently the atomic density is changed in a well controlled manner by a factor 1/2, allowing for precise extrapolation to zero density.

Practically, a combination of the adiabatic passage and differential measurement method leads to a real time measurement of the cold collision frequency shift. The differential method is based on alternating between a high density configuration (100%, full AP) and low density configuration (50%, interrupted AP) each

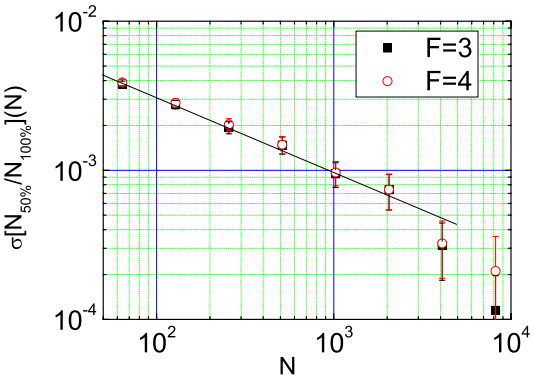


FIG. 3: Stabilities of the density ratios (in relative values) with respect to the number of cycles (N), in FO1 fountain.

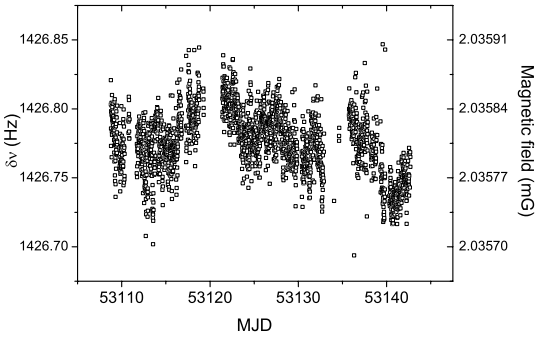


FIG. 4: Measurement of the bias magnetic field in the interrogation region of the FO2 fountain. Magnetic field fluctuations are smaller than ±20 pT peak-to-peak over a one month duration.

60 s, this duration being chosen to optimize the overall stability of the measurement.

During this process the detected atom number in F = 3 and F = 4 for both configurations are stored and analyzed, allowing for a real time monitoring of the adiabatic population transfer. We monitor the ratio of the detected atom number between low and high configurations for a given state.

Figure 3 displays the stability of these ratios as a function of the number of fountain cycles. Both ratios average down as white noise below 10^{-3}. For FO1, the measured ratios are :

$$\frac{N_{F=4}^{50\%}}{N_{F=4}^{100\%}} = \frac{1}{2}(1 + (2 \pm 0.2) \times 10^{-3}) \quad (1)$$

$$\frac{N_{F=3}^{50\%}}{N_{F=3}^{100\%}} = \frac{1}{2}(1 + (6 \pm 0.2) \times 10^{-3}) \quad (2)$$

The deviation from 1/2 of this ratio is accounted for as one of the uncertainty. For atoms in F = 3, the deviation from one half has shown slightly larger due to spurious population in $|\mathrm{F} = 3, m_\mathrm{F} \neq 0\rangle$ states due to imperfections of the selection.

As we have shown [12], the collision shift due to these $|F = 3, m_\mathrm{F} \neq 0\rangle$ atoms is at most 1/3 of the shift caused by $|F = 3, m_\mathrm{F} = 0\rangle$ and $|F = 4, m_\mathrm{F} = 0\rangle$ for the same atomic density. The total population of these $|F = 3, m_\mathrm{F} \neq 0\rangle$ states represents 6×10^{-3} of the selected atoms. The uncompensated shift of these $m_\mathrm{F} \neq 0$ atoms is thus $\sim 2 \times 10^{-3}$ of the total collision shift. This contribution is taken into account in the uncertainty of the collisional shift.

With the short term stability of FIG. 2, the linear extrapolation to zero density is performed with a statistical uncertainty of 10^{-16} in only two days.

Notice that the adiabatic passage allows to measure all atom number dependent effects without distinction. Both the collisional shift and the cavity pulling [13] are accounted for.

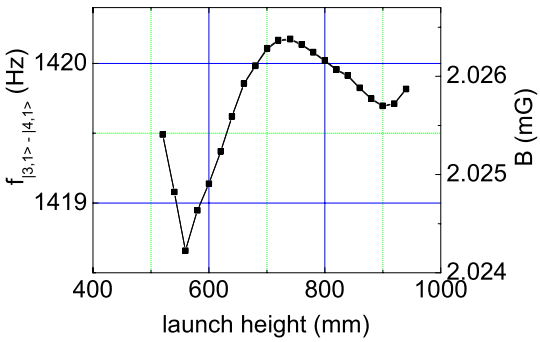


FIG. 5: Map of the magnetic field in the FO2 fountain.

Quadratic Zeeman effect

In the past, the uncertainty related to the quadratic Zeeman effect has been a significant contribution to the overall uncertainty, partly due to the strong magnetic field fluctuations (∼ 30 mG peak-to-peak) in the Paris urban environment. Recently, significant improvements in the control of this effect have been achieved with an active stabilization where the corrections applied to several coils are suitably weighted. The resulting magnetic field instability is low enough that an automatic measurement every 15 minutes of the central fringe of the Ramsey pattern of the $|F = 3, m_F = 1\rangle \leftrightarrow |F = 4, m_F = 1\rangle$ transition can be performed. As an example, Figure 4 shows the result of such measurements over a month. The magnetic field exhibits peak-to-peak fluctuations smaller than ±20 pT over this entire period, resulting in an uncertainty of only 4×10^{17} for FO2. Figure 5 shows the magnetic field as a function of the launch height. The magnetic field is homogeneous at the 0.1% level.

For FO1 we observe unexplained jumps whose mag-

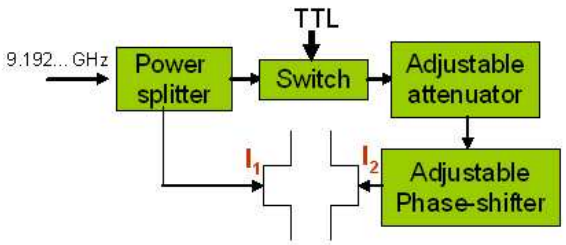


FIG. 6: Schematic of the symmetric feeding of the interrogation microwave cavity.

nitude constrains us to extend the uncertainty to one fringe (± 1 Hz or ± 0.14 nT) in the measurement of the magnetic field.

Distributed cavity phase shift

The interrogation cavity is fed using two oppositely located feedthroughs as displayed in FIG. 6. The input interrogation signal is splitted in two paths, one of which is switchable and adjustable in phase and amplitude. Using atoms as a probe, it is possible to control and balance the fields inside the cavity both in phase and amplitude to better than $\sim 1\%$. Additionally, repeated measurements show that this adjustment is stable at the 1% level over a few weeks. At this level, electrical asymmetry in the couplings can become the limiting factor for the cancellation of the linear component of the phase distribution.

In order to evaluate the effect of the linear component of the phase distribution, the clock frequency for symmetric (Sym) and asymmetric (I_1 or I_2) feeds are compared. With FO2, the measured shifts are :

$$\nu_{I1} - \nu_{Sym} = (-25.3 \pm 1.1) \times 10^{-16} \quad (3)$$
$$\nu_{I2} - \nu_{Sym} = (+24.0 \pm 1.2) \times 10^{-16} \quad (4)$$

Given the cavity quality factor in FO2 ($Q = 7700$), these values are consistent with the prediction of [14] assuming a tilt of the launch direction $\lesssim 1$ mrad with respect to gravity. Given the good coupling symmetry measured in FO2, the symmetric feed cancels the linear component of the phase distribution to better than 5%. At this level, the second order term in the phase distribution becomes the limiting factor. Based on the predictions of [14] and [15] concerning this quadratic term, numerical computations indicate that the corresponding shift does not exceed 3×10^{-16} for the two fountains. Note that for FO1, the effect of the linear component is found 5 times smaller, due to a higher quality factor and a better alignment of the launch direction. Despite a larger asymmetry between the couplings, the symmetric feed reduces the effect of the linear component below the quadratic term.

Note that [15] has recently revisited the issue of the distributed cavity phase shift. In future work, this improved model will be included in our evaluation, allowing for a better assessment of this effect.

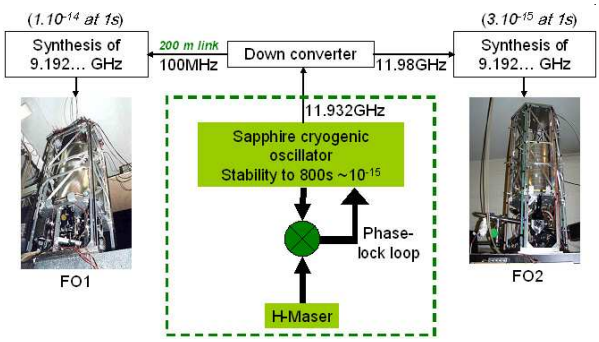


FIG. 7: Schematic of the frequency synthesis and distribution system.

Other frequency shifts

The blackbody radiation shift [16] [17] is evaluated based on the Stark measurement of [18]. A 10% uncertainty is attributed to the T^6 term [16] . In FO2, an ensemble of thermistors monitors the absolute temperature and its gradient leading to a ± 1 K uncertainty. Similarly, the entire set-up of FO1 is at the laboratory temperature which is monitored with an uncertainty of ± 1 K.

In order to evaluate the effect of spectral impurities of the interrogation signal and microwave leakages, we measure the clock frequency as a function of the microwave power. For each microwave power the cold collision shift is measured and extrapolated to zero, using the above described adiabatic passage method. For both fountains the measured shift between $\pi/2$ and $3\pi/2$ pulses is less than the typical resolution of 3.3×10^{-16}. Since this measurement alone does not allow to distinguish between the effects of synchronous perturbations, microwave leakages and spectral impurity, we conservatively use this measurement uncertainty with no extrapolation as a function of the microwave power.

Additionally after the comparison described in the following, we have operated and compared the two microwave synthesizers on the same fountain (FO2). This measurement has shown no frequency shift between the two synthesizers with a resolution of 1 part in 10^{16} [10].

DIRECT COMPARISON BETWEEN FO1 AND FO2

For the first time we have operated two atomic fountains using an ultra stable cryogenic sapphire resonator oscillator as interrogation oscillator. Figure 7 describes the frequency synthesis and distribution system developed at SYRTE in order to use the CSO.

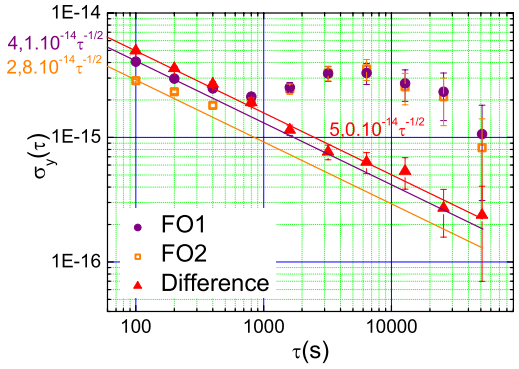


FIG. 8: Fractional frequency instability of FO1 (•) and FO2 (□) including high density and low density data and after all corrections are applied. The ▲ represents the fractional frequency instability of the difference between the two clocks.

These systems are designed to induce minimal degradations of the remarkable stability of the CSO [10]. For FO2, while staying in the microwave domain, the output signal from the CSO is converted into a 9.192 GHz signal. The resulting interrogation signal stability is 3×10^{-15} at 1 s. For delivering the reference signal to FO1 at $\sim$ 200 m distance, the output of the CSO is converted down to a 100 MHz signal which is distributed through a high stability RF cable and multiplied again up to 9.192 GHz. These steps degrade the stability of the interrogation signal to $\sim 1 \times 10^{-14}$ at 1 s. Additionally, the 100 MHz signal is used to phase-lock the CSO and all the related signals to the 100 MHz output of the hydrogen maser with a time constant of $\sim$ 1000 s. With this scheme, we realize a high quality metrological reference signal combining the short term stability of the CSO and the long term behavior of the H-maser. When comparing the two clocks, all data are time-tagged using an NTP protocol and subsequently referenced to the same local time scale (UTC(OP)), allowing for an efficient common mode suppression of the frequency fluctuations of the reference signal in the comparison.

During the comparison, each fountain operates in a differential mode for evaluating and cancelling the collision shift in real time. For the two clocks, both high density and low density data are used to compute the average frequency over synchronous time intervals of 100 s, including corrections for all systematic shifts. Figure 8 shows the fractional frequency instability for the two clocks, including high density and low density data fully corrected. As we can see, one of the two fountains, after all corrections are included, has a remarkable stability of 2.8×10^{-14} at 1 s, better than a H-maser for all timescales. For averaging times higher than 100 s, each clock measures the frequency fluctuations of the reference signal. Figure 8 also shows the fractional frequency instability of the difference between the two clocks. This combined instability is found equal to 5.0×10^{-14} at 1 s and decreases with a white frequency noise behavior down to 2.2×10^{-16} at 50 000 s. Hence, least one of the two fountains reaches a stability smaller than 1.6×10^{-16} at the same averaging time. This is the first time that such a stability is observed between two primary frequency standards.

The average frequency difference is found equal to $(4 \pm 2.2) \times 10^{-16}$. This value is fully consistent with table I and therefore strongly confirms the stated accuracies. Additionally, this demonstration of a measurement capability at the 10^{-16} level in a reasonable time ($\sim$ 1 day) leaves an excellent prospect to further investigate systematic shifts at this level and to bring the accuracy down to $(1-2)$ parts in 10^{16}.

CONCLUSION

For the first time, a direct comparison between two primary standards using cryogenic oscillator has been realized. An unprecedented stability of 2.2×10^{-16} at 50 000 s has been achieved. During this comparison, one of the fountains has shown a short term stability of $2.8 \times 10^{-14}\tau^{-1/2}$ after corrections for all frequency shifts are included. The measured fractional frequency difference between the two clocks of only $(4 \pm 2.2) \times 10^{-16}$ strongly confirms the clock accuracies (± 7.5 and $\pm 6.5 \times 10^{-16}$ respectively).

To further improve these results we have already upgraded the frequency synthesis and distribution system [9]. Future work on atomic fountain will focus on a thorough investigation of frequency shifts related to the interrogation signal (synchronous phase perturbations, ...) as well other currently most limiting systematic shifts (distributed cavity phase shift, ...). Achieving a combined fractional frequency instability of $(2-3) \times 10^{-14}$ at 1 s and an accuracy of $(1-2)$ part in 10^{16} appears as a realistic goal.

Acknowledgment

The authors wish to thank the technical and electronics staffs of the SYRTE. This work was supported in part by the LNE and the CNRS. SYRTE and Laboratoire Kastler-Brossel are Unités Associées au CNRS, UMR 8630 and 8552.

* Present address: ESA, Fundamental Physics Missions Division, SCI/SP Keplerlaan 1 - P.O.Box 299 2200, AG Noordwijk ZH, The Netherlands. Luigi.Cacciapuoti@esa.int

† Present address: CNES, 18, Av. Edouard Belin, 31401, Toulouse cedex 9, France. Michel.Abgrall@alten.cnes.fr

[1] A. Clairon *et al*, Proceedings of the Fifth Symposium on Frequency Standards and Metrology, p. 49-59, *World Scientific* (1995)

[2] S. Weyers *et al*, "Uncertainty evaluation of the atomic caesium fountain CSF1 of the PTB", *Metrologia* 38, 343-352 (2001)

[3] S. R. Jefferts *et al*, "Accuracy evaluation of NIST-F1", *Metrologia* 39, 321-336 (2002)

[4] F. Levi *et al*, "Systematic shift uncertainty evaluation of IEN CSF1 primary frequency standard",*IEEE trans. on Instr. and Meas.*, vol. 52, p. 267, (2003)

[5] J.Y. Richard, P. Uhrich, A. Clairon, S. Bize, H. Marion, Ph. Laurent, M. Abgrall, I. Maksimovic, J. Grunert, A. Bauch, S. Weyers and R. Wynands, "Comparison of remote cesium fountains using GPS P3 and TWSTFT links", Proc. of 18th European Freq. and Time Forum, Guildford, United-Kingdom, (2004)

[6] T.E. Parker, P. Hetzel, S.R. Jefferts, S. Weyers, L.M. Nelson, A. Bauch, and J. Levine, "First Comparison of Remote Cesium Fountains", Presented at the 15th European Frequency and Time Forum, Neuchâtel Switzerland, (2001)

[7] H. Marion *et al*, "Search for Variations of Fundamental Constants using Atomic Fountain Clocks", *Phys. Rev. Lett.*, vol. 90, p. 150801, (2003)

[8] S. Chang *et al*, *IEEE trans. on Instr. and Meas.*, vol. 50, p. 519, (2001)

[9] D. Chambon *et al*, "High performance flywheel source for atomic fountains and advanced metrology applications", Proceeding at the 18th European Frequency and Time Forum, Guildford United-Kingdom, 5-7 April (2004)

[10] D. Chambon, S. Bize, F. Chapelet, F. Narbonneau, M. Lours, A. Clairon, G. Santarelli, "Design and Characterization of New Microwave Synthesizers for Metrological Applications", presented at the 19th European Frequency and Time Forum, Besançon France, 21-24 March (2005)

[11] F. Pereira Dos Santos *et al*, "Controlling the Cold Collision Shift in High Precision Atomic Interferometry", *Phys. Rev. Lett.*, vol. 89, p. 233004, (2002)

[12] H. Marion *et al*, "First observation of Feshbach resonances at very low magnetic field in a ^{133}Cs fountain", Proceeding at the 18th European Frequency and Time Forum, Guildford United-Kingdom, 5-7 April (2004)

[13] S. Bize *et al*, "Cavity Frequency Pulling in Cold Atom Fountains", *IEEE Trans. Instrum. Meas.*, vol. 50, p.503, (2001)

[14] R. Schröder, U. Hübner and D. Griebsch, "Design and Realization of the Microwave Cavity in the PTB Caesium Atomic Fountain Clock CSF1" *IEEE Trans. Instrum. Meas.*, vol. 49, p.383, (2002)

[15] R. Li and K. Gibble, "Phase variations in microwave cavities for atomic clocks", *Metrologia*, vol. 41, pp. 376-386, (2004)

[16] W. Itano *et al*, "Shift of $^2S_{1/2}$ hyperfine splittings due to blackbody radiation", *Phys. Rev. A*, vol. 25, pp. 1233-1235, (1982)

[17] A. Bauch and R. Schröder, "Experimental verification of the shift of the caesium hyperfine transition frequency due to blackbody transition radiation", *Phys. Rev. L*, vol. 78, p. 622, (1997)

[18] E. Simon, P. Laurent and A. Clairon, "Measurement of the Stark shift of the Cs hyperfine splitting in an atomic fountain", *Phys. Rev. A*, vol. 57, p. 436, (1998)

FREQUENCY BIASES ASSOCIATED WITH DISTRIBUTED CAVITY PHASE AND MICROWAVE LEAKAGE IN THE ATOMIC FOUNTAIN PRIMARY FREQUENCY STANDARDS IEN-CSF1 AND NIST-F1†

S.R. Jefferts, J.H. Shirley, N. Ashby, T.P. Heavner, E.A. Donley
NIST –Time and Frequency Division,
325 Broadway
Boulder , CO USA 80305

F. Levi
Istituto Elettrotecnico Nazionale – Torino IT

E.A. Burt, G. J. Dick
Jet Propulsion Laboratories, Pasadena, CA USA

ABSTRACT†

We show that the frequency bias caused by distributed cavity phase has a strong dependence on microwave power. We also show that frequency biases associated with microwave leakage have distinct signatures in their dependence on microwave power and the physical location of the leakage interaction with the atom.

1. INTRODUCTION

The subject of frequency shifts in atomic frequency standards caused either by distributed cavity phase or microwave leakage goes back to the earliest days of the thermal beam standards [1,2], and has been the subject of continuing theoretical and experimental work over the last fifty years. [3-8]. Laser-cooled fountain frequency standards pose problems with respect to both distributed cavity phase and microwave leakage different from those associated with thermal beam standards. This is due both to the very different microwave structure used in fountains as well as the low center-of-mass velocity and very narrow velocity distribution, which allows operation at significantly elevated microwave power in fountain standards.

We discuss the dependence on microwave power of the frequency biases induced in an atomic fountain by the distributed cavity phase (DCP) and microwave leakage. The entire discussion takes place within the limit of small detuning because the velocity of the atoms, typically a maximum of 3m/s, results in a maximum Doppler shift of only 100 Hz.

The analysis of the phase of the microwave field within the typically TE_{011}, microwave cavities used in cold-atom fountain frequency standards has been the subject of a large body of work. DeMarchi provided the seminal contributions of proving a correspondence between the phase and power-flows within the microwave cavity, identifying the preferred cavity configuration and modeling the phase of the microwave field within the cavity [9-11]. Three-dimensional analyses of the microwave field within the cavity have been completed by several authors [12-16]. In many of these studies large phase excursions are predicted. At least one author has used these phase excursions to predict frequency shifts in the $\delta f/f \simeq 10^{-16}$ range as a result of distributed cavity phase [14], larger than the current value assigned in the error budget of many operational fountain type primary frequency standards. We note that the theory presented here allows one to test that prediction.

We begin by briefly reviewing the properties of the microwave field and within a resonant TE_{011} cavity. We then solve the Schrödinger equation for a two-level atom in the case where the field within the microwave cavity has both real and imaginary parts, which is the case required to understand both the distributed cavity phase frequency shift and the microwave leakage frequency shift. We then examine the dependence on microwave field amplitude of the frequency bias caused either by distributed cavity phase or microwave leakage.

2. MICROWAVE CAVITIES AND PHASE

The typical microwave cavity used in cesium (and rubidium) fountain frequency standards is a cylindrical cavity resonating in the TE_{011} mode at the hyperfine frequency of the atom, 9.193 GHz in the case of cesium. See Fig. 1. The "z-axis" of the cavity is aligned with the gradient of the gravitational potential and atoms enter and leave the cavity via below-cutoff waveguides. This cavity has the crucial property of allowing relatively large diameter ($2r_a \sim 1$cm at $\nu_0 = 9.2$ GHz.) cylindrical waveguides for atoms to enter and leave the cavity without unduly influencing the TE_{011} mode of the cavity and thereby causing large phase gradients in the microwave field [10]. To lowest order the field within the cavity is describable as purely TE ($E_Z \equiv 0$) and all field components are derivable from the longitudinal magnetic field $H_z(x,y)$ [17]. Under these assumptions and using a trivial extention of the notation of DeMarchi in [11], the longitudinal field can be written as

$$H_Z(x,y,z) = \left|H(x,y)\right| e^{i\varphi(x,y)} f(z)\hat{z}, \quad (1)$$

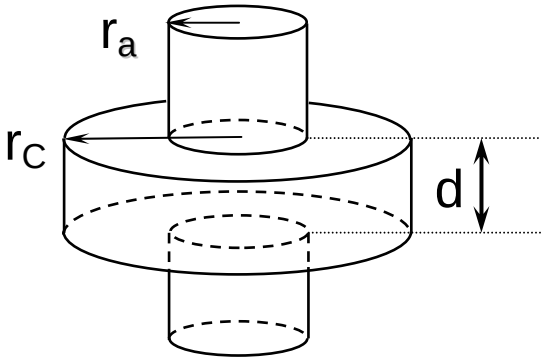


Figure 1. A schematic diagram of the TE_{011} microwave cavity used in most cesium fountain frequency standards.

where $\varphi(x,y)$ is the distributed cavity phase under discussion here. The expression for the real part of the field within the cavity is given (with a change to cylindrical coordinates ρ,ϕ and z) to lowest order by

$$H_Z(\rho,\phi,z)=\left(\frac{\pi}{2}\right)H_0 J_0\left(\frac{p'_{0,1}\rho}{r_c}\right)\sin\left(\frac{\pi z}{d}\right), \quad (2)$$

where J_n is the nth Bessel function and $p'_{n,m}$ is defined as the mth solution of $\frac{dJ_n(x)}{dx}=0$, r_c is the radius of the cavity, d is the height of the cavity, and $\pi H_0/2$ is the field amplitude at the center of the cavity (see Fig. 1 and [17]). The cavity phase is defined by the relation

$$\varphi\equiv\tan^{-1}\left(\frac{\mathrm{Im}(H_Z)}{\mathrm{Re}(H_Z)}\right), \quad (3)$$

where the real part of H is approximated by Eq (2) and the imaginary part is, roughly, approximated by

$$\mathrm{Im}(H_Z)=\frac{\pi H_0}{2Q}\left[\left(1+\left(\frac{\rho}{r_c}\right)^2+\left(\frac{\rho}{r_c}\right)^4\cos 4\phi\right)\left(\frac{z}{d}-\frac{1}{2}\right)^2\right]. \quad (4)$$

Here Q is the quality factor of the cavity. The second term in Eq. (4) comes from power flow to the walls from the standing mode of the field, and the third term comes from power flow associated with the four-feed structure of the particular cavity investigated here [12]. The terms proportional to $\frac{\rho}{r_c}$ and $\frac{z}{d}$ are negligible for a symmetric, well constructed cavity and are not included here. The overall cavity phase has been set equal to zero at the geometric center of the cavity.

Several things can be seen by inspection of the previous equations. First, the phase of the microwave field within the cavity is independent of microwave power (as it should be). The effect on the atom of the imaginary part of H_z is, however, unlikely to be independent of microwave power. Second, the microwave phase can become relatively large when the real part of H_z is sufficiently small, or equivalently, the Rabi frequency of the atom is sufficiently small. What is ultimately of interest is not the value of the phase angle of the microwave field, but rather the value of the phase angle imposed on the atomic wave function, and any resulting frequency biases. In order to quantify the effect of the imaginary part of the microwave field upon the atom, we now obtain a solution to the time-dependent Schrödinger equation as the atoms pass through a cavity with fields described by Eq. (2) & (4).

3. SCHRÖDINGER EQUATION AND RAMSEY LINESHAPES

We extensively employ the theoretical framework developed by Shirley [18-20] and present here the extensions required to handle both the real and imaginary phases of microwave field.

The Hamiltonian for the system can be written as (cf. (7) of [20])

$$\mathcal{H}=\hbar\begin{pmatrix}\omega_a & 2b\cos\omega t+2b'\sin\omega t\\ 2b\cos\omega t+2b'\sin\omega t & \omega_b\end{pmatrix}, \quad (5)$$

where $\hbar\omega_a$ and $\hbar\omega_b$ are respectively the energies of the upper and lower states. The interaction Rabi frequency for the real part of the microwave field is given by $2b=\mu_B g\,\mathrm{Re}(H_Z)/\hbar$, where μ_B is the Bohr magneton, g the Landé g-factor and H_Z the microwave magnetic field parallel to the quantization axis imposed by the external c-field. A similar expression applies for b', $2b'=\mu_B g\,\mathrm{Im}(H_Z)/\hbar$, the Rabi frequency due to the imaginary part of the microwave field. Both $b(t)$ and $b'(t)$ are time-dependent, owing to the atoms motion in the cavity. We also define $b_0\equiv\frac{g\mu_B}{2\hbar}H_0$.

In the rotating wave approximation, the Hamiltonian in 1.5 is written

$$\mathcal{H}=\hbar\begin{pmatrix}\omega_a & (b+ib')e^{-i\omega t}\\ (b-ib')e^{i\omega t} & \omega_b\end{pmatrix}. \quad (6)$$

Note the sign change in $b'(t)$ in the off-diagonal couplings. This comes about because the rotating wave approximation selects one exponential from $\sin\omega t$ in one coupling and the other exponential in the other coupling (Compare to (7) and (8) in [20]). Using the "phase factored" solutions, α and β, (cf (9 & 10) of [20]) gives us, finally, the Schrödinger equation for the system,

$$i\hbar\frac{d}{dt}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}=\hbar\begin{pmatrix}-\Delta & b+ib'\\ b-ib' & \Delta\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}. \quad (7)$$

With the initial condition $\alpha(0)=1$ and $\beta(0)=0$ then α is the probability amplitude that the system remains in its initial state and β the probability amplitude that the

system changes state. Δ is half the detuning $\delta\omega$ from the atomic resonance ω_0: $\Delta = \frac{1}{2}(\omega - \omega_0) = \frac{\delta\omega}{2}$, where $\omega_0 = \omega_a - \omega_b$ is the hyperfine splitting of the atom. Δ, b, and b' are all real, possibly time-dependent, quantities.

We have given a solution to (7), valid through first order in Δ and b', under the assumption that the detuning is small compared to the Rabi frequency, in [21]. We quote here only the final results.

Using our solutions mentioned above and assuming that the real part of the excitation is the same for both Ramsey pulses leads to a transition probability of

$$P = \frac{\sin^2(2b_0\tau)}{2}\left[1 + \cos\delta\omega T_R + \left\{\begin{matrix}(\varepsilon_2 - \varepsilon_1)\csc b_0\tau + \\ (\eta_1 + \eta_2)\sec b_0\tau\end{matrix}\right\}\sin\delta\omega T_R\right]. \quad (8)$$

The term $b_0\tau = \int_0^\tau b(t)\,dt$, that is, b_0 is the average value of $b(t)$ during the excitation time τ, which amounts to $\pi/4$ at optimum power, and $\varepsilon_{1,2}$ and $\eta_{1,2}$ are proportional to the imaginary part of the microwave field b'. The transition probability, P, is clearly a normal Ramsey fringe (the $\cos\delta\omega T_R$ term) plus an underlying fringe of small amplitude and $\pi/2$ displacement (the $\sin\delta\omega T_R$ term). The latter asymmetrically distorts the Ramsey curve and leads to a frequency bias. The frequency bias is proportional to the difference between the transition probabilities P_L and P_R on the left and right sides of the central Ramsey fringe at equal detunings. This difference is given by

$$P_L - P_R = \sin^2(2b_0\tau)\left[\frac{\varepsilon_2 - \varepsilon_1}{\sin(b_0\tau)} + \frac{\eta_2 + \eta_1}{\cos(b_0\tau)}\right], \quad (9)$$

where $P_{L,R}$ is the probability given by (8) on the left and right sides of the central Ramsey fringe respectively (that is with $\delta\omega = \mp\pi/2T_R$, respectively). Eq (9) is plotted below in Fig 2. The complicated nature of the power dependence of the frequency bias associated with distributed cavity phase is immediately apparent in Fig. 2. This signature should allow measurement of the frequency bias associated with distributed cavity phase, or at least allow the placing of an upper limit on the effect.

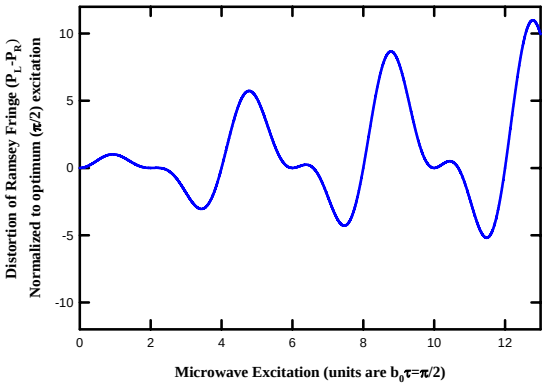


Figure 2. The distortion of the Ramsey fringe plotted caused by distributed cavity phase with respect to the microwave excitation (number of $\pi/2$ pulses). The distortion is proportional to the frequency bias and the data shown here are normalized to the shift at optimum power (one $\pi/2$ pulse).

4. MICROWAVE LEAKAGE

We can use the Hamiltonian, Eq. (6), also to investigate microwave leakage, except that in this case both b and b' are presumably small. We divide our analysis into three cases: leakage before the two Ramsey interactions, leakage between the Ramsey pulses, and leakage after the Ramsey interactions. Because the state-selection mechanism in NIST-F1 and IEN-CSF1 destroys any microwave coherence and projects the atoms into a pure F=3 state we can ignore the case of leakage before the first Ramsey interaction. We note however that the solution to the case of leakage before the first Ramsey interaction can be obtained from the case of leakage after the second Ramsey interaction by the substitution of $-t$ for t.

Some general results can be obtained [22]:

1. Microwave leakage in phase (the *b* term) with the field in the Ramsey cavity does not cause a frequency shift.
2. Microwave leakage in quadrature phase (the b' term) does not cause a frequency shift if it is applied symmetrically with respect to the center of the Ramsey interaction at $T_R/2$.
3. Microwave leakage in quadrature phase applied asymmetrically with respect to $T_R/2$ excites a Ramsey fringe shifted with respect to the central fringe ($a\sin(\delta\omega T_R)$ term) much like that in Eq. (8), which can cause a frequency bias.

We investigate the power dependence of these frequency biases next. Full solutions to the Schrödinger Equation for all three leakage cases are given in [22]. The various specific cases of microwave leakage are examined below.

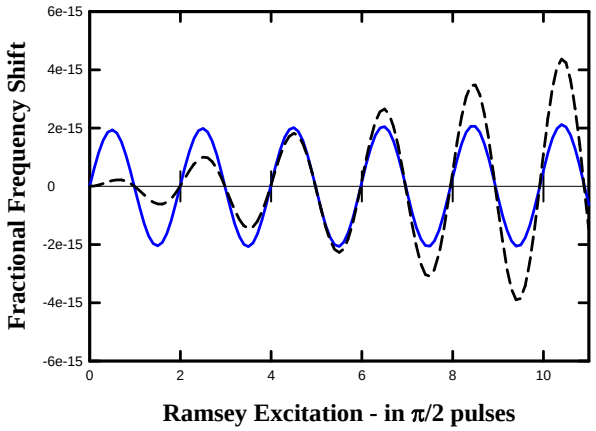


Figure 3. The modeled frequency shift induced by leakage above the Ramsey cavity in NIST-F1 and IEN-CSF1. The solid curve is the case where the leakage amplitude is constant while the field in the Ramsey cavity is increased. The broken curve is the case where the leakage field amplitude is proportional to the amplitude of the field within the Ramsey cavity. The microwave leakage power of each curve was been chosen for convenience and has no particular significance.

5. MICROWAVE LEAKAGE BETWEEN THE TWO RAMSEY PULSES

The dependence of the frequency bias on the amplitude of the microwave field in the Ramsey cavity is shown in Fig 3. It is immediately apparent that the frequency bias caused by microwave leakage is zero at integer multiples of the optimum excitation amplitude, $b_0\tau = (2n+1)\pi/2$, n=0,1,2... . A measurement of the residual frequency bias can be obtained with some leverage by measuring well away from optimum power, as illustrated in Fig. 4. In NIST-F1 the microwave power is set to within ± 0.1 dB of optimum, denoted by the dotted lines in Fig. 4. The frequency shift induced by microwave leakage at ± 3 dB from optimum excitation is some 30 times greater than the shift expected within 0.1 dB of optimum. The measurement at ± 3 dB therefore gives a "leverage" of 30 over the standard operating conditions, at least for the specific case of leakage above the Ramsey cavity.

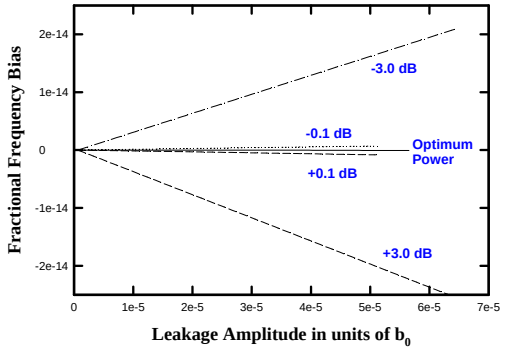


Figure 4. The frequency bias from microwave leakage above the Ramsey cavity can be measured in a "leveraged" fashion by operating well away from optimum microwave power in the Ramsey cavity. Assuming that the normal operating conditions keep the microwave power within 0.1 dB of optimum, a leverage of about 30 can be obtained by operating 3 dB above or below optimum instead.

6. MICROWAVE LEAKAGE AFTER THE RAMSEY INTERACTION

In the case of both NIST-F1 and IEN-CSF1 this is the most likely location for microwave leakage to interact with cesium atoms. The power dependence of this shift is shown in Fig. 5. The behavior of this shift is distinct from that caused by leakage above the Ramsey cavity in that its frequency with respect to the Rabi frequency is half that of the case of leakage above the Ramsey cavity. This signature can be used to identify the source of a frequency bias as being caused by leakage either above or below the Ramsey cavity. Unfortunately this shift is difficult to distinguish from the frequency shift caused by distributed cavity phase, as shown in Fig 6. Various approaches to separate these two effects are discussed in [22].

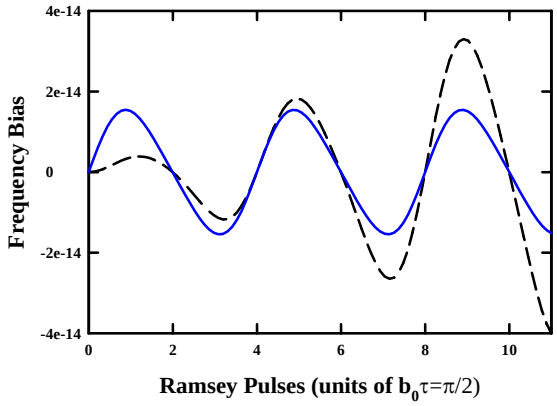


Figure 5. The modeled frequency bias induced by leakage below the Ramsey cavity. The solid curve is the case where the leakage amplitude is constant while the field in the Ramsey cavity is increased. The broken curve is the case where the leakage field amplitude is proportional to the amplitude of the field within the Ramsey cavity. In both cases the microwave power level has been chosen purely for convenience and has no particular significance.

7. DISCUSSION

The frequency biases caused by distributed cavity phase and microwave leakage have distinct signatures when measured as a function of the amplitude of the microwave excitation in the Ramsey cavity. The various shifts can be measured and either "fixed" by eliminating the source of the leakage (best), or perhaps calibrated and corrected.

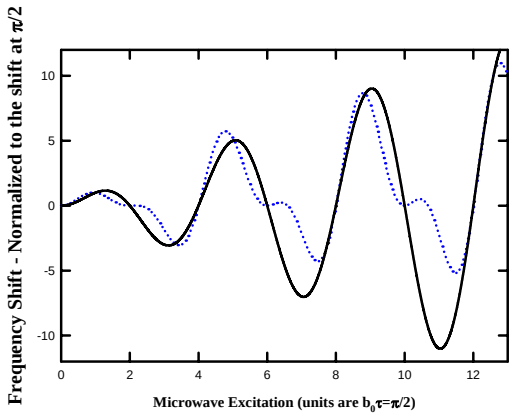


Figure 6. A comparison of the power dependence of the shift caused by microwave leakage (dotted curve) and the shift caused by microwave leakage below the Ramsey cavity (solid curve) as a function of the microwave excitation of the atoms in the Ramsey cavity.

It has been assumed here that the leakage terms must be treated in the small detuning limit. Various other approaches such as using the nonresonant AC Stark shift have been applied to the case of microwave leakage in a cesium fountain [23]. We hold that these cases are not valid approximations in a cesium fountain; the Doppler shift caused by atomic motion is, at 100 Hz maximum, much too small to be considered within the framework of an off-resonant AC Stark shift. The frequency shift caused by microwave leakage in a fountain depends on the amplitude of the microwave field and not on the power (square of the amplitude).

ACKNOWLEDGEMENTS

The authors gratefully acknowledge many useful discussions of microwave field effects in primary frequency standards with Andrea DeMarchi, Bob Drullinger, Bill Klipstein, Stefania Romisch and Tom Parker. We also thank David Smith, Mike Lombardi and Mark Weiss for their many useful suggestions on the manuscript.

REFERENCES

[1] N. Ramsey, Molecular Beams, Oxford – Clarendon Press 1956.

[2] R.C. Mockler, *"Atomic Beam Frequency Standards"*, Adv. Electron El. Phys., **15** (1962) pp.1-72.

[3] S. Jarvis Jr., "Molecular Beam Tube Frequency Biases due to Distributed Cavity Phase Variations", NBS Tech. Note 660 (1975).

[4] A. DeMarchi, J.H. Shirley, D.J. Glaze, and R.E. Drullinger, *"A New Cavity Configuration for Cesium Beam Primary Frequency Standards"*, IEEE Trans. Instrum. Meas. **37** (1988) pp. 185-190.

[5] A. Bauch, B. Fischer, T. Heindorff and R. Schröder, "Performance of the PTB reconstructed primary clock CS1 and an estimate of its current uncertainty", Metrologia **35** (1998) pp. 829-845.

[6] K. Dorenwendt and A. Bauch, "Spurious Microwave Fields in Caesium Atomic Beam Standards: Symmetry Considerations and Model Calculations", Proc. EFTF, 1999.

[7] B. Boussert, G. Théobald, P. Cérez and E. deClercq, *"Frequency Shifts in Cesium Beam Clocks Induced by Microwave Leakages"*, IEEE Trans. UFFC **45** (1998)728-738.

[8] M. Abgrall, "Evaluation des Performances de la Fontaine Atomique PHARO, Participation à l'étude l'horloge spatiale PHARO", PhD Thesis, University de Paris VI, January 2003.

[9] A. DeMarchi, "The Optically Pumped Caesium Fountain: 10^{-15} Frequency Accuracy?", Metrologia **18** (1982) pp. 103-116.

[10] G. Vecchi, A. DeMarchi, "Spatial Phase Variations in a TE_{011} Microwave Cavity for use in a Cesium Fountain Primary Frequency Standard", IEEE Trans. Instrum. Meas. **42** (1993) pp. 434-438.

[11] A Khursheed, G. Vecchi, A. DeMarchi, "Spatial Variations of Field Polarization in Microwave Cavities: Application to the Cesium Fountain Cavity", IEEE Trans UFFC. **43** (1996) pp. 201-210.

[12] S.R. Jefferts, R.E. Drullinger and A. DeMarchi, *"NIST Cesium Fountain Microwave Cavities"*, Proc. 1998 Intl. Freq. Cont. Symp. pp. 6-8.

[13] P. Laurent et al., *"Interrogation of Cold Atoms in a Primary Frequency Standard"*, Proc. 1999 Joint Mtg. IEEE Intl. Freq. Cont. Symp. and EFTF Conf. pp. 152-155.

[14] C. Fertig, R. Li, J. Rees, K.Gibble, "Distributed Cavity Phase Shifts and Microwave Photon Recoils", Proc. 2002 IEEE Intl. Freq. Cont. Symp. pp. 469-472.

[15] G.J. Dick, W.M. Klipstein, T.P. Heavner, S.R. Jefferts, "Design Concept for the Microwave Interrogation Structure in PARCS", Proc. 2003 Joint Mtg. IEEE Intl. Freq. Cont. Symp. and EFTF Conf. pp. 1032-1036.

[16] N. Ashby, S. Römisch, S.R. Jefferts, *"Endcaps for TE_{011} Cavities in Fountain Frequency Standards"*, Proc.

2003 Joint Mtg. IEEE Intl. Freq. Cont. Symp. and EFTF Conf. pp. 1076-1083.

[17] Robert E. Collin, Foundations of Microwave Engineering, 2nd Ed.,McGraw-Hill 1998.

[18] J.H. Shirley, *"Some Causes of Resonant Frequency Shifts in Atomic Beam Machines I. Shifts Due to Other Frequencies of Excitation"*, J. Appl. Phys. **34** (1963) pp.783-788.

[19] J.H. Shirley, *"Solution of the Schrödinger Equation with a Hamiltonian Periodic in Time"*, Phys. Rev. **138** (1965) pp. B979-B987.

[20] J.H. Shirley, W.D. Lee, R.E. Drullinger, *"Accuracy Evaluation of the Frequency Standard NIST-7"*, Metrologia **38** (2001) pp. 427-458.

[21] S.R. Jefferts, J.H. Shirley, N. Ashby, E.A. Burt, G.J. Dick, *"Power Dependence of Distributed Cavity Phase Induced Frequency Biases in Atomic Fountain Frequency Standards"*, IEEE Trans UFFC, in press.

[22] F. Levi, J. Shirley, S.R. Jefferts, *Microwave Leakage Induced Frequency Shifts in the Primary Frequency Standards NIST-F1 and IEN-CSF1, to be submitted.*

[23] K. Szymaniec, W. Chalupczak, P.B. Whibberley, S.N. Lea, D. Henderson, *Evaluation of the Primary Frequency Standard NPL_CsF1*, Metrologia 42 (2005) pp. 49-57.

MICROWAVE CAVITIES FOR ATOMIC CLOCKS: PHASE VARIATIONS AND ASSOCIATED POWER DEPENDENCES

Ruoxin Li and Kurt Gibble*

The Pennsylvania State University, Physics Department, University Park, PA, 16802 USA

ABSTRACT

We analyze the power dependent effects of phase variations in cylindrical TE_{011} cavities in atomic clocks. For a cavity with no holes, the dominant phase variations produce frequency shifts but have no power dependence. We calculate the power variation for cavities with holes in the endcaps and see significant power dependent frequency shifts due to the phase variations. For the simplest density variations, the power variations can be larger for two 2π, 4π, and 6π pulses, than for two half integer π pulses.

1. INTRODUCTION

In this paper we discuss the phase variations in microwave cavities, their effect on the transition probability and the apparent frequency of an atomic clock. In Ref. [1] we have given a detailed analysis of the phase variations in TE_{011} cylindrical cavities. We will summarize key results and use them to calculate the power dependent effects of the phase variations, which may be helpful for the evaluation of the distributed cavity phase shift.[2]

We begin by discussing the power variation of the transition probability for a cylindrical cavity with no holes in the endcap. We start with this cavity because the three-dimensional phase variations are analytic.[1] We first ignore the variation of the strength of the microwave field over the cavity aperture. We then include this variation of the tipping angle and show that it has a significant effect on the power dependences. Finally, we use our finite element solutions to analyze the power dependence in real cavities with holes in the endcaps.

2. PHASE VARIATION IN CYLINDRICAL CAVITIES

The phase is not uniform throughout a microwave cavity because a traveling wave delivers power from the feed to the walls to account for the losses in the metallic surfaces.[1,3] We first consider a TE_{011} cylindrical cavity with no holes in the endcaps. The phase of the field near the cylindrical axis of the cavity is:

$$\Phi(\vec{r}) \approx \frac{\delta}{2d}\left[k_1^2 r^2 - 2k_1 z \tan(k_1 z)\right] + \sum_{m,p}^{\infty} c_{mp} r^m \frac{\cos(k_p z)}{\cos(k_1 z)} \cos(m\phi) \quad (1)$$

where δ is the skin depth, d is the height of the cavity, $k_p = p\pi/d$, and the coefficient c_{mp} is from Ref. [1]. For m=0 and p=1, $c_{01}=0$ and, for m=0 and p>1, the r^2 term in the Taylor series should also be included.

As an example, we consider a TE_{011} cylindrical cavity with a 30 mm radius and a height of 26.075 mm that is fed by 4 feeds around the cavity midsection.[4]

$$\Phi(\vec{r}) \approx 32.9(r^2 - 2z^2) - 7.2r^4 \cos(4\phi) \quad (2)$$

Here Φ is in micro radians and r and z are in cm.

The first term in Eqs. 1 and 2 is azimuthally symmetric and caused by endcap losses. It has a quadratic variation in r and z near the center of the cavity and is independent of the number of feeds. The second term represents the power delivered at 4 feeds which is deposited on the side walls and endcaps. Note that the p=1 terms in Eq. 1 have no dependence on z. The higher order terms, p=3, 5, 7 …, have an oscillating variation in z that grows near the endcaps.

We note that concurrent work has treated a phase variation of $\Phi(\vec{r}) \approx r^2 z^2 / QR^2 d^2$ where Q is the quality factor of the cavity.[5] In Eq. 1, the terms r^2z^0 and r^0z^2 are much larger than the r^2z^2 terms.[1] For this reason, the results in Ref. [5] do not accurately describe the dominant behavior of atomic clocks with TE_{011} cylindrical cavities.

3. MODEL OF POWER DEPENDENCE

We analyze the effects of the phase variations using the sensitivity function s(t) for two pulses of arbitrary strength that have a relative phase of $\pi/2$. The effect of a small phase variation is:[6,7]

$$\delta P = \frac{1}{2}\int_{-\infty}^{\infty} s(t)\frac{d\Phi(t)}{dt}dt \quad (3)$$

During the first cavity passage, $s(t) = -\sin[\theta(r_2)]\times \sin[\theta(r_1,t)]$ where $\theta(r_2)$ is the total tipping angle during the second cavity passage, $\theta(r_1,t) = b_p\frac{\pi}{2} \times\int_{-\infty}^{t} H_{0,z}(r_1,t')dt'$ is the tipping angle throughout the first passage, $H_{0,z}$ is the standing wave field, and b_p is a constant we will use to vary the power. The sensitivity function is similar for the second passage.[6,7]

To evaluate distributed cavity phase shifts, it is helpful to integrate Eq. 3 by parts.

$$\delta P = -\frac{1}{2}\int_{-\infty}^{\infty}\frac{ds(t)}{dt}\Phi(t)dt \quad (4)$$

We change variables from time to position in the cavity

and substitute $\Phi(r,\phi,z)=\Phi_m(r,z)\cos(m\phi)$. We then calculate δP for an atom that traverses the cavity at r_1 on the upward passage and returns downward through the cavity at r_2. During the cavity transit, we neglect the small transverse velocity.

$$\delta P(\vec{r}_1,\vec{r}_2)=\frac{\pi}{4}\left(\sin[\theta(r_2)]\delta\Phi_{eff}(\vec{r}_1)-\sin[\theta(r_1)]\delta\Phi_{eff}(\vec{r}_2)\right) \quad (5)$$

$$\delta\Phi_{eff}(\vec{r})=b_p\cos(m\phi)\int_{z_{-\infty}}^{z_{+\infty}}\cos[\theta(r,z)]\times H_{0,z}(r,z)\frac{\Phi_m(r,z)}{v(z)}dz \quad (6)$$

Here $z_{\pm\infty}$ are the distances beyond which the fields are negligible. Because the effect of gravity is small for all cases studied here, we use v(z)=v.

In Eq. 5, it is clear that if the effective phase throughout the cavity is the same on both passages then there is no effect of the phase variations.[8] Further, one must consider the density distribution of the cold atoms because there is no population change, or frequency shift, if the density distribution is the same on both cavity passages. As in Ref. [1], we consider an expanding ball of atoms whose initial radius is larger than the radius of the hole in the cavity $r_a\approx 5$mm. The lowest order variation is a non-uniform density on the upward passage and, because the ball expands significantly, a uniform distribution for the second downward passage. With N as the total number of atoms that survive both cavity passages, the fractional population change is:

$$\overline{\delta P}=\frac{1}{N^2}\int^{r_a}\int^{r_a}n_1(r_1)n_2(r_2)\delta P(\vec{r}_1,\vec{r}_2)d\vec{r}_1d\vec{r}_2=\frac{1}{N\pi r_a^2}\int^{r_a}\int^{r_a}n_1(r_1)\delta P(\vec{r}_1,\vec{r}_2)d\vec{r}_1d\vec{r}_2 \quad (7)$$

From Eqs. 5 and 6, a distributed cavity phase shift can only occur when the density distribution has a cos(m′ϕ) dependence that matches the cos(mϕ) dependence of the phase variations. The lowest order variation is an m=0 azimuthally symmetric quadratic variation on the upward passage $n_1(r_1)=N/\pi r_a^2\left[1+\alpha\left(r_1^2/r_a^2-1/2\right)\times\cos(m\phi)\right]$, where α is small and negative. A fountain tilt or a displacement of the ball from the cavity axis on the upward passage leads to an m=1 term $n_1(r_1)=N/\pi r_a^2\left[1+\alpha\cos(\phi)r_1/3r_a\right]$ that has a linear variation over the aperture.[9] For m=2 and 4, we take the first quadratic variation of n_1 with r_1 as in the above for m=0. For m>0, the second term in Eq. 5 vanishes if the density on the second passage is homogeneous in Eq. 7.

From Eq. 7 we get the frequency shift of a clock by dividing by the slope of the Ramsey fringes. The slope for any atom is:

$$\frac{\partial P(r_1,r_2)}{\partial\nu}=-\frac{\pi}{\Delta\nu}\sin[\theta(r_1)]\sin[\theta(r_2)] \quad (8)$$

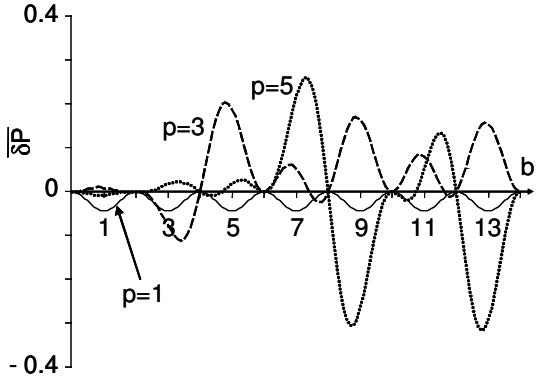


Figure 1. The change in transition probability $\overline{\delta P}$ versus the amplitude of the microwave field b_p, where $b_p=1$ corresponds to two π/2 pulses. Here the transverse variation of the tipping angle is neglected. For p=1 (solid line) there is no power variation of the clock's frequency. For p=3 (dashed) and 5 (dotted), there are significant power dependences. For large cavity radii, the scale of the phase shift Φ_{r_a} is very small for p>1. Note that δP =0 for b_p=2, 4, 6,…, corresponding to two integer π pulses.

where Δν is the Ramsey fringe width. For the ensemble, the slope of the Ramsey fringes is:

$$\overline{\partial P/\partial\nu}=\frac{1}{N^2}\int^{r_a}\int^{r_a}n_1(r_1)n_2(r_2)\frac{\partial P(r_1,r_2)}{\partial\nu}d\vec{r}_1d\vec{r}_2 \quad (9)$$

In the following we show $\overline{\delta P}$ because it is never singular. On the other hand, the frequency shift can be singular because $1/\overline{\partial P/\partial\nu}$ has a series of zeros as the power is increased.

4. ANALYTIC RESULTS FOR CYLINDRICAL CAVITIES

It is instructive to use the analytic solution of Eq. 1 to analyze the power dependence in a cylindrical cavity. To understand the effect of the different dependences of the phase on z, we first analyze a cavity where we ignore the variation in tipping angle near the cavity axis. We then include the tipping angle variation and show that it changes the character of the power variations. In the following section we show similar behaviors for cavities with holes in the endcaps.

The above equations simplify when we ignore the variation in the tipping angle for trajectories near the center of a cylindrical cavity. The standing wave field is $H_{0,z}(r,z)=\nu k_1 J_0(\gamma_1 r)\cos(k_1 z)/2$ and the tipping angle $\theta(r,z)=b_p\pi/4\left[1+\sin(\pi z/d)\right]$, where $\gamma_1=\sqrt{k^2-k_1^2}$ and k is the free space wave vector. For m=0, p=1, the first term in Eqs. 1 and 2, we only consider the r^2 term because the $z\tan(k_1 z)$ term has no transverse variation (although it does produce power dependent effects[1]).

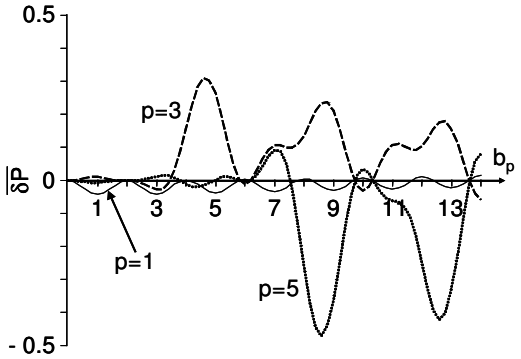


Figure 2. The change in transition probability δP as in Fig. 1 where the variation of the tipping angle is included. Here the cavity radius is R=3 cm. Again p=1 terms (solid line) have little power variation. For p=3 and 5 there are large power dependences. Note that the power sensitivity is significant for b_p=4, 8, and 12, corresponding to two 2π, 4π and 6π pulses.

We get:

$$\overline{\delta P} = \alpha \frac{\Phi_{r_a}}{24} \sin^2\left(b_p \frac{\pi}{2}\right) \qquad (10)$$

where $\Phi_{r_a} = \delta k_1^2 r_a^2 / 2d$. It is striking that this term produces no power dependent frequency shift because Eq. 8 has the same form. This is clear because the phase for the p=1 terms in Eq. 1 is independent of z and therefore constant throughout the cavity for any atomic trajectory. For m=1, $\Phi_{r_a} = c_{11} r_a$ and, for m=2 and 4, $\Phi_{r_a} = c_{m1} r_a^m / 2$. Because the p=1 terms dominate for large cavity radii and since they have no power dependence, it seems that power dependences cannot be used to evaluate distributed cavity phase shifts. However, the tipping angle variations near the cavity axis and the effects of the endcap holes produce significant power dependences.

The power dependence for p>1 phase variations are more interesting. We get:

$$\overline{\delta P} = \alpha \frac{\Phi_{r_a}}{24} \sin\left(b_p \frac{\pi}{2}\right) \times \int_{-\pi/2}^{\pi/2} b_p \frac{\pi}{4} \cos\left(b_p \frac{\pi}{4}[1+\sin(x)]\right)\cos(px)dx \qquad (11)$$

where $\Phi_{r_a} = c_{0p} r_a^2, c_{1p} r_a$ and $c_{mp} r_a^m / 2$ for m=0, m=1, and m=2 and 4 as above. We note that Eq. 11 reduces to Eq. 10 when p=1. We show the power variations in Fig. 1 for p=1, 3, and 5, $\alpha = -1$, and $\Phi_{r_a} = 1$. The transition probability is most sensitive to phase variations for half integer π/2 pulses, b_p=1, 3, 5, …. For integer π pulses, Eq. 11 has no effect due to phase variations, when the variation of the tipping angle is ignored.

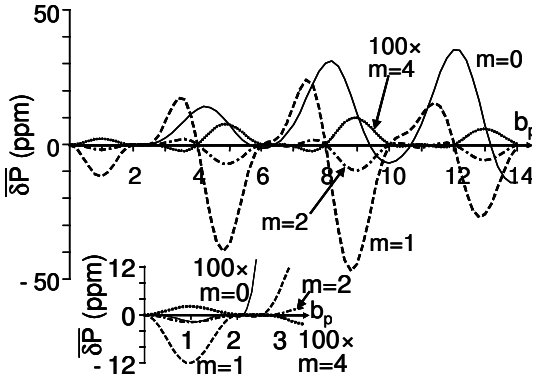


Figure 3. The power dependence for a cavity with R=3cm, an endcap hole of r_a=0.5 cm, and a quadratic density variation (α=−1). The m=0, 1, 2, and 4 terms correspond to phase variations which vary as cos(mϕ). The inset shows a magnified view for small amplitudes. For m=0, the sensitivity is largest for b_p=4, 8, and 12, corresponding to two integer 2π, 4π, and 6π pulses.

In Fig. 2 we show the change in transition probability when we include the variation of the tipping angle. For a cavity radius of R=3cm, the tipping angle varies by 11% over a centered 1 cm diameter. For p=1 phase variations, it has a small effect on the population differences. The population differences and the frequency shift decrease slowly for high power. This variation in the tipping angle leads to a significant sensitivity for p>1 and b_p=4, 8, and 12 as shown in Fig. 2. The p>1 terms are very much smaller than the p=1 modes for large cavity radii.[1]

5. RESULTS FOR CAVITIES WITH ENDCAP HOLES

The cavities used in fountain clocks must have holes in the endcaps to allow the atoms to enter and exit the cavities. Near the aperture in the endcaps, the microwave field is large near the surface of the endcap and the phase of the field approaches −π/4.[1] In addition, the standing wave field has a π phase shift away from the center of the cut-off waveguide sections that are added to prevent microwave leakage.

In Fig. 3 we calculate the power dependence for a cavity[4] with endcap holes using the finite element solutions from Ref. [1]. Here modes with definite m are clear but the different p modes are not easily separable. The endcap holes excite a number of p modes near the endcap holes.

As in Fig. 2, we see that the effect of the non-uniform tipping angles are large because the shifts are large at b_p=4, 8, and 12. In Fig. 3 we see that the ratio of the shift for 2 π/2 pulses versus higher powers is very different depending on the mode. Therefore the magnitude of distributed cavity phase shifts cannot be

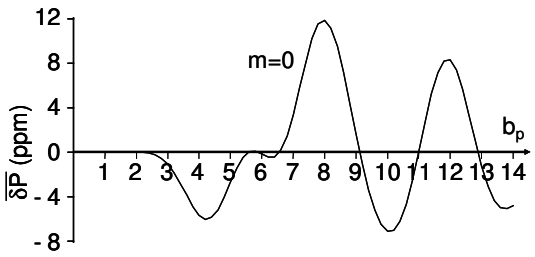


Figure 4 The m=0 power dependence for the cavity in Ref. [1] which has an endcap hole with a radius of r_a=0.6 cm. This cavity is designed to have no m=0 distributed cavity phase shift for two π/2 pulses and a quadratic atomic density variation (α=−1).

evaluated by only looking at power dependences. A combination of limits on density variations and an accurate calculation of the phase variations in the cavity are helpful. Clearly the m=0 and m=1 phase variations are most important. Higher m variations seem to be less important because, for most clocks, we expect the cos(mϕ) density variation to usually be very small.

In Fig. 4, we show the m=0 power dependence for the cavity design in Ref. [1]. This design avoids the large field and phase shift near the endcap holes and has no m=0 frequency shift for a quadratic density variation at b_p=1. It too shows larger sensitivity for b_p=4, 8, and 12, corresponding to two 2π, 4π and 6π pulses than for two half integer π pulses.

6. SUMMARY

We analyze the effect of phase variations in cylindrical TE_{011} cavities in atomic clocks. For a cavity with no holes, the dominant phase variation produces a frequency shift but has no power dependence. It is therefore unclear whether power variation alone is adequate to evaluate distributed cavity frequency shifts. We calculate the power variation for cavities with holes in the endcaps. The endcap holes, as well as the variation of the tipping angle for different atomic trajectories, lead to significant power dependent frequency shifts due to the phase variations. This work shows that, for the simplest density variations, one should also measure the population changes for two 2π, 4π, and 6π pulses, in addition to two half integer π pulses.

7. ACKNOWLEDGEMENTS

We gratefully acknowledge and financial support from NASA, ONR, and Penn State University. We thank the authors of Ref. [5] for an advance copy of their paper.

REFERENCES

[1] R. Li and K. Gibble, Phase variations in microwave cavities for atomic clocks, Metrologia, vol. **41**, pp. 376-386, 2004.

[2] A. Bauch, T. Heindorff, and R. Schroeder, Measurement of the frequency shift due to distributed cavity phase difference in an atomic clock, IEEE Trans. Instrum. Meas., vol. **34**, pp. 136-138, 1985.

[3] A. Khursheed, G. Vecchi, and A. DeMarchi, Spatial Variations of Field Polarization and Phase in Microwave Cavities; Application to the Cesium Fountain Cavity, IEEE Trans. UFFC, vol. **43**, pp. 201-210, 1996.

[4] S. R. Jefferts, J. Shirley, T.E. Parker, T.P. Heavner, D.M. Meekhof, C.W. Nelson, F. Levi, G. Costanzo, A. DeMarchi, R.E. Drullinger, L. Hollberg, W.D. Lee, and F.L. Walls, Accuracy Evaluation of NIST-F1, Metrologia, vol. **39**, pp. 321-336, 2002.

[5] See paper by S.R. Jefferts, J.H. Shirley, N. Ashby, E.A. Burt, and G.J. Dick, in these proceedings.

[6] G. Santarelli, P. Laurent, A. Clairon, G.J. Dick, C.A. Greenhall and C. Audoin, Theoretical description and experimental evaluation of the effect of the interrogation oscillator frequency noise on the stability of a pulsed atomic frequency standard, in Proceedings of European Frequency Time Forum, 1996, vol. **45**, pp. 66-71.

[7] P. Lemonde, G. Santarelli, Ph. Laurent, F. Pereira, A. Clairon and Ch. Salomon, The sensitivity function: a new tool for the evaluation of frequency shifts in atomic spectroscopy, in Proceedings of Frequency Control Symp., 1998, pp. 110-115.

[8] If the phase throughout the cavity is constant but the power is different during the two passages, there is also no change in population because $\delta\Phi_{eff}(\vec{r}) \propto \sin[\theta(r)]$.

[9] M. Abgrall, Ph.D. thesis, Evaluation des performances de la fontaine atomique PHARAO**,** Participation à l'étude de l'horloge spatiale PHARAO 2003, p. 147.

*Kurt Gibble, 104 Davey Laboratory, 232, Dept. of Physics, The Pennsylvania State University, University Park, PA, 16802, USA, Phone: 814-863-5343 Fax: 814-865-3604, Email: kgibble@phys.psu.edu

COMPARISON BETWEEN REMOTE CS FOUNTAIN PRIMARY FREQUENCY STANDARDS

D. Calonico, L. Lorini, F. Levi
IEN "G. Ferraris", Strada delle Cacce 91, 10135 Torino, Italy

S. Bize, H. Marion, F. Chapelet, J.-Y. Richard, A. Clairon, C. Salomon*,
LNE-SYRTE, Observatoire de Paris 61 Avenue de l'Observatoire 75014 Paris, France.
*Laboratoire Kastler Brossel, Ecole Normale Supérieure, 24 rue Lhomond, F-75231 Paris, cedex 05, France

K. Szymaniec, W. Chalupczak, D. Henderson
National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 0LW, UK

ABSTRACT

Results of the synchronous comparison between the Cs fountain primary frequency standards of IEN, SYRTE, and NPL are reported. The comparison lasted 20 days and made use of three different satellite techniques for time and frequency transfer (TWSTFT, GPS CP and GPS P3). The performance of the transfer techniques in the comparison experiment have been exploited.

1. INTRODUCTION

Ultra-cold atomic primary frequency standards realize the SI unit of time (and frequency) at the 1×10^{-15} accuracy level [1-5]. Up to now, five laboratories have operated Cs fountains and reported data for the calibration of the International Atomic Scale (TAI). These laboratories, IEN, SYRTE, PTB, NIST, and NPL, have demonstrated the operation of metrological systems composed of stable oscillators such as hydrogen masers, accurate and stable atomic primary references (fountains) and high performance methods to transfer time and frequency between remote laboratories.
The fountains at IEN, SYRTE, and NPL were compared during the period from 26 October 2004 to 15 November 2004 (Modified Julian Date 53304-53324),.
Each fountain used a hydrogen maser (HM) as local flywheel oscillator; the frequencies of the HMs were compared using remote satellite time and frequency transfer methods: Two Way Time and Frequency Satellite Transfer (TWSTFT) [6], and the Global Positioning System (GPS), with the GPS P3 code [7] and the GPS Carrier Phase (GPSCP) techniques [8].
The Time and Frequency Division of the BIPM and the Astronomical Institute of the University of Bern (AIUB) were also involved in this comparison, providing the GPGP3 and GPSCP data analysis [7,9]. PTB and NIST have participated in the comparison with hydrogen masers and time and frequency transfer tools to investigate the stabilities of all the possible links. The results of this analysis are presented in an accompanying paper [10].

The principal aim of the experiment was to compare the remote Cs fountain primary frequency standards at an uncertainty level of parts in 10^{15}. Moreover, the comparison provided an analysis of the stability of the different remote transfer techniques, and verified whether the links set a limit for the comparison due to their noise, or through the introduction of biases.
The experiment was also intended to establish whether frequency fluctuations of the local oscillators (masers) would be a significant limit to the accuracy of the comparison, when the fountains operate with significant dead time. With this aim, during the comparison period TWSTFT was operated with an intensified schedule (12 measurements per day, spaced by 2 hours) and all the laboratories agreed to provide fountain vs. maser measurements averaged every two hours.
Finally, the data obtained during this experiment have provided calibration of the TAI unit by these three primary frequency standards over the period MJD 53304-53324.

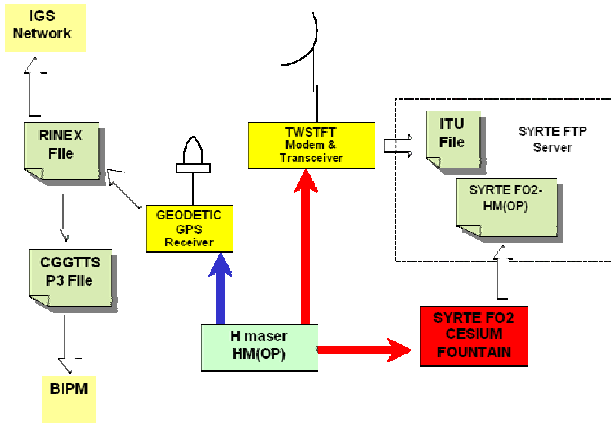


Figure 1. Block diagram of the time and frequency transfer scheme at SYRTE.

2. LINK SCHEME

In the experiment the TWSTFT, GPSP3 and GPSCP methods were used as the time and frequency transfer tools. As regards the GPSCP method, each lab participates in the International GPS Service [11], while

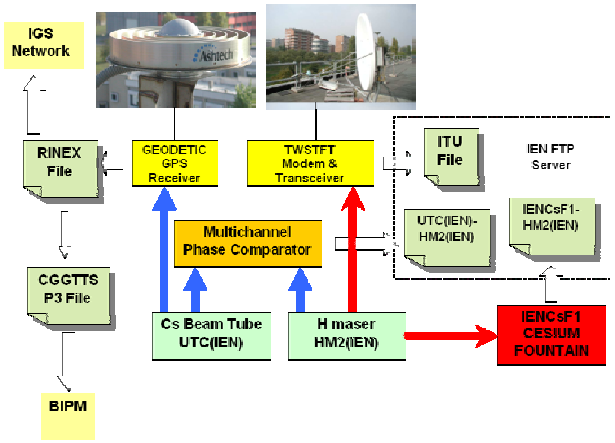


Figure 2. Block diagram of the time and frequency transfer scheme at IEN.

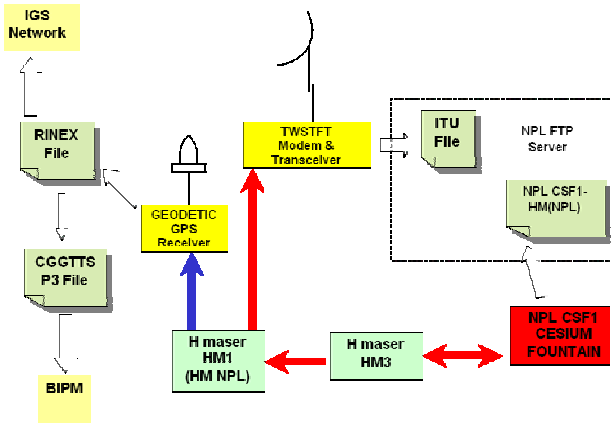


Figure 3. Block diagram of the time and frequency transfer scheme at NPL.

TWSTFT stations have been used in the past for other comparisons [12,13]. Figures 1–3 show the schemes by which the transfer receivers and stations were referred to the fountains in each laboratory.

The systems implemented at IEN, SYRTE and NPL differ at some points. At SYRTE, the hydrogen maser is the direct reference for both the GPS receiver and the TW station, its frequency being measured versus the fountain SYRTE-FO2.

At IEN, the hydrogen maser is the reference for the TW station only, whilst the GPS receiver is referred to UTC(IEN) generated by a commercial beam Cs clock. A further phase comparison is used to link UTC(IEN) to the hydrogen maser and subsequently to the fountain.

At NPL, both the GPS receiver and the TW station are referred to a hydrogen maser (HM1), while the fountain is the reference for another maser (HM3), which is continuously compared to HM1.

Other details of the implementation of transfer techniques in the three laboratories can be found in [10].

3. FOUNTAINS

The fountains participating in the experiment were SYRTE-FO2, IEN-CsF1 and NPL-CsF1. Their accuracies range between 0.7×10^{-15} and 1.1×10^{-15}; their uncertainty budgets are reported in Tables I-III.

Although the three fountains were corrected for the same physical effects, they were operated at different values of the magnetic field (about 0.2 μT in SYRTE-FO2 and NPL-CsF1 and about 0.1 μT in IEN-CsF1), and at different structure temperature, (about 300 K in SYRTE-FO2 and NPL-CsF1, about 340 K in IEN-CsF1), resulting in different magnitudes of the 2nd order Zeeman shift and the blackbody radiation corrections.

The methods used to evaluate the accuracy of the three fountains are not identical, in particular regarding the Zeeman and the collisional shifts [1-5].

EFFECT	BIAS (x 10^{-15})	uB (x 10^{-15})
2nd order Zeeman	192.17	0.03
Blackbody Radiation	-16.8	0.2
Microwave leakages	0	0.4
Pulling by other lines	0	0.1
Gravitational Red Shift	62.5	0.02
Cold collisions	7	0.3
TOTAL	244.9	**0.7**

Table I. Accuracy budget of the SYRTE-FO2 fountain.

EFFECT	BIAS (x 10^{-15})	uB (x 10^{-15})
2nd order Zeeman	45.9	0.2
Blackbody Radiation	-29.8	0.4
Microwave spectrum and leakage	0	0.1
Pulling by other lines	0	0.1
Gravitational Red Shift	26.4	0.1
Cold collisions	-1.2	1.0
TOTAL	42.3	**1.1**

Table II. Accuracy budget of the IEN-CsF1 fountain.

The stability analysis of the fountain vs. HM frequency measurements in the three laboratories are presented in Figure 4. Figures 5 to 7 report fountain vs. HM raw data, averaged every 2 hours.

EFFECT	BIAS (x 10^{-15})	uB (x 10^{-15})
2nd order Zeeman	142.4	0.1
Blackbody Radiation	-16.8	0.4
Microwave spectrum	0	0.1
Microwave leakages	1.1	0.3
Pulling by other lines	0	0.1
Gravitational Red Shift	1.3	0.1
Cold collisions	(-7.7)	0.8
TOTAL	128.0	**1.0**

Table III. Accuracy budget of the NPL-CsF1 fountain.

The SYRTE data show excellent stability in the short term, due to the high signal to noise ratio in SYRTE-FO2 and also to the local oscillator, a cryogenic sapphire oscillator phase-locked to the hydrogen maser [4]. In the medium and long term, the stability is limited to $1-2\times10^{-15}$ by fluctuations of the hydrogen maser (figure 5).

The IEN data show a stability of 3×10^{-13} for 1 s of integration time, limited by the signal to noise ratio of the IEN-CsF1 fountain, which uses a BVA quartz phase-locked to the HM as a local oscillator. The IEN data (figure 6) exhibit good long-term stability as the HM shows a stable linear frequency drift as low as 3×10^{-16} per day.

NPL-CsF1 was also operated using a BVA quartz phase-locked on the HM, but the short term Allan deviation was higher (7×10^{-13} at 1 s) owing to the way the collisional shift was evaluated [5]. However, the data in figure 7 show that the NPL HM was very stable in the long term, with a linear frequency drift as low as 1×10^{-16} per day.

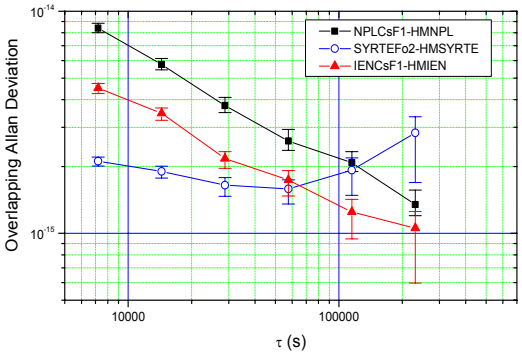


Figure 4. Allan deviation of the fountains SYRTE-FO2, IEN-CsF1 and NPL-F1 during the comparison.

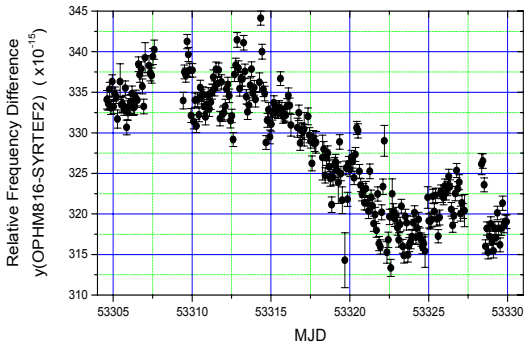


Figure 5. Frequency of the Hydrogen Maser measured versus the Cs fountain at SYRTE.

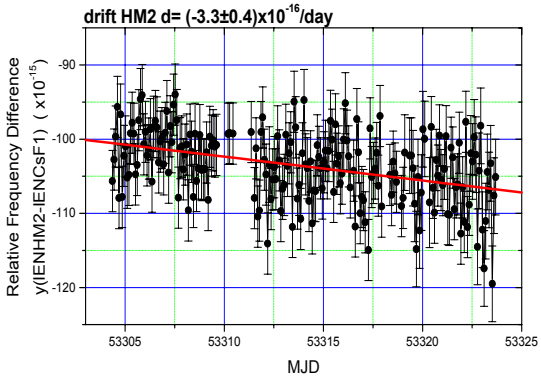


Figure 6. Frequency of the Hydrogen Maser measured versus the Cs fountain at IEN.

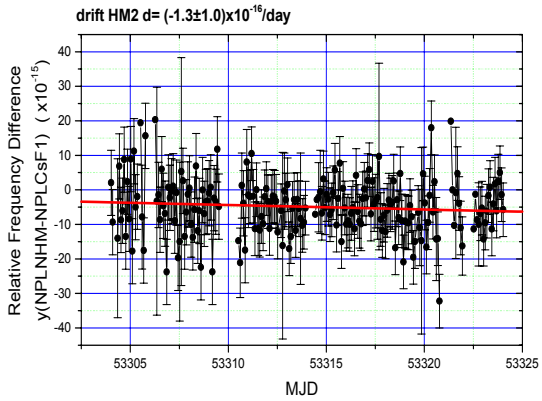


Figure 7. Frequency of the Hydrogen Maser measured versus the Cs fountain at NPL.

4. COMPARISON SCHEDULE AND DATA ANALYSIS

The whole comparison period (20 days) was divided into intervals of 2 hours. For each interval, a measurement of the maser frequency with respect to the fountains and a measurement of the frequency

difference between the masers via the transfer method was provided. The measurement interval for TWSTFT was 2 hours and the GPSP3 and GPSCP data were available even more frequently: it was thus possible to synchronize the time tags of all of these measurements (with 30 minute accuracy due to the TWSTFT schedule for different pairs of laboratories).

Once the time tags of the fountain and link measurements were synchronized, the frequency differences between the pairs of fountains (IEN-CsF1 – NPL-CsF1), (IEN-CsF1 – SYRTE-FO2), (NPL-CsF1 – SYRTE-FO2) were evaluated every 2 hours. This ensemble of values was then averaged.

This synchronous method enabled only those measurements intervals in which each pair of fountains to be compared was running at the same time to be retained (with 2 hour resolution). This approach enabled the potential limit to the comparison set by the dead time in the fountain measurements to be overcome, at the cost of considering only those measurement intervals during which each pair of fountains were running at the same time.

The common method used consisted of using the fountains to evaluate the average values of the maser frequencies in the two labs. By comparing these values with the average of the maser frequency differences obtained by the remote transfer method, the average frequency difference between the fountains was obtained. With this last method, the evaluation of the frequency difference between the remotely compared fountains requires a model for the maser's frequency fluctuation during the time intervals when it is not measured (dead time). This could be a critical point, especially in the presence of significant dead time or of fluctuations which are not as easy to model as linear frequency drift.

	IEN-SYRTE	IEN-NPL	SYRTE-NPL
	($x10^{-15}$)		
	3	5	1
	5	5	0
	3(4)	6(4)	1(4)
	4	5	1
	$\sigma = 2$ x10^{-15}		

Table IV. Results of the comparison (details in the text)

The results of the comparison are shown in Table IV. Each column reports the differences of the two labelled fountains. The first three lines report the results using the three transfer methods, using a differential data analysis. The fourth line reports the results of a standard data analysis based on the average of the collected data over the whole period of comparison, using in the example the TW link.

The uncertainty of these results is 2×10^{-15} for the GPSCP or TWSTFT methods and 4×10^{-15} with the GPSP3 transfer. It appears that a difference of about 4×10^{-15} is present between IEN-CsF1 and the other two fountains, whilst NPL-CsF1 and SYRTE-FO2 agree at the 1σ uncertainty level. This discrepancy (figure8), is yet under investigation. Further comparisons will be useful to measure it with greater accuracy.

The uncertainty budget for TWSTFT and GPSCP is reported in Table V, and for GPSP3 in Table VI.

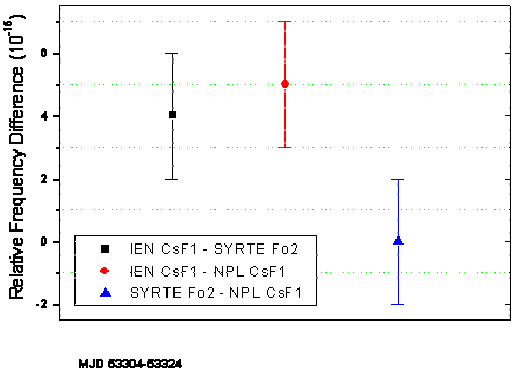


Figure 8. Relative frequency differences between IEN-CsF1, NPL-CsF1, and SYRTE-FO2 over the 20 day comparison

	IEN-SYRTE	IEN-NPL	NPL-SYRTE
	x10^{-15}		
CP link	0.7	0.7	0.7
TW link	1.5	1.5	1.0
Fountain-HM uA	0.4 - 0.2	0.4 - 0.6	0.6 - 0.2
Fountain-HM uB	1.1 - 0.7	1.1 - 1.0	1.0 - 0.7
TOTAL	**2 x10^{-15}**		

Table V. Uncertainty budget for the comparison using TWSTFT and GPSCP transfer methods.

We evaluated the contribution of the transfer method to the uncertainty by differentiating in pairs the maser frequency differences coming from thedifferent transfer methods. In this way, each "double difference" cancels out the contribution of the masers themselves and the contribution of the link alone is present. The Allan deviation of these double differences are reported in

Fig. 9. The value of the Allan Deviation at $\sim 4 \times 10^5$ s has been chosen as the contribution to the uncertainty, avoiding any hypotheses in extrapolating the instabilities reported in the graph to 20 days. This ensures a conservative evaluation of the uncertainty.

	IEN-SYRTE	IEN-NPL	NPL-SYRTE
	x10^{-15}		
GPSP3 link	4	4	4
Fountain-HM uA	0.4 - 0.2	0.4 - 0.6	0.6 - 0.2
Fountain-HM uB	1.1 - 0.7	1.1 - 1.0	1.0 - 0.7
TOTAL	4.2	4.3	4.2

Table VI. Uncertainty budget for the comparison using GPSP3 transfer method.

It should be noted that GPSP3 appears to show poorer short-term stability than TW and GPSP3. Moreover, there is evidence that the differences TWSTFT-GPSCP involving IEN are less stable than the those between SYRTE and NPL. This is due to the TWSTFT modem used at IEN, that in use at SYRTE and NPL being more stable.

From this comparison, the conclusion could be drawn that the TWSTFT and GPSCP transfer methods are equivalent at the present accuracy level and their results are in agreement, and that the same is true of GPSP3, which exhibits poorer short-term stability. This should be investigated further in future comparisons.

Once the Allan deviation slope for the TWSTFT and GPSCP are verified to be constant, it will be demonstrated that the contribution of the link via TWSTFT and GPSCP is not a limit for this 20 day comparison, compared to the uncertainty limit set by the accuracies of the fountains.

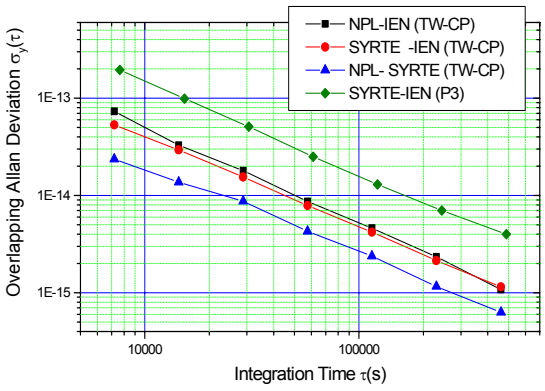


Figure 9. Allan deviation of the relative frequency differences between IEN-CsF1, NPL-CsF1, SYRTE-FO2 over the 20 days comparison

6. CONCLUSIONS

In this experiment, three Cs fountain primary frequency standards were directly compared for the first time. A synchronous schedule was used for the measurement, providing maximal rejection of the noise of the hydrogen masers used as flywheel oscillators.

Cs fountains have again been proven to be reliable metrological systems. This direct comparison has confirmed their realization of the TAI unit at 2×10^{-15} (1σ) uncertainty level.

Frequency differences larger than 1σ, but smaller than 3σ, were found between the fountains. This issue is under investigation. At the current accuracy level, the GPSCP and TWSTFT links have equivalent stability for 20 day comparisons, while GPSP3 exhibits poorer short-term stability. Further investigation is required to evaluate the real long-term limit set by the links.

As a forward look, further direct and synchronous comparisons, possibly with more fountains participating, will be useful to study both the agreement between the standards and the ultimate performance of the transfer methods.

7. ACKNOWLEDGEMENTS

The authors wish to thank all the colleagues who have been involved in the experiment, in particular with the transfer methods: J.Y. Richard , J. Achkar, P. Uhrich, P. Merck (SYRTE), V. Pettiti, D. Orgiazzi, (IEN), J. Davis, R. Hlavac (NPL), A. Bauch, D. Piester (PTB), T. Parker (NIST), G. Petit (BIPM), R. Dach (AIUB).

8. REFERENCES

[1] S. Jefferts, et al., Metrologia 39, 321-336 (2002).
[2] S. Weyers et al., Metrologia, 38, 343-352 (2001).
[3] F. Levi et al., IEEE UFFC., 51 (2004).
[4] C. Vian et al., IEEE Trans. Instrum. Meas. 54, 33 (2005).
H. Marion et al., Phys. Rev. Lett., 90, 150801 (2003).
[5] K. Szymaniec et al., Metrologia, 42, 49-57 (2005).
[6] D. Kirchner, Rev. of Radio Sc. 1996-1999, Ed. W. R. Stone, Oxford Univ. Press, Oxford, England, 1999.
[7] P. Defraigne, G. Petit, Metrologia 40, 184, (2003).
[8] K. Larson, J. Levine, IEEE Tr. UFFC, 45 (1998).
[9] R. Dach et al., these Proc. EFTF 2005.
[10] J. Ray, K. Senior, Metrologia 40, S270 (2003).
[11] T. Parker et al., Proc. IEEE Freq. Control Symp. (2001) and Proc. EFTF, Neuchâtel (2001).
[12] J.-Y. Richard et al., Proc. EFTF, Guildford (UK), (2004).
[13] A. Bauch et al., these Proc. EFTF 2005.

Session 5A / Time scales, algorithms and analysis

A Prediction Method Applicable to Steered Time Scales

Laurent-Guy Bernier

METAS - Swiss Federal Office of Metrology and Accreditation
Lindenweg 50, CH-3003 Bern-Wabern, Switzerland
e-mail: laurent-guy.bernier@metas.ch fax: +41-31-3233573

Keywords: time scales, steering, statistical prediction, optimal linear prediction

Abstract

The approach to the statistical prediction of a time scale process reported in this paper follows a sequence of three operations: pre-processing, prediction and post-processing. The pre-processing operation consists into removing both the voluntary (steering corrections) and the involuntary (accidental rate steps) statistical discontinuities present in the original process. The output from the pre-processing is a well behaved process with stationary statistical properties to which a statistical prediction algorithm can be applied. The post-processing consists into adding back the removed rate steps to transform the prediction of the pre-processed time scale into a prediction of the original time scale. The paper discusses the algorithms used in the prediction of UTC(CH) and reports the performance obtained.

1 Introduction

Contrary to the common practice, UTC(CH) is not a hardware clock but a computed time scale. Currently the Swiss time scales are computed on the basis of 3 cesium clocks and 1 hydrogen maser. Both UTC(CH) and TA(CH) are computed using the same time scale algorithm on the basis of the same clock measurements. The only difference is that TA(CH) is left free running while UTC(CH) is steered monthly to track UTC. A small drift of TA(CH) and of UTC(CH) is observed versus UTC because of the presence of the hydrogen maser in the clocks ensemble. However a constant drift is not a problem with respect to predictability. It is the presence of statistical discontinuities in the time scale process that makes prediction difficult.

In principle TA(CH) should be a non-stationary process with Stationary Second Increment (SSI). SSI processes constitute the class of well behaved noise processes typical of atomic clocks for which the Allan deviation statistic is stationary. SSI processes are also tractable regarding prediction because it is possible to construct linear prediction operators for which the error on the prediction is stationary. On the other hand UTC(CH), like any other steered process, is very difficult to predict because its statistical properties are disrupted every time it is steered. For a similar reason, accidental events affecting the rate of atomic clocks are a major problem regarding prediction because every step introduces a disruptive change in the statistical properties of the time scale process.

The approach reported in this paper follows a sequence of three operations: pre-processing, prediction and post-processing. The pre-processing consists into removing both the voluntary (steering corrections) and the involuntary (spontaneous rate steps) statistical discontinuities present in the original process. The output from the pre-processing is a well behaved SSI process to which a statistical prediction algorithm can be applied. The post-processing consists into adding back the removed rate steps to transform the prediction of the pre-processed time scale into a prediction of the original time scale. This paper discusses the three processing operations used in the prediction of UTC(CH) and reports the performance obtained.

2 Statistical Prediction

Any statistical prediction method, be it sophisticated or simple, fails at actually predicting the future. Statistical prediction only tries to model, in a more or less optimal way, how the consequences of the past propagate into the future. For example optimal linear prediction [4] uses an adapted whitening filter to transform the original process to be predicted into a white noise process of minimal amplitude. The white noise process is then fed to an inverse filter that reconstitutes the original process. The white noise process is a causal source able to generate the original process in the sense that, since the inverse filter is causal, the present value of the original process is exclusively a

function of the past values of the white noise process. Then how can the optimal linear prediction be performed if the (future) values of the white noise process are unknown within the prediction interval ? Simply by setting the unknown values to zero. In other words it is the transient of the inverse filter, when the input is forced to zero, that performs the prediction.

The corollary of the above observation is that statistical prediction methods work as intended only if the statistical properties of the process to be predicted are stationary. If a statistical discontinuity occurs within the prediction interval, there is no way a statistical prediction method can anticipate this future event and the quality of the prediction is degraded. On the other hand, if statistical discontinuities have occurred in the past and have not been removed, then the statistical prediction parameters will be computed on the basis of recorded data that does not have homogeneous statistical properties. The resulting prediction parameters will be a compromise between the optimum prediction parameter sets adapted to each statistically homogeneous segment of the recorded process.

For the above reasons it is important to pre-process the original time scale before applying a prediction algorithm. The goal of the pre-processing is to remove the statistical discontinuities. In the case of atomic clocks and time scales, the goal is to transform the original process into a well behaved SSI process.

3 Pre-proc A: *Unfolding*

In the case of a steered time scale such as UTC(CH) the first part of the pre-processing consists into removing the statistical discontinuities introduced by the rate corrections applied each month to steer UTC(CH) vs UTC. This operation is called the *unfolding* of the steered process. A linear transformation of the free running process $x_f(t)$ exists [1]

$$x_u(t) = x_f(t) + x_0 + r_0 \times (t - t_0) \qquad (1)$$

for which the resulting *unfolded* process $x_u(t)$ is identical to the steered process $x_s(t)$ between the epoch t_0 of the last steering and the present epoch t

$$x_u(t)u(t - t_0) = x_s(t)u(t - t_0). \qquad (2)$$

In the above $u(t - t_0)$ is the unit step function, x_0 is a constant state offset and r_0 is a constant rate offset. To achieve the *unfolding* operation it is necessary to record both the free running process $x_f(t)$, e.g. TA(CH), and the steered process $x_s(t)$, e.g. UTC(CH). The *unfolded* process is identical to the steered process within the last segment and has the same statistical properties as the free running process. Figure 1 shows the original UTC-UTC(CH) process as published by BIPM. Figure 2 shows the *unfolded* process. A small frequency drift

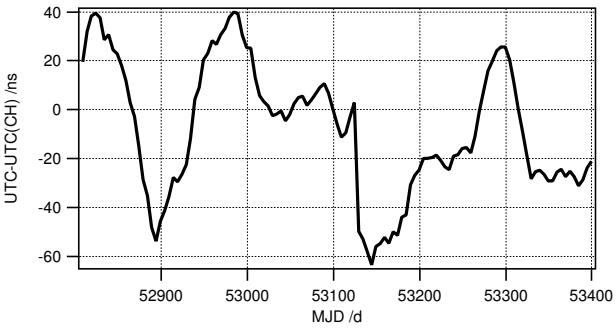


Figure 1: Original UTC-UTC(CH) process.

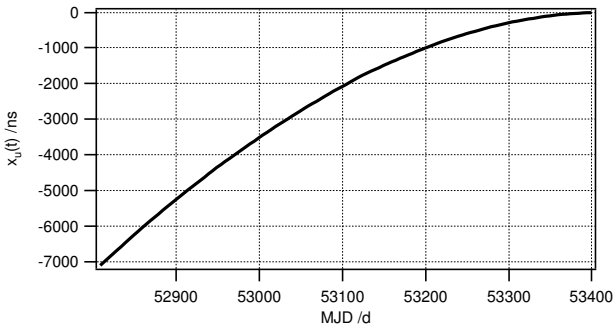


Figure 2: *Unfolded* UTC-UTC(CH) Process.

($-3.5 \times 10^{-16} d^{-1}$) is obvious when one looks at the parabolic shape of the *unfolded* process. This drift comes from the hydrogen maser which participates to the clocks ensemble.

4 Pre-proc B: Steps Removal

The *unfolded* process is already much easier to predict than the original process. However when we look at the drift removed residual, as shown on Figure 3, we see that the residual is not random and clearly diverges at the ends. This is a clue indicating that the model of a combination of noise and constant drift (parabola)

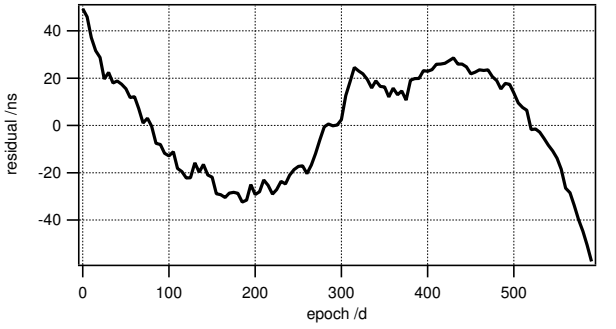


Figure 3: Drift removed residual of *unfolded* UTC-UTC(CH) process.

MJD	ns/d
53130	-0.85
53255	-1.00
53343	-0.92

Table 1: Rate steps removed from the *unfolded* UTC-UTC(CH) process.

does not adequately describe the observed data. Our first hypothesis was to assume that the drift is not constant and varies erratically. In that case the drift would be almost impossible to model. After investigation however it appeared that the drift is extremely stable and that the non-stationary perturbations have their origin in rate steps that occur spontaneously. It is very difficult theoretically to distinguish between a *random* rate step that could be explained by a flicker noise or a random walk (which over an infinite observation time averages to zero) and a genuine *non-stationary single event* rate step that changes the frequency forever. Nevertheless in practice we observe that after a long period of observation during which the residual shows a random characteristic, the residual starts to diverge and if a rate step correction is not performed the residual diverges more and more. This phenomenon really looks like a genuine single event rate step.

This assumption is confirmed by the fact that over an observation interval of 590 d it is sufficient to assume only three small rate steps in addition to a constant drift in order to produce a random residual with very small amplitude. Each rate step is defined by two parameters: an epoch of occurrence and a rate offset. The parameters of each removed step are estimated by minimizing the RMS drift removed residual. The three rate steps that were estimated and removed are listed in Table 1. The drift and steps removed residual is shown on Figure 4. The RMS value of the residual is only 3.6 ns over an observation interval of 590 d. With such a small residual it is difficult not to believe in the model. The drift appears to be extremely stable and the rate steps hypothesis fits very well the deviation of the residual from a parabola. The Allan deviation of the drift removed residual is shown on Figure 5, with and without steps removal. With the rate steps removed, the Allan deviation of the residual goes down to 2×10^{-15} for an averaging interval of 45 d.

5 DGSF-1 Prediction

It should be noted that the accuracy of the estimation of the rate steps does not affect the accuracy of the prediction because the effect of each step removal during the pre-processing is cancelled out by the addition of an identical step during the post-processing. On the other hand the *unfolding* operation is deterministic and

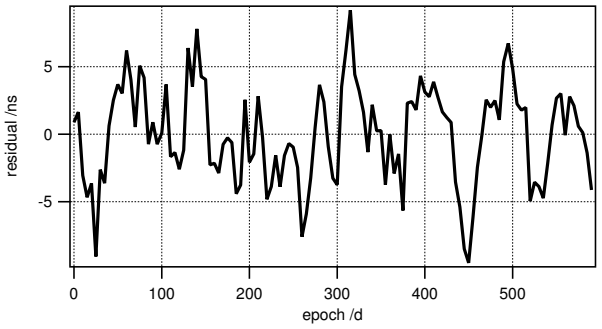


Figure 4: Drift and steps removed residual of the *unfolded* UTC-UTC(CH) process.

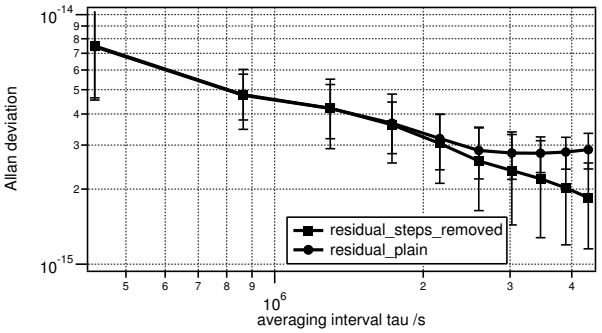


Figure 5: Allan deviation of the drift removed residual of the *unfolded* process, with and without steps removal.

does not contribute to the uncertainty. Therefore the uncertainty on the prediction of the original process UTC-UTC(CH) is equal to the uncertainty on the prediction of the steps removed residual. Again the goal of the *unfolding* operation and of the steps removal operation is to make the residual as close as possible to a SSI process so that statistical prediction can be applied under good conditions.

We use the DGSF-1 prediction algorithm for the prediction of the steps removed residual. The prediction algorithm is reported in details in [3]. The DGSF-1 prediction operator is a generalization of the simple second difference SF-1 prediction operator [2]

$$\hat{x}(t+\tau_1) = x(t) + \tau_1\, y(t,\tau_1), \tag{3}$$

where the time process $x(t)$ is the integral of the instantaneous frequency process $y(t)$

$$x(t) = \int_{-\infty}^{t} y(u)du, \tag{4}$$

and where $y(t,\tau)$ is the moving average of $y(t)$ over an averaging interval τ

$$y(t,\tau) = \frac{1}{\tau}\int_{t-\tau}^{t} y(u)du. \tag{5}$$

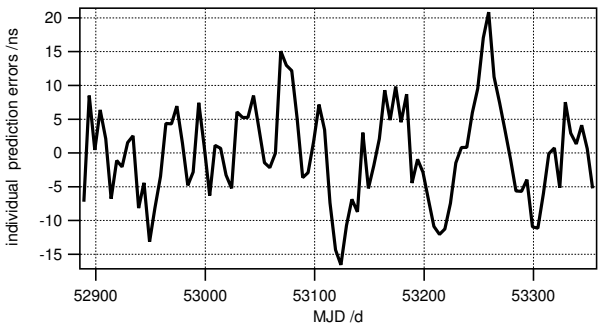


Figure 6: Individual DGSF-1 prediction errors for a 45 d prediction interval.

In the SF-1 prediction the prediction $\hat{x}(t+\tau_1)$ is simply the present state $x(t)$ of the time scale process extrapolated into the future, assuming that the initial average rate $y(t,\tau_1)$ will not change during the prediction interval τ_1. In the SF-1 prediction the averaging interval of the rate is identical to the prediction interval and the RMS error on the prediction as a function of the prediction interval is given by the Allan deviation of the time scale process

$$\text{rms}\,\{\epsilon(t,\tau_1)\} = \sqrt{2}\tau_1\sigma_y(\tau_1). \qquad (6)$$

The DGSF-1 prediction operator, on the other hand, is a simple generalization of the SF-1 prediction

$$\hat{x}(t+\tau_1) = x(t) + \tau_1\, y(t,\tau_2) + \frac{1}{2}\, d\tau_1^2 \left(1+\frac{\tau_2}{\tau_1}\right), \qquad (7)$$

where the averaging interval of the rate τ_2 is different from the prediction interval τ_1. This allows an optimization of the averaging interval τ_2 to minimize the RMS error on the prediction. d is a drift parameter that is also optimized in order to minimize the RMS prediction error. Its function is to compensate a deterministic drift in the time scale process. Figure 6 shows the sequence of all the individual prediction errors, for a prediction interval of 45 d, that can be computed on the basis of the recorded steps removed residual. For a given prediction interval τ_1 it is the RMS value of all past predictions that is minimized for the estimation of the optimum prediction parameters d and τ_2. In the present case the optimum prediction parameters are $d = -3.035 \times 10^{-3}$ (ns/d)/d and $\tau_2 = 80$ d yielding an RMS value of the prediction error of only 7 ns for a prediction interval of 45 d.

6 Post-Processing

Since the last segment of the *unfolded* UTC-UTC(CH) process is identical to the actual UTC-UTC(CH) process and since the DGSF-1 prediction algorithm takes care of the drift, the only post-processing necessary to produce a prediction of UTC-UTC(CH) is to add the removed rate steps to the prediction of the steps removed residual.

7 Discussion

An RMS prediction error of 7 ns for a prediction interval of 45 d is a very small value. So is the Allan deviation of the drift and steps removed residual which is 2×10^{-15} for a 45 d averaging interval. Moreover these results were achieved on the basis of only 3 cesium clocks and one hydrogen maser. The main comment to be made is that to achieve a good prediction the pre-processing, i.e. the removal of statistical discontinuities, is even more important than the choice of a particular statistical prediction algorithm among the set of available methods.

In practice the main problem we have for the prediction of UTC-UTC(CH) is the fact that spontaneous rate steps are detected only after a steering correction that does not take the new step into account has been performed. Circular T is published by BIPM 15 d after the last published epoch and the prediction is performed up to the next publication a month later. This explains why the prediction interval is 45 d. If a new rate step has occurred after the last published epoch, it cannot be detected so it it not taken into account in the current prediction. The rate step is detected only a month later when the next Circular T is published. By that time the UTC-UTC(CH) accumulated error due to the undetected rate step can become much larger than the 7 ns RMS error obtained when there is currently no undetected rate step.

One possible solution (if one can afford it) is to operate a large number of clocks, instead of only 4, so that a rate step experienced by one of the hardware clocks will have a negligible impact on the ensemble time scale.

The physical origin of the rate steps experienced by hardware clocks is not always easy to determine. We have observed rate steps in both our hydrogen maser and our cesium clocks. The time scale algorithm that we use gives a larger weight to the hydrogen maser because it has a better short-term stability than the cesium clocks in the ensemble. As a result a rate step of the hydrogen maser has a larger impact on the time scales than a rate step of one of the cesium clocks.

8 Conclusion

This paper reports a simple but practical algorithm used to steer UTC(CH) versus UTC. An essential part of the algorithm is the pre-processing step which consists into removing both the voluntary (steering corrections) and the involuntary (accidental rate steps) statistical discontinuities present in the original time

scale process. The output from the pre-processing is a well behaved SSI process with stationary statistical properties to which a statistical prediction algorithm can be applied successfully. Using the DGSF-1 prediction algorithm, which is capable of compensating a constant frequency drift, a RMS prediction error of only 7 ns is obtained for a prediction interval of 45 d when there is currently no undetected rate step.

References

[1] L. G. Bernier, "Experimentation at METAS with a Simple Steering Algorithm based on Linear Prediction," in *Proc. IVth International Time Scale Algorithms Symposium*, BIPM, Sèvres, March 18-29, 2002.

[2] L. G. Bernier, "Use of the Allan Deviation and Linear Prediction for the Determination of the Uncertainty on Time Calibrations Against Predicted Time scales," in *IEEE Trans. on Instrum. and Meas.*, vol. 52, no. 2, pp. 483–486, April 2003.

[3] L. G. Bernier, "Application of the GSF-1 Algorithm to the Near-Optimal Time scale Prediction of the Hydrogen Maser," in *Proc. PTTI 2003*, San-Diego, December 2003.

[4] J. Makhoul , "Linear Prediction: A Tutorial Review," *Proc. IEEE* , vol. 63, pp. 561–580, April 1975.

LONG-TERM STABILITY AND ACCURACY OF TAI

G. Petit

Bureau International des Poids et Mesures, 92312 Sèvres France gpetit@bipm.org

Keywords: EAL, TAI, TT.

ABSTRACT

International Atomic Time TAI gets its stability from some 250 atomic clocks worldwide that generate the free atomic scale EAL and its accuracy from a small number of primary frequency standards (PFS) which frequency measurements are used to steer the EAL frequency. While the long-term (above one year) stability of TAI is mostly driven by the PFS and therefore directly related to its accuracy, its stability for averaging times from one to a few months is mostly driven by EAL. Based on the analysis of 6 years (1999-2004) of TAI data, the paper presents several results. We estimate the one-month stability of EAL and TAI from the statistics and the estimated stability of weighted clocks. We study different classes of clocks which form EAL, the most important of these classes being Agilent 5071A and Hydrogen masers, and evidence systematic rate variations in clock ensembles. Their effect on the long-term stability of EAL is discussed. The long-term stability and accuracy of TAI is estimated with respect to the latest realization of TT(BIPM).

1. INTRODUCTION

International Atomic Time TAI, the time scale established by the BIPM, is a realization of Terrestrial Time TT, a coordinate time of a geocentric reference system. TAI gets its stability from some 250 atomic clocks kept in some 50 laboratories worldwide and its accuracy from a small number of primary frequency standards (PFS) developed by a few metrology laboratories. To be more specific, in the computation of TAI, a free-running time scale, EAL, is first established from a weighted average of some 250 atomic clocks, then the frequency of EAL is compared with that of the primary frequency standards using all available data processed with the algorithm presented in [1], and a frequency shift (frequency steering correction) is applied to EAL to ensure that the frequency of TAI conforms to its definition. Changes to the steering correction are expected to ensure accuracy without degrading the long-term (several months) stability of TAI, and these changes are announced in advance in the BIPM Circular T.

Because TAI is computed in "real-time" every month and has operational constraints (e.g. no re-computation on a given time interval even if new data become available), it does not provide an optimal realization of TT. The BIPM therefore computes another realization TT(BIPM) in post-processing [2,3], which is based on a weighted average of the evaluations of TAI frequency by the PFS. The latest realization is TT(BIPM2004) and provides the best reference which allows to estimate the achieved accuracy of TAI.

Since 1999, we have at our disposal the complete set of electronic files relative to the computation of TAI in a consistent form (see ftp://62.161.69.5/pub/tai). We use this database to study the long term stability of EAL over 72 months, from January 1999 to December 2004. In section 2, the 1-month instability of EAL is determined while, in section 3, some long-term trends in the EAL frequency are studied and possible causes are discussed. Section 4 provides the estimation of the accuracy of TAI.

2. EVALUATION OF THE 1-MONTH STABILITY OF EAL OVER 1999-2004

The database contains the series (w_{ij}).of weights attributed to clock i over the month j and their rates (r_{ij}).with respect to EAL over this month and the past 10. From these we compute the standard deviation over one month with 11 samples $\sigma(11, 1\text{month})$ and we may infer the Allan deviation $\sigma_y(1\text{month})$ by $\sigma_y^2(1\text{month}) = \sigma^2(11, 1\text{month}) / B_1$, where B_1 is a constant (e.g. about 1.9 for flicker frequency modulation) [4]. Because EAL is more stable than each individual clock, this provides a good estimate of the relative frequency stability of each clock for the month j. Then, assuming no correlation between the clocks, we may estimate the stability Σ_j of EAL over month j by $\Sigma_j^2 = \Sigma_i(w_{ij}\ \sigma_i^2) / \Sigma_i(w_{ij})$. In this study, we only consider clocks that were continuously operating for the 11-month period while the actual EAL computation takes into account all clocks that were present for at least the past 4 months. The difference is small, however, because the set of 11-month clocks generally accounts for about 90% of the total weight in EAL computation (see figure 1).

In the period considered in this paper, there were several changes that affect EAL: Until December 2000, the maximum weight of a clock was set to a fixed value (0.7%) but from January 2001, the maximum weight was set to 2/N, where N is the number of weighted

clocks, then it was set to 2.5/N from July 2002. To these discrete changes, we can add the significant and steady increase in the number of reported clocks. In addition the clocks are found to be significantly more reliable since 2001: In four years, the number of clocks continuously operating for 11 months and gaining at least one tenth of the maximum weight nearly doubled from about 120 end 2000 to about 220 end 2004 (Figure 1).

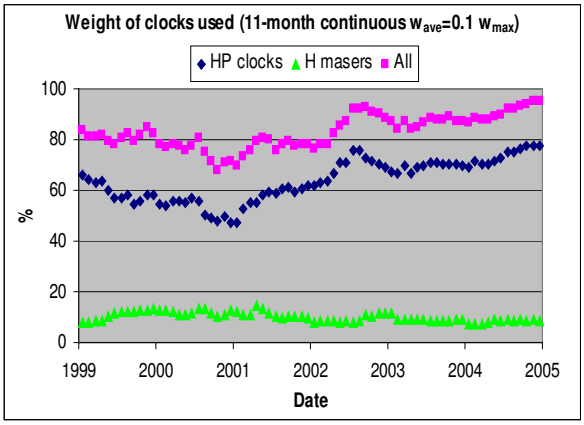


Figure 1: Cumulated weight of sets of clocks with continuous 11-month operation (Diamond: HP clocks; Triangle: H masers; Square: All types).

Figure 2 shows the 1-month frequency instability of EAL, Σ_i, computed as mentioned above for all months i from 1999 to 2004. We see that the change of weighting strategy in January 2001 yielded a big improvement while that on July 2002 yielded a smaller improvement. The computed value is at or below 5×10^{-16} since year 2001 and is at or below 4×10^{-16} since year 2003. This improvement is in line with the previous evaluation [5].

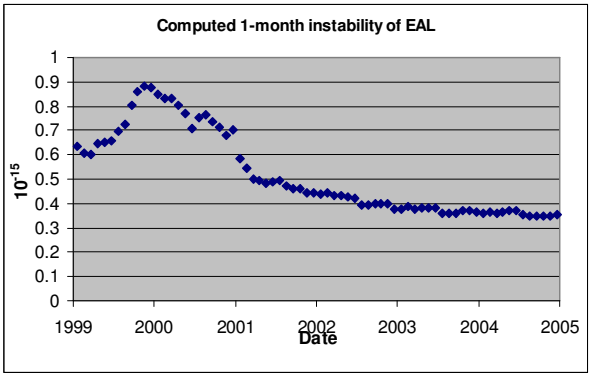


Figure 2: The 1-month frequency instability of EAL.

3. LONG-TERM TRENDS IN EAL

It has been noticed for years that some systematic long-term trends seem to occur in EAL: Indeed the rate difference between EAL and TAI (steering correction) has almost constantly decreased since the inception of this procedure in 1977 (see file ealtai03.ar in ftp://62.161.69.5/pub/tai/scale). However, this rate only approximates the actual rate of EAL because the steering correction is decided in real-time (every month) based on measurements of primary frequency standards available at the time of computation. It may also be affected by external factors such as the Recommendation S2 of CCDS(1996) to apply the Blackbody frequency shift to all primary standards, in practice shifting the rate of TAI by about 2×10^{-14}. A more reliable estimation of the true behavior of EAL is with respect to TT(BIPM), and is represented in Figure 3 for TT(BIPM2004). We see that, although the general tendency is to a decrease of the rate of EAL, some significant changes do occur e.g. from early-1997 to mid-1998, and also, apparently, since 2004.

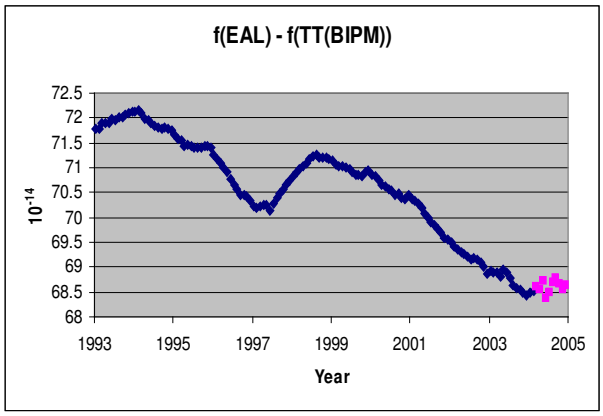


Figure 3: Rate of EAL with respect to TT(BIPM2004). Values extrapolated since April 2004, after the computation interval of TT(BIPM2004).

Because the rate of EAL is that of a weighted ensemble of clocks, we investigate whether groups of clocks may display some systematic behavior. As a first approach, we distinguish two groups: The Cs clocks Agilent 5071A (denoted HP in the following) and the hydrogen masers (denoted HM in the following). For each group, we compute the average rate drift with respect to TT(BIPM). This is computed each month using the past 11 months, i.e. the drift value for month m is the mean value of the drift over the year ending with m. For each computation, we keep only clocks which had a continuous operation over the whole year and whose average weight was at least 10% of the maximum weight. The results are presented in Figures 3 and 4.

We see (Figure 3) that the average drift for HM seems to be both quite large (of order 1×10^{-15} per month) and significantly biased, having the same sign for nearly the whole period.

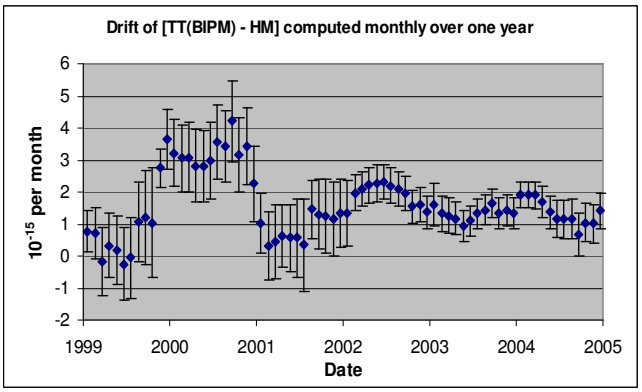


Figure 3: Average drift vs. TT(BIPM) for the HM set of clocks, computed monthly over the past year of continuous operation.

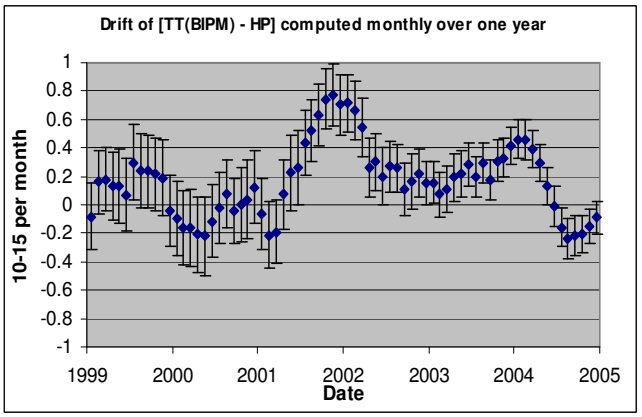


Figure 4: Average drift vs. TT(BIPM) for the HP set of clocks, computed monthly over the past year of continuous operation.

On the other hand, the average drift for HP (Figure 4) seems to be both smaller (of order a few 10^{-16} per month) and less biased, its sign changing occasionally over the period. However HM make 10-15% of EAL while HP make typically 70% of EAL (Figure 1). Thus the net result on EAL of the observed drifts in HM and HP is represented by an average drift with a significant bias (about $2\text{-}3\times10^{-16}$ per month) of which a good part may be originating from the masers. The drift also varies significantly (peak-to-peak variations of 1×10^{-15} per month) on time scales of years.

Such long-term trends correspond to a level of instability of about 2×10^{-15} for 6-month averaging time. This is very significant when comparing different evaluations of a Cs fountain at different periods, or when using different evaluations to estimate TAI accuracy.

4. ACCURACY OF TAI

TT(BIPM2004) has been computed in April 2004 using all primary frequency standards available following [3]. From this we compute the monthly rate difference between TAI and TT(BIPM2004), represented in Figure 5, where the bars represent the uncertainty of each monthly estimate following [1]. After April 2004, the monthly estimation of the TAI scale unit, published in the BIPM Circular T, have been used.

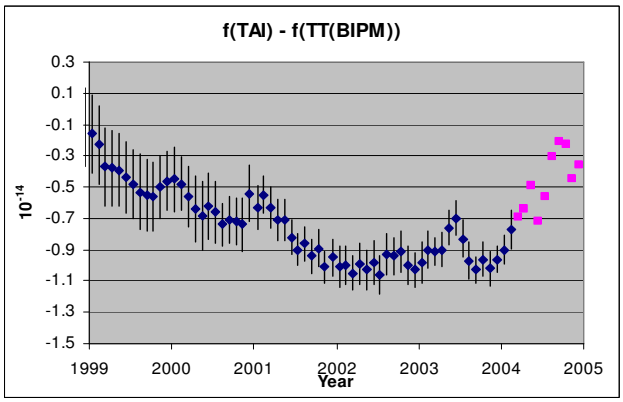


Figure 5: Rate difference between TAI and TT(BIPM2004) over 1999-2004

Primary Standard	Type /selection	u_B typical	Operation period
NICT-01	Beam /Opt.	$<4\times10^{-15}$	1998->....
IEN CSF1	Fountain	$<2\times10^{-15}$	2003->....
NIST7	Beam /Opt.	$X\times10^{-15}$	1995->2000
NIST-F1	Fountain	$<1\times10^{-15}$	1999->....
NPL-CSF1	Fountain	1×10^{-15}	2004->....
PTB CS1	Beam /Mag.	8×10^{-15}	1998->....
PTB CS2	Beam /Mag.	12×10^{-15}	1993->....
PTB CSF1	Fountain	1×10^{-15}	2000->....
SYRTE-JPO	Beam /Opt.	8×10^{-15}	1999->....
SYRTE-FO2	Fountain	$<1\times10^{-15}$	2002->....
SYRTE-FOM	Fountain	$<1\times10^{-15}$	2002->....

Table 1: Some characteristics of the main primary frequency standards reported over 1999-2004. The second column indicates Type (beam or fountain) and type of selection (Magnetic or Optical). Values for u_B may vary over the period and are only indicative.

The period used (1999-2004) corresponds to the first regular evaluation of Cs fountains, of which six now contribute to the estimation of TAI frequency (see Table 1 for the main characteristics of the PFS). As a result the uncertainty in the estimation decreased over the period from about 2.5×10^{-15} in 1999 to about $1.\times10^{-15}$ since 2003 (see the bars in Figure 5). This uncertainty is mostly driven by the values of the combined uncertainty from systematic effects, u_B, reported for the PFS (see Table 1). For some years, the steering strategy has failed to bring the TAI frequency close to that of the PFSs probably because the systematic frequency drift in EAL (see previous section) was counter-acting the steering

corrections. Since 2004, this is no more the case, and TAI is now conform to its definition within 3 standard deviations, as recommended by Recomendation 3 of the Consultative Committee on Time and Frequency in 2004.

4. CONCLUSION

By a reanalysis of all clock data used in the computation of TAI over 1999-2004, we have estimated the 1-month instability of EAL which is at or below $4x10^{-16}$ since 2003. The 1-month instability of TAI is similar because the frequency steering steps are kept at an average level below $4x10^{-16}$ per month.

Some long-term trends are evident when comparing EAL to TT(BIPM). A background frequency drift of about $1\text{-}2x10^{-16}$ / month may be due to H-masers, and random fluctuations with characteristic times of years may be due to the ensemble of Cs clocks. These trends correspond to a level of instability of about $2x10^{-15}$ for 6-month averaging time.

The accuracy of TAI is known to within $2x10^{-15}$ or below over the recent years. Although the scale unit of TAI has sometimes significantly diverged from its definition, this is no more the case since mid-2004 and TAI is now conform to its definition within 3 standard deviations.

REFERENCES

[1] Azoubib J., Granveaud M., Guinot B., *Metrologia* 13, 87, 1977.

[2] Guinot B., *Astron. Astrophys.*, 192, pp. 370-373, 1988.

[3] Petit G., A new realization of Terrestrial Time, *Proc. 35th PTTI*, pp. , 2003.

[4] Barnes and Allan, NIST Tech. Note 1318, 1990.

[5] Azoubib J., Proc. 32nd PTTI meeting, 195, 2000.

ON THE EVALUATION OF UNCERTAINTIES IN [*UTC – UTC(k)*]

W. Lewandowski[1], D. Matsakis[2], G. Panfilo[3,4], P. Tavella[4]

[1] Bureau International des Poids et Mesures, Sèvres, wlewandowski@bipm.org
[2] United States Naval Observatory, USA, matsakis.demetrios@usno.navy.mil
[3] Politecnico di Torino, Italy, panfilo@ien.it
[4] Istituto Elettrotecnico Nazionale Galileo Ferraris, Italy, tavella@ien.it

ABSTRACT

In this work, we present a study of the uncertainty of [*UTC – UTC*(*k*)][1]. In the first part of the paper, we consider an analytical solution considering the law of the propagation of uncertainty. In the second part, we verify the analytical results numerically, using the software used for the generation of UTC. In the third part the first use of this procedure in the official publication of BIPM Circular T is presented.

1. INTRODUCTION

Coordinated Universal Time (UTC), the worldwide time standard, is disseminated monthly through the publication of [*UTC – UTC*(*k*)] in *Circular T* by the Bureau International des Poids et Mesures (BIPM). Until now, these values have been published without their uncertainties, but the rules of the CIPM Mutual Recognition Arrangement (CIPM MRA) require the evaluation of this uncertainty. This paper reports the computing of these uncertainties [1], and their first use in *Circular T*.

UTC is derived from International Atomic Time (TAI) by the addition of leap seconds, while TAI is derived from the Free Atomic Timescale (EAL) through the addition of pre-announced frequency steers determined by comparison with a weighted set of primary frequency standards. EAL is a worldwide weighted average of a large number of free-running, effectively uncalibrated, frequency standards [2-4]. The uncertainty in the determination of EAL, TAI, and UTC, as steps in the realization of Terrestrial Time, is affected by three major elements: clock variations, the means of comparisons of remote clocks (time transfer), and the time-scale algorithm. The uncertainties of time transfer are particularly significant over averaging times of up to a few tens of days, and also influence the uncertainty of the access of the participating laboratories *k* to UTC, in other words the uncertainty of [*UTC – UTC* (*k*)]. In this work we present a study of the determination of the uncertainties of [*UTC – UTC* (*k*)]. In the first part of the paper, we develop an analytical solution using the law of the propagation of uncertainty. In the second part, we carry out a numerical verification of the analytical results using the software used for the generation of UTC. In the third part the first use of this procedure in the official publication of BIPM Circular T is presented.

[1] This paper is mostly based but not exclusively on a previous paper [1] presented at last PTTI.

2. THE ALGORITHMS OF EAL, TAI, AND UTC AND WORKING HYPOTHESIS

In order to obtain the uncertainty of [*UTC – UTC*(*k*)], we consider the general equation of the free atomic time scale (EAL). EAL is defined using the ALGOS algorithm [2-4] as:

$$EAL(t) = \sum_{i=1}^{N} w_i \left[h_i(t) + h_i'(t) \right] \qquad (1)$$

where N is the number of the atomic clocks, w_i the weight of the clocks, $h_i(t)$ is the reading of clock H_i at time t, and $h_i'(t)$ is the prediction of the reading of clock H_i to guarantee the continuity of the time scale. The weight attributed to a given clock reflects its long-term stability, since the objective is to obtain a weighted average that is more stable in the long term than any of the contributing elements [5, 6]. The weights of the clocks obey the relation:

$$\sum_{i=1}^{N} w_i = 1 \qquad (2)$$

Subtracting the same quantity from both sides of (1), we obtain:

$$EAL(t) - \sum_{i=1}^{N} w_i h_i(t) = \sum_{i=1}^{N} w_i \left[h_i(t) + h_i'(t) \right] - \sum_{i=1}^{N} w_i h_i(t)$$

Using (2) and rearranging we obtain:

$$\sum_{i=1}^{N} w_i \left(EAL(t) - h_i(t) \right) = \sum_{i=1}^{N} w_i h_i'(t). \qquad (3)$$

Setting

$$x_i(t) = EAL(t) - h_i(t), \qquad (4)$$

it is clear that equation (3) is of the form:

$$\sum_{i=1}^{N} w_i x_i(t) = \sum_{i=1}^{N} w_i h_i'(t). \qquad (5)$$

The software package termed ALGOS is used in the Time Section of the BIPM in order to generate UTC. Weights are determined from the variance of monthly average frequencies, subject to a maximum value [6]. The data

used by ALGOS take the form of the time differences between readings of clocks, written as:

$$x_{i,j}(t) = h_j(t) - h_i(t). \quad (6)$$

Equation (5) in conjunction with the $N - 1$ equations (6), results in a system with N equations and N unknowns:

$$\begin{cases} \sum_{i=1}^{N} w_i x_i(t) = \sum_{i=1}^{N} w_i h_i'(t) \\ x_i(t) - x_j(t) = x_{i,j}(t) \end{cases} \quad (7)$$

The solution is:

$$x_j(t) = EAL(t) - h_j = \sum_{i=1}^{N} w_i \left[h_i'(t) - x_{i,j}(t) \right]. \quad (8)$$

If we choose a particular clock H_j, we can see that the difference between that clock and EAL (8) depends on weights, clock prediction, and measured clock differences. The clock H_j may also represent a *UTC (j)* time scale; therefore, we may also interpret $x_j(t)$ as:

$$x_j(t) = EAL(t) - UTC(j)$$

The predictions and the weights are fixed by appropriate algorithms based on the past clock behavior and in (8) they can be considered as time-varying deterministic parameters. Sub optimal estimation of these parameters would affect the uncertainty of *TAI* as realization of the Terrestrial Time (*TT*), but they do not affect the knowledge of the difference between *EAL* and clock H_j; the measures $x_{i,j}$ are thus the only contributors to the uncertainties in x_j. In particular, we consider negligible the contribution of the uncertainty given by measures of clocks located inside the same laboratory.

Moreover, [*UTC – EAL*] depends only on pre-determined leap seconds and frequency steers that do not add uncertainty. The uncertainties of [*UTC – UTC*(*k*)] are therefore close to the uncertainties of [*TAI – UTC* (*k*)] and [*EAL – UTC* (*k*)].

We conclude considering the uncertainties of the links among laboratories as the only source of the uncertainty of [*UTC – UTC* (*k*)].

2.1 The law of propagation of uncertainty

According to [7], the uncertainty in the $x_j(t)$ can be found using the law of propagation of uncertainty. Let's indicate with *y* a generic quantity indirectly measured by means of direct measurements of the input quantity x_i,:

$$y = f(x_1, x_2, ..., x_M).$$

The expression of the law of the propagation of uncertainty [7] is given by:

$$u_y^2 = \sum_{i=1}^{M} \left(\frac{\partial f}{\partial x_i} \right)^2 u_{x_i}^2 + 2 \sum_{i=1}^{M-1} \sum_{k=i+1}^{M} \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_k} u_{(x_i, x_k)} \quad (9)$$

where the first term corresponds to the effect of the uncertainties on the input quantities x_i, and the second term accounts for the correlation between them.

In our case, we are interested in evaluating the uncertainty of the quantity $x_j(t) = EAL(t) - UTC(j)$ here playing the role of the indirect quantity *y*, and the uncertainty contributions are only due to the measurement noise of the links $x_{i,j}(t)$. Applying (9) to our model (8) yields:

$$u_{x_j}^2 = \sum_{i=1}^{N} \left(\frac{\partial x_j}{\partial x_{i,j}} \right)^2 u_{x_{i,j}}^2 + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} \frac{\partial x_j}{\partial x_{i,j}} \frac{\partial x_j}{\partial x_{k,j}} u_{(x_{i,j}, x_{k,j})} = \quad (10)$$

$$= \sum_{i=1}^{N} w_i^2 u_{x_{i,j}}^2 + 2 \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} w_i w_k u_{(x_{i,j}, x_{k,j})}$$

where $u_{x_j}^2 = u_{EAL - h_j}^2$.

The weights of the clocks are available from the BIPM Web site, and the uncertainties of links between the clocks [8] are published in *Circular T* (see the BIPM Time Section's FTP server at www.bipm.org).

The propagation of uncertainty (10) could as well be expressed by a matrix formulation using a multivariate weighted approach [12,13]; here, we present the scalar approach that allows one to clearly understand how the different components contribute to the combined uncertainty.

It can be demonstrated that the ALGOS algorithm would generate the same results if each laboratory's clocks were replaced by a single "equivalent" clock whose reading was the weighted average of the individual clocks and whose weight in EAL was the sum of the individual clock weights. Therefore, we can simplify the computations by summing the weights of the clocks at each lab as follows:

$$W_{Lab_k} = \sum_{i=1}^{N_{Lab_k}} w_i \quad (11)$$

where N_{Lab_k} is the number of clocks of the considered lab *k*.

Since we neglect the uncertainty of the measurement between clocks located inside the same lab, at any lab we can consider a single equivalent clock used as reference for the time transfer with the other labs. We note that the formalism of equation (10) could still be applied without these simplifying assumptions, as the double summation would account for the 100% correlation of the time transfer noise between clocks in the same lab.

2.2 Extension of the computation

It is not necessary to apply equation (10) to every laboratory. If the uncertainty for a clock is known, for example considering $x_j = EAL - h_j$, the evaluation of uncertainty on $x_i = EAL - h_i$ may be obtained by using the second equation in (7) and applying the property of the variance to obtain:

$$u_{x_i}^2 = u_{x_j}^2 + u_{x_{i,j}}^2 + 2u_{x_j, x_{i,j}} \quad (12)$$

The last term is the covariance between the measures $x_{i,j}$ and the quantity $x_j = EAL - h_j$. Since in EAL all the clocks are included, the measure $x_{i,j}$ will produce a nonzero covariance term by coupling to the same measures $x_{i,j,}$

which enters in EAL definition as many times as are the clocks inside or behind the lab *i*. We obtain:

$$u_{x_j,x_{i,j}} = u_{\left(\sum_{\ell=1}^{N} w_\ell (h'_\ell - x_{\ell,j}), x_{i,j}\right)} = -u_{\left(\sum_{\ell=1}^{N} w_\ell x_{\ell,j}, x_{i,j}\right)} = \\ = -\sum_{\ell=1}^{N_{eq\ lab_i}} w_\ell u^2_{x_{i,j}} = -W_{eq_i} u^2_{x_{i,j}} \quad (13)$$

where $N_{eq\ lab_i}$ is the equivalent number of the clocks in the laboratory *i*, also including the clock external to that lab but that are connected to the UTC through the lab *i*. We call equivalent weight W_{eq_i} the sum of the $N_{eq\ lab_i}$ clock weights. This will be better explained in the next section, where examples are given. In this case, equation (12) becomes:

$$u^2_{x_i} = u^2_{x_j} + u^2_{x_{i,j}} - 2\ u^2_{x_{i,j}} W_{eq_i} . \quad (14)$$

Equation (14) gives the uncertainty of each clock H_i, given the uncertainty of any clock H_j and the uncertainties of the chains of measures linking clock H_i to clock H_j.

2.3 Correlations and anticorrelations

In this work we have considered the measures $x_{i,j}$ affected by an uncertainty that we named "uncertainty of the link," which is the one reported in BIPM *Circular T* (Sec. 6). Any time a link $x_{i,j}$ appears in a multiple link such as $x_{i,k} = x_{i,j} + x_{j,k,}$ it gives raise to a correlation term with respect to $x_{i,j,}$ and this is the meaning of the last covariance term in (10).

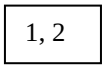
A more refined evaluation should consider that in case of TWSTFT, GPS/Glonass CV, or Melting Pot, also termed or All-in-View [9,10]; the uncertainty of the link is due to different contributions whose correlation properties may differ. For example, in same cases the uncertainty is dominated by systematic calibration uncertainties between lab *i* and lab *j*, while in other cases the link uncertainty reported in BIPM *Circular T* is mostly due to the noise of the receivers in both sites.

Correlations will always occur in situations wherein the same receiver or system is used to link between two different external labs. The analysis of these effects requires more details than are readily available about correlation of the links. Further study is in progress and a preliminary evaluation indicates that the final uncertainty would change by a very small extent. Some details are reported in the last section devoted to future work. In this work, we will assume that the noise affecting different non-overlapping links is uncorrelated, and that correlation will appear only when the same intermediate link appears in multiple links, as will be shown in the example below.

2.4 Examples of application of the method

In order to illustrate the applied procedure, we provide some illustrative examples. In every example, one can see that the uncertainty of each clock H_k with respect to UTC increases as the noise of the lab *k* link increases, but also decreases as the weight of the clock increases.

Case A1: Let's consider only two clocks (labelled 1 and 2) in the same lab *k* with no external links, as depicted in the sketch below:



k

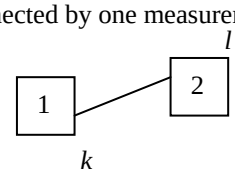
The ensemble time similar to UTC would be computed with these two clocks, and let's suppose clock 1 realizes *UTC*(*k*). The uncertainty of x_k = [*UTC* – *UTC*(*k*)], by means of equation (11), would be:

$$u^2_{UTC-UTC(k)} = u^2_{UTC-h_1} = w_2^2\, u^2_{x_{1,2}} . \quad (15)$$

The uncertainty of the measure $x_{1,2}$ was considered negligible because the clocks are within the same laboratory. Thus

$$u^2_{UTC-UTC(k)} \approx 0 . \quad (16)$$

Case A2: Let's consider again two clocks (labelled 1 and 2), but maintained in two different labs (*k* and *l*, respectively) connected by one measurement link $x_{k,l}$:

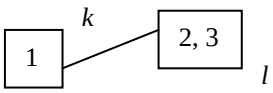


In this case, the uncertainty of the measurement link would be non-negligible and the final uncertainty for *UTC*(*k*), realized in lab *k* with clock number 1, would be

$$u^2_{UTC-UTC(k)} = w_2^2 u^2_{x_{l,k}} = (1 - w_1)^2 u^2_{x_{l,k}} . \quad (17)$$

From (17) we can see that increasing the weight of clock 1 from lab *k* decreases the uncertainty of [*UTC* – *UTC* (*k*)].

Case B1: Let's consider three clocks (labelled 1, 2, and 3), maintained in two different labs (*k* and *l*) connected by a measurement link $x_{k,l}$ according to the sketch below:

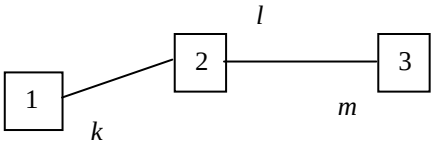


For site *k*, with *UTC*(*k*) realized by clock 1, the uncertainty on [*UTC* – *UTC* (*k*)] would be:

$$u^2_{UTC-UTC(k)} = w_2^2 u^2_{x_{2,k}} + w_3^2 u^2_{x_{3,k}} + 2w_2 w_3 u_{(x_{2,k}, x_{3,k})} \\ = (w_2 + w_3)^2 u^2_{x_{l,k}} = (1 - w_1)^2 u^2_{x_{l,k}} \quad (18)$$

since $u^2_{2,k} = u^2_{3,k} = u_{(x_{2,k}, x_{3,k})} = u^2_{l,k}$, because clocks 2 and 3 are in the same laboratory.

Case B2: Let's consider three clocks (labelled 1, 2, and 3), located in three different labs (*k*, *l*, and *m*). We are interested in the uncertainty of *UTC* (*l*) realized with clock 2 inside the lab *l*, and at first assume the links between lab (*l*,*k*) and lab (*l*,*m*) are uncorrelated:



Using (10), where the double summation is zero since the links are not correlated, we obtain,

$$u^2_{UTC-UTC(l)} = w_1^2 u^2_{x_{k,l}} + w_3^2 u^2_{x_{m,l}} \qquad (19)$$

For the case where the links between lab (*l*,*k*) and lab (*l*,*m*) were correlated through site-based noise in the receiver in lab *l*, but the total noise in both links was the same and equally contributed by the two labs, we would have $u_{(x_{l,k},x_{l,m})} = .5u^2_{l,k} = .5u^2_{l,m}$ and

$$\begin{aligned} u^2_{UTC-UTC(l)} &= w_1^2 u^2_{x_{l,k}} + w_3^2 u^2_{x_{l,m}} + 2w_1 w_3 u_{(x_{l,k},x_{l,m})} \\ &= (w_1^2 + w_1 w_3 + w_3^2)\ u^2_{x_{l,k}} \end{aligned} \qquad (20)$$

Case B3: Let's consider three clocks (labelled 1, 2, and 3), located in three labs (*k*, *l* and *m*), topology as for Case B2 above. The uncertainty is evaluated for lab *k* and lab *l* acts as "pivot" to connect lab *m* to lab *k*. The clocks in lab *m* are connected to lab *k* by a double link $x_{k,m} = x_{k,l} + x_{l,m}$. We can see that the noise of the link $x_{k,l}$ affects either the measures of clocks of lab *l* or clocks of lab *m*, and this noise correlates the clock estimates.

Assuming the two links $x_{k,l}$ and $x_{l,m}$ are uncorrelated we obtain, from (10):

$$\begin{aligned} u^2_{UTC-UTC(k)} &= w_2^2 u^2_{x_{l,k}} + w_3^2 \left[u^2_{x_{l,k}} + u^2_{x_{m,l}}\right] + \\ &+ 2w_2 w_3 u_{(x_{l,k},x_{m,k})} = (w_2 + w_3)^2 u^2_{x_{l,k}} + w_3^2 u^2_{x_{m,l}} \end{aligned} \qquad (21)$$

where the correlation between $x_{k,l}$ and $x_{k,m}$ is due the common path (*l*,*k*):

$$u_{(x_{l,k},x_{m,k})} = u^2_{x_{l,k}}\ . \qquad (22)$$

This example shows that, for uncorrelated links, the weight of labs behind a pivot lab is added to the weight of the pivot lab itself. We call "equivalent weight" W_{eq_i} the weight of the pivot laboratory *i*, which includes also the weights of the clocks behind that pivot lab. W_{eq_i} was also introduced in (13). The last term in (21) takes into account the noise from the pivot (lab *l*) to the remote lab *m*, weighted by the weight of clocks in lab *m*.

If we consider the case of links correlated through site-based noise, the equations leading to (21) would apply, and:

$$u^2_{UTC-UTC(k)} = (w_2^2 + w_2 w_3 + w_3^2)\ u^2_{x_{1,2}} \qquad (23)$$

For links dominated by site-based noise, this would be expected, since the dominant noise of that site implies that all ways to combine the external links would yield the same results.

2.5 Formulas for the UTC newtork

Starting from the example above, we evaluated the uncertainty corresponding to the network of links currently used for the computation of UTC.

USNO, NIST, NICT[2], and NTSC[3] act as intermediate pivots and that all the measurements are referenced in the end to the PTB laboratory.

Since the PTB plays a central role, we can use (10) to obtain the uncertainty for [*UTC* – *UTC*(PTB)] considering the "equivalent weights" for the intermediate pivot laboratories USNO, NIST, NICT, and NTSC. The final expression turns out to be:

$$\begin{aligned} u^2_{UTC-UTC(PTB)} &= \sum_{i=2}^{N_1+1} W_i^2 u^2_{x_{i,UTC(PTB)}} + \sum_{k=N_1+2}^{(N_1+N_2)+1} W_k^2 u^2_{x_{k,UTC(USNO)}} \\ &+ \sum_{k=(N_1+N_2)+2}^{(N_1+N_2+N_3)+1} W_k^2 u^2_{x_{k,UTC(NIST)}} + \sum_{k=(N_1+N_2+N_3)+2}^{(N_1+N_2+N_3+N_4)+1} W_k^2\ u^2_{x_{k,UTC(NICT)}} + \\ &+ W^2_{eq_{USNO}} u^2_{x_{UTC(PTB),UTC(USNO)}} + W^2_{eq_{NIST}} u^2_{x_{UTC(PTB),UTC(NIST)}} + \\ &+ W^2_{eq_{NICT}} u^2_{x_{UTC(PTB),UTC(NICT)}} + W^2_{eq_{NTSC}} u^2_{x_{UTC(PTB),UTC(NTSC)}} + \\ &+ W^2_{JACT} u^2_{x_{UTC(NTSC),UTC(JACT)}} \end{aligned} \qquad (24)$$

Here $N_1 = 25$ is the number of laboratories directly linked to PTB with one single link (IEN, AOS, SP, LT, etc.), excluding the four labs acting as intermediate pivots; $N_2 = 4$ is the number of laboratories linked to USNO (namely NRC, CNMP, APL, ONBA); $N_3 = 4$ is the number of laboratories linked to NIST (namely TCC, ONRJ, CNM, IGMA); $N_4 = 13$ is the number of laboratories linked to NICT (SG, TL, NAO, MSL, etc.); W_{eq_i} are "equivalent weights" of laboratories as introduced in (13).

The "equivalent weight" of the USNO is, for example:

$$W_{eq_{USNO}} = W_{NRC} + W_{CNMP} + W_{APL} + W_{ONBA} + W_{USNO}\ .$$

Similar expressions hold for the other intermediate pivots, NIST, NICT, and NTSC, and W_{lab} was defined in (11).

Once the uncertainty of [*UTC* – *UTC* (PTB)] is evaluated, the expression (14) can be used to compute the uncertainty of [*UTC* – *UTC* (*k*)] for every other laboratory *k*.

It has to be stressed that the results obtained by the application of equations (24) and (14) hold till all the other quantities involved are constant, in particular the clocks number and weights are to be constant and this is true only within the month of UTC evaluation. On a subsequent month, the clock number and weights may change leading to a different uncertainty evaluations.

3. A MONTE CARLO SIMULATION

The analytical approach presented in the previous section was tested by a Monte Carlo simulation using the BIPM's software ALGOS, on the full set of clock and time transfer data used to generate the *Circular T* of July 2004. For that

[2] NICT is the new name of CLR lab.

[3] All the acronyms appearing in this paper are in agreement with the list reported in BIPM Time Section Annual Reports.

month, the pivot laboratories were PTB, USNO, NIST, NICT, and NTSC.

Since the source of uncertainty in [*UTC* – *UTC* (*k*)] are the links whose uncertainty $u^2_{x_{i,j}}$ is listed in the BIPM *Circular T*, we performed a simulation by assuming that every link measure is described by a random variable with Gaussian distribution, mean value equal to the obtained measurement value *a*, and standard deviation equal to $u^2_{x_{i,j}}$.

The Monte Carlo simulation consists in picking up different values for the measure $x_{i,j}$ coming from its statistical distribution and computing [*UTC* – *UTC* (*k*)] with the simulated measure values. This gives an indication of the variability of the results [*UTC* – *UTC* (*k*)] due to the variability of the measures $x_{i,j}$, which are, in our assumption, the only source of variability and, hence, of uncertainty.

We proceed in the simulation step by step firstly evaluating the effect on [*UTC* – *UTC* (*k*)] of only one noisy link (all the other links are considered with negligible noise); then the noise is inserted link by link and we compare this obtained "experimental" variability with respect to the expected theoretical results coming from the analytical estimation (11).

3.1 One link

We began by studying the effect of the uncertainty of the link between the PTB and the LT labs using the data of July 2004. Laboratory LT has one clock of percentage weight $w_{LT} = 0.119$; the uncertainty of the link [*UTC* (LT) – *UTC* (PTB)] is $u^2_{UTC(LT)-UTC(PTB)} = 27.25\ ns^2$, and on MJD 53144 [*UTC* (LT) – *UTC* (PTB)] measurement value was equal to –242.8 ns. We generated a large number *N* of experimental values of [*UTC*(LT) – *UTC*(PTB)], normally distributed around the central value [*UTC* (LT) – *UTC*(PTB)] = –242.8 ns and with a variance of $u^2_{UTC(LT)-UTC(PTB)} = 27.25\ ns^2$.

- Theory

Using (11), with only the link (*PTB,LT)* corrupted with noise, we have

$$u^2_{UTC-UTC(PTB)} = w^2_{LT} u^2_{UTC(LT)-UTC(PTB)}$$
$$= \left(\frac{0.119}{100}\right)^2 (27.25) = 3.85x10^{-5}\ ns^2$$

Thus

$$u_{UTC-UTC(PTB)} = 0.0062\ ns\,,$$

and from (17)

$$u_{UTC-UTC(LT)} = 5.21\ ns\,.$$

Alternatively, using equation (11) directly referenced to *LT* lab, which correspond to having the total clock weight in PTB with the exception of the LT clock, we obtain the same theoretical result:

$$u_{UTC-UTC(LT)} = \left(1 - \frac{0.119}{100}\right) \cdot \sqrt{27.25} = 5.21\ ns\,.$$

- Experimental results

Using *N* = 10 000 simulated measurement values, we observe a normal distribution of the obtained values [*UTC* – *UTC* (PTB)] with a mean value equal to –16.7087 ns and standard deviation equal to 0.0063 ns; the value [*UTC* – *UTC* (PTB)] for that date published in *Circular T* was in fact equal to –16.7 ns. Therefore, the mean value and the standard deviation of [*UTC* – *UTC*(PTB)] obtained by simulating the noisy link *PTB,LT* correspond to the published value and to the expected theoretical uncertainty.

For the laboratory *LT* we obtained a normal distribution of mean value equal to 226.161 ns with a standard deviation of 5.2 ns. The published value in *Circular T* was [*UTC* – *UTC* (LT)] = 226.1 ns.

The same results are depicted in the figures below, where the histogram reports the obtained results from the Monte Carlo simulation, while the solid line (in red) represent the expected statistical distribution of the values according to the theoretical analysis.

The analytical and simulation results agree perfectly.

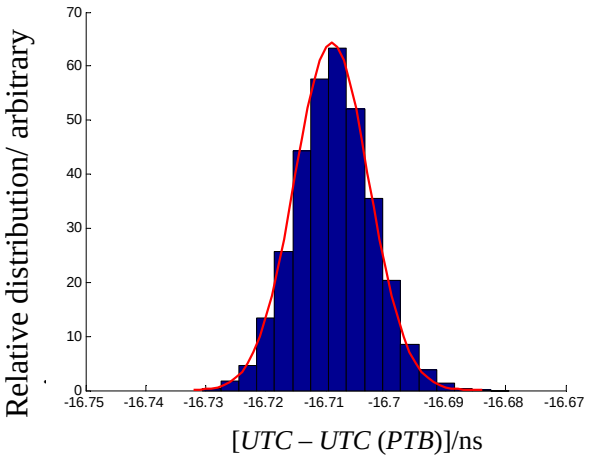


Figure 2. Theoretical and experimental distribution of the values [*UTC* – *UTC* (*PTB*)] considering only one link with noise.

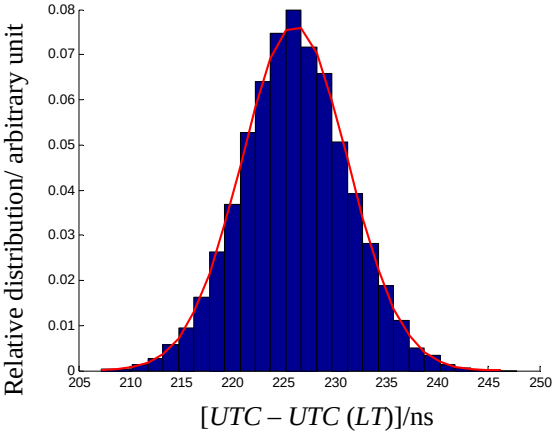


Figure 3. Theoretical and experimental distribution of the values [*UTC* – *UTC* (*LT*)] considering only one link with noise.

3.2 Complete system

The Monte Carlo results are here reported for the complete system of links, using N = 20 000 simulated data for any measure. Figure 4 shows the close agreement for example between the theoretical and experimental results for the values of [*UTC* – *UTC* (NIST)]. In Table 1, we can see the analytical compared to the simulation results obtained considering every links with noise. The agreement between the two estimation methods is quite impressing giving complete confidence on the analytical development.

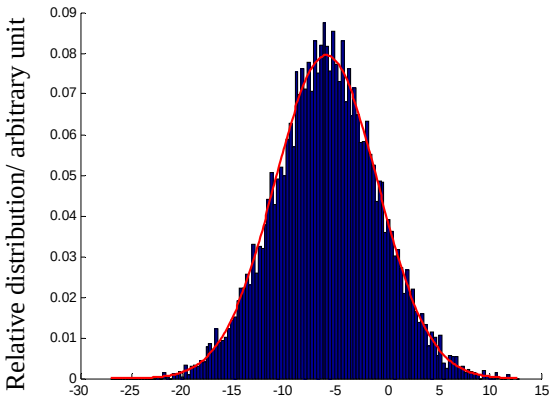


Figure 4. Theoretical and experimental distribution of the values of [*UTC* – *UTC* (NIST)].

Table 1. Analytical and numerically estimated uncertainties for July 2004 of all the UTC participating laboratories, considering every link with noise.

k	*u*[*UTC* – *UTC* (*k*)]/ns	
	Analytical method	Numerical method
AOS	5.6	5.6
APL	5.7	5.7
AUS	7.3	7.3
BEV	5.5	5.5
BIRM	20.6	20.6
CAO	21.2	21.0
CH	5.4	5.3
CNM	21.0	21.1
CNMP	8.4	8.4
CSIR	20.3	20.6
DLR	5.4	5.4
DTAG	10.6	10.5
HKO	7.3	7.3
IEN	2.4	2.4
IFAG	5.3	5.3
IGMA	20.8	20.9
INPL	10.9	11.0
JATC	21.2	21.3
JV	20.7	20.6
KRIS	7.1	7.2
LDS	20.3	20.2
LT	5.6	5.6
MSL	20.7	20.8
NAO	20.6	20.7
NICT	4.4	4.4
NIM	20.4	20.3
NIMT	20.7	20.9
NIS	20.3	20.6
NIST	5.0	5.0
NMC	20.7	20.8
NMIJ	7.0	7.0
NPL	5.3	5.2
NPLI	20.3	20.4
NRC	15.2	15.1
NTSC	7.1	7.0
OMH	20.2	20.5
ONBA	8.9	8.9
ONRJ	21.2	21.3
OP	5.4	5.4
ORB	5.3	5.4
PL	5.4	5.3
PTB	1.9	1.9
ROA	5.4	5.4
SCL	11.6	11.6
SG	20.5	20.6
SMU	20.7	20.7
SP	10.4	10.5
SU	6.1	6.1
TCC	21.2	20.9
TL	6.7	6.7
TP	5.8	5.8
UME	25.1	25.1
USNO	2.3	2.3
VSL	5.3	5.3

4. APPLICATION TO *CIRCULAR T* COMPUTATION

The formulas developed in Section 1.5 for UTC network have been applied to the link uncertainties as listed in Section 6 of BIPM *Circular T*. They are used operationally for the computation of *Circular T* since January 2005. Results for *Circular T* 205 of January 2005 are reported in Figure 5 below.

In the BIPM Circular T the last column values denoted by *u* correspond to the combined standard uncertainty of [*UTC* – *UTC* (*k*)] for any lab *k*. The values *uA* and *uB* are the components of the combined standard uncertainty *u* originating from type A and type B evaluations listed in Section 6, respectively. The following relationship holds:

$$u = \sqrt{uA^2 + uB^2}$$

Again, it has to be reminded that the application of equations (24) and (14) separately to the type A and type B evaluations is permitted only when all the other quantities involved are constant, in particular the clocks number and weights are to be constant and this is true only within the month of UTC evaluation.

5. CONCLUSIONS

This paper presents a preliminary study of the determination of the uncertainties of [*UTC – UTC* (*k*)]. An analytical solution is derived from the law of the propagation of uncertainty, taking into account that leap seconds and deterministic frequency steering of EAL do not affect these uncertainties. The analytical results were verified through Monte Carlo simulations using the software used to generate UTC, and a perfect agreement was found, giving complete confidence on the analytical estimation. A more detailed analysis is in progress, including full inclusion of all available calibration information, more details for the correlation between the links, methods for optimizing the link structure, given uncertainty information, non-Gaussian behaviour, and different correlation properties of uncertainties due to calibration or due to random noise.

REFERENCES

[1] W. Lewandowski, D. Matsakis, G. Panfilo, P. Tavella, "First evaluation and experimental results on the determination of [*UTC – UTC*(*k*)] uncertainties", 2004, *Proc. 36th PTTI*, Washington DC, pp. 247-261.

[2] C. Thomas, P. Wolf, and P. Tavella, 1994, *"Time Scales"*, BIPM Monograph 94/1.

[3] B. Guinot and C. Thomas, 1988, *"Establishment of International Atomic Time"*, Annual Report of the BIPM Time Section, 1.

[4] P. Tavella and C. Thomas, 1991, *"Comparative study of time scale algorithms"*, Metrologia, **28**, 57-63.

[5] C. Thomas and J. Azoubib, 1996, *"TAI computation: study of an alternative choice for implementing an upper limit of clock weights"*,Metrologia, **33**,227-240.

[6] J. Azoubib, *"A revised way of fixing an upper limit to clock weights in TAI computation,"* Working Document CCTF/01-14 of the 15th meeting of the CCTF (see www.bipm.org).

[7] *Guide to the Expression of Uncertainty in Measurement* (International Organization for Standardization, Switzerland), 1993.

[8] J. Azoubib and W. Lewandowski, 2003, *"Uncertainty of time links used for TAI,"* in Proceedings of the 34th Annual Precise Time and Time Interval (PTTI) Meeting, 3-5 December 2002, Reston, Virginia, USA (U.S. Naval Observatory, Washington, D.C.), pp. 413-424.

[9] M. Miranian and W. Klepczynski, 1991, *"Time Transfer via GPS at USNO,"* in Proceedings of the 4th International Technical Meeting of ION-GPS-91 (Institute of Navigation, Alexandria, Virginia), pp. 215-222.

[10] Z. Jiang and G. Petit, 2004, *"Applying GPS satellites all in view in Circular T,"* BIPM Time Section, Technical Memorandum TM132.

[11] J. Davis and G. de Jong, private communication.

[12] *GUM – Supplement 2: Models with any Number of Output Quantities* (Joint Committee for Guides in Metrology) (in progress).

[13] J. Hahn and P. Tavella, 2000, *"A time scale for satellite navigation systems: why and how?"* International Journal of Satellite Communication, 18, No. 4 & 5, 305-324.

CIRCULAR T 205 ISSN 1143-1393
2005 FEBRUARY 15, 16h UTC

BUREAU INTERNATIONAL DES POIDS ET MESURES
ORGANISATION INTERGOUVERNEMENTALE DE LA CONVENTION DU METRE
PAVILLON DE BRETEUIL F-92312 SEVRES CEDEX TEL. +33 1 45 07 70 70 FAX. +33 1 45 34 20 21 tai@bipm.org

1 - Coordinated Universal Time UTC and its local realizations UTC(k). Computed values of [UTC-UTC(k)].
From 1999 January 1, 0h UTC, TAI-UTC = 32 s.

Date 2004/05 0h UTC	DEC 30	JAN 4	JAN 9	JAN 14	JAN 19	JAN 24	JAN 29	Uncertainty/ns			Notes
MJD	53369	53374	53379	53384	53389	53394	53399				
Laboratory k				[UTC-UTC(k)]/ns				uA	uB	u	
AOS (Borowiec)	-21.2	-14.8	-18.7	-13.6	-11.2	-12.1	-17.9	1.6	5.4	5.6	
BEV (Wien)	84.9	87.5	90.2	90.2	94.8	96.9	99.9	1.5	5.3	5.5	
CH (Bern)	-27.4	-25.2	-27.4	-31.1	-28.6	-23.8	-21.1	0.8	5.3	5.4	
DLR (Oberpfaffenhofen)	1.4	-3.1	-7.8	-14.1	-19.1	-22.3	-31.3	0.8	5.4	5.5	
IEN (Torino)	-58.2	-71.8	-77.9	-79.8	-78.5	-84.6	-87.5	0.7	2.2	2.3	
LT (Vilnius)	486.0	519.1	547.4	549.6	528.4	533.6	542.7	1.6	5.3	5.5	
NICT (Tokyo)	4.5	-0.7	6.8	6.1	-1.7	1.1	3.9	1.2	4.3	4.5	
NIST (Boulder)	3.5	2.6	1.3	1.5	0.8	-0.6	-1.7	0.6	5.1	5.1	
NPL (Teddington)	8.0	11.2	14.6	18.5	22.7	25.7	33.7	0.7	2.2	2.3	
NTSC (Lintong)	17.7	20.4	26.7	31.1	27.4	19.6	19.1	2.7	6.5	7.0	
ONBA (Buenos Aires)	-923.0	-967.0	-1200.8	-1407.0	-1666.2	-3418.7	-3392.1	5.0	7.3	8.8	
ONRJ (Rio de Janeiro)	155.1	153.1	162.1	159.6	161.7	158.4	176.0	5.0	20.6	21.2	
OP (Paris)	41.7	45.8	49.4	43.6	40.3	39.8	36.4	0.6	2.1	2.2	
ORB (Bruxelles)	-42.9	-41.2	-39.3	-37.5	-35.2	-34.5	-32.3	0.8	5.3	5.4	
PL (Warszawa)	-3.9	-1.8	-5.7	-6.5	-4.5	-4.1	-7.8	1.5	5.2	5.4	
PTB (Braunschweig)	-0.6	1.2	2.5	3.0	5.5	5.9	4.7	0.4	1.9	1.9	
ROA (San Fernando)	-63.1	-62.0	-68.3	-67.9	-67.1	-65.8	-74.1	0.8	5.3	5.4	
UME (Gebze-Kocaeli)	599.8	595.1	593.2	603.0	608.3	611.2	612.5	15.0	20.1	25.1	
USNO (Washington DC)	-4.9	-4.7	-5.1	-4.6	-3.5	-3.5	-2.5	0.5	2.2	2.3	
VSL (Delft)	-76.9	-77.0	-77.9	-77.4	-56.4	-41.8	-21.2	0.7	2.2	2.3	

Figure 5: A sample of first section of BIPM *Circular T*.

THE DEVELOPMENT OF A KALMAN FILTER CLOCK PREDICTOR

J A Davis, C A Greenhall* and R Boudjemaa

National Physical Laboratory,
Hampton Road, Teddington, Middlesex, TW11 0LW

*Jet Propulsion Laboratory, California Institute of Technology,
4800 Oak Grove Dr. MS 298-100, Pasadena, CA 91109

ABSTRACT

A Kalman filter based clock predictor is developed, and its performance evaluated using both simulated and real data. The clock predictor is shown to possess a near to optimal Prediction Error Variance (PEV) when the underlying noise consists of one of the power law noise processes commonly encountered in time and frequency measurements. The predictor's performance in the presence of multiple noise processes is also examined. The relationship between the PEV obtained in the presence of multiple noise processes and those obtained for the individual component noise processes is examined. Comparisons are made with a simple linear clock predictor. The clock predictor is used to predict future values of the time offset between pairs of NPL's active hydrogen masers.

1. INTRODUCTION

National Measurement Institutes and Global Navigation Satellite System (GNSS) operators have both shown considerable interest in constructing atomic timescales that are maintained as close as possible to UTC. Because of the well-known latency in the publication of BIPM's Circular T, it is necessary to predict the future performance of an atomic timescale over periods of up to 50 days in order to maintain a local timescale that is close to UTC. Several authors have developed clock predictors. These range from the simple yet effective [1] to more elaborate designs for specific applications, e.g. GNSS [2].

Most clock predictors will provide a close to optimal prediction in the presence of a single well-specified noise process. The aim of the work described in this paper is to develop a clock predictor that will provide a close to optimal prediction in the presence of all noise processes commonly encountered in time and frequency measurements and in the presence of linear frequency drift. This is achieved by employing a two-stage analysis process. Initially, an adaptive Kalman filter is used to estimate the magnitude of the noise processes present within the clock measurements. The Kalman filter now possesses accurate noise parameters. This filter is then used to estimate the clock drift parameters, from which predictions of the future clock offsets may be obtained.

A description of the noise processes encountered in atomic clocks and the approach used in this paper to model them is described in section 2. The design of the Kalman filter is described in section 3. The clock predictor's performance in the presence of both single and multiple noise processes is discussed in sections 4 and 5, respectively. The relationship between the Prediction Error Variance (PEV) obtained in the presence of multiple noise processes and those obtained for individual component noise processes is examined. The Kalman filter predictor's performance is compared directly against that of a simple linear predictor. An example of the application of the clock predictor is described in section 6, and concerns the prediction of future values of the time offset between pairs of NPL's active hydrogen maser.

An adaptive Kalman filter method has been employed to estimate the magnitude of the noise parameters. The method is based on the work of Meyer [3], the technique being modified and updated so as to estimate the noise parameters used in the clock models used in this paper. A detailed account of the noise parameter estimation method will be published shortly.

2. CLOCK NOISE AND DRIFT PARAMETERS

The noise processes occurring in atomic clocks and the associated measurement systems are traditionally modelled as a linear combination of five well known power law noise processes [4], these being White Phase Modulation (WPM), Flicker Phase Modulation (FPM), White Frequency Modulation (WFM), Flicker Frequency Modulation (FFM), and Random Walk Frequency Modulation (RWFM). An exact model of either FPM or FFM cannot be incorporated into a Kalman filter.

In this paper we describe the noise processes intermediate in structure between WPM and WFM and between WFM and RWFM as a linear combination of Markov and integrated Markov noise processes, respectively [5]. This model enables us to construct a good approximation to all FPM and FFM power law

noise processes intermediate in character between WPM and RWFM. In addition, we are able to describe noise processes observed in atomic clocks that are not well modelled as a linear combination of power law noise processes but may be described well using the Markov and integrated Markov noise processes.

The deterministic characteristics of atomic clocks may be described in terms of time offset, frequency offset, and in many situations linear frequency drift offset, drift parameters. In addition the relaxation "memory" associated with both the FPM and FFM noise processes is described approximately as a finite number of Markov drift parameters.

3. DESIGN OF THE CLOCK PREDICTOR

3.1 Kalman filter design

A clock predictor is used to estimate the future offsets between two clocks, using both current and previous measurements of the time difference. The Kalman filter clock predictor provides estimates of the deterministic parameters of the offset along with the PEV.

The clock predictor is implemented using the standard Kalman filter equations [6]. The measurements are assumed to be noiseless.

The state vector at measurement n is given by x_n where

$$x_n = \left(\xi_n,\ \eta_n,\ \zeta_n,\ m_{1n}\ \Lambda\ m_{jn},\ p_{1n}\ \Lambda\ p_{kn}\right)^{\mathrm{T}}. \quad (3.1)$$

Here, ξ_n, η_n, and ζ_n are the time offset, frequency offset, and linear frequency drift offset components of the state vector, and m_{in} and p_{in} are the frequency and time offset of the ith Markov frequency and phase components respectively.

The design matrix H is given by

$$H = \left(1,\ 0,\ 0,\ 0\ \Lambda\ 0,\ 1\ \Lambda\ 1\right). \quad (3.2)$$

The state propagation matrix Φ and process covariance matrix Q are based on equations (3.3) and (3.4) of [5]. The matrices described in [5] are extended to include components that operate on the Markov phase components of the state vector. The off-diagonal elements of these additional rows and columns are zero, and the vectors describing the diagonal elements are given by

$$\Phi_p = \left(\exp(-R_{p1}\tau_0)\ \ \Lambda\ \ \exp(-R_{pk}\tau_0)\right)^{\mathrm{T}} \quad (3.4)$$

and

$$Q_m = \left(\tau_0 a_{22}(R_{p1}\tau_0)\sigma_{p1}^2\ \ \Lambda\ \ \tau_0 a_{22}(R_{pk}\tau_0)\sigma_{pk}^2\right)^{\mathrm{T}}, \quad (3.5)$$

respectively, where R_{pi} is the time constant of the ith Markov process, τ_0 is the spacing between successive measurements, and σ_{pi}^2 is the ith Markov noise parameter.

3.2 Initialisation

To initialise the Kalman filter, the updated parameter covariance matrix P^+ is set to P_0^+ where P_0^+ contains only diagonal elements described by the vector T_0 where

$$T_0 = \left(\infty,\ \infty,\ \infty,\ \frac{\sigma_{m1}^2}{2R_{m1}}\ \Lambda\ \frac{\sigma_{mj}^2}{2R_{mj}}\ \frac{\sigma_{p1}^2}{2R_{p1}}\ \Lambda\ \frac{\sigma_{pk}^2}{2R_{pk}}\right)^{\mathrm{T}}. \quad (3.6)$$

The $T_{0\xi}$, $T_{0\eta}$ and $T_{0\zeta}$ components are set to very large values, and the T_{0mi} and T_{0pi} components are set to the "steady state" variances of the Markov processes. All elements of the state vector estimates $\hat{x}_0^+$ are set to zero.

In the situation where we are confident that linear frequency drift is not present in the clock measurements, e.g., in the long term comparisons of two caesium fountain clocks, then the ζ linear frequency drift parameter component of the state vector may be removed or alternatively $T_{0\zeta}$ is set to zero.

3.3 Prediction

To obtain a prediction of the clock offsets and the corresponding uncertainties we need to run the Kalman filter beyond the last available measurement y_c. The state vector estimates $\hat{x}_\delta^-$, and parameter covariance matrix P_δ^- at prediction length δ are obtained using P_c^+ where

$$\hat{x}_\delta^- = \Phi(\delta)\hat{x}_c^+, \quad (3.7)$$

and

$$P_\delta^- = \Phi(\delta)P_c^+\Phi^{\mathrm{T}}(\delta) + Q(\delta). \quad (3.8)$$

The time offset prediction $\hat{y}_\delta$ and the associated PEV V_δ^K are then given by

$$\hat{y}_\delta = H\hat{x}_\delta^- \quad (3.9)$$

and

$$V_\delta^K = HP_\delta^- H^{\mathrm{T}}. \quad (3.10)$$

Combining equations (3.8) and (3.10) gives the following expression for the PEV:

$$V_\delta^K = H\Phi(\delta)P_c^+\Phi^{\mathrm{T}}(\delta)H^{\mathrm{T}} + HQ(\delta)H^{\mathrm{T}}. \quad (3.11)$$

The first term on the right of equation (3.11) depends on the uncertainty in the extrapolation of the current estimate of the clock offset drift parameters, while the second term depends completely upon the stochastic processes occurring within the clocks.

4. PERFORMANCE IN THE PRESENCE OF A SINGLE NOISE PROCESS

4.1 Comparison with simple linear and optimal predictors

We compare the performance of the Kalman filter clock predictor against that of a simple linear predictor and where possible against that of an optimal clock predictor. The minimum possible PEV V^M has been previously calculated in the case of several power law noise processes from the power spectral density of those noise processes [1, 7]. Where possible we relate the PEV to the Allan variance $\sigma_y^2(\delta)$ of the noise process.

4.2 WFM Noise

In the presence of only WFM noise, assuming an infinite length data set and known linear frequency drift parameter, both the Kalman filter and simple linear clock predictors provide optimal predictions. This prediction is obtained from two infinitely spaced data points resulting in exact knowledge of the frequency offset parameter. The PEV and $\sigma_y^2(\delta)$ are then related by

$$V_{WFM}^{M} = \delta^2 \sigma_y^2(\delta) = \delta \sigma_{WFM}^2 \quad . \qquad (4.1)$$

4.3 RWFM Noise

In the presence of RWFM noise and known linear frequency drift parameter, the optimal PEV is given by

$$V_{RWFM}^{M} = \delta^2 \sigma_y^2(\delta) = \frac{\delta^3}{3} \sigma_{RWFM}^2 \quad . \qquad (4.2)$$

The PEV of the Kalman filter predictor is given by

$$V_{RWFM} = \frac{\delta^3}{3} \sigma_{RWFM}^2 + \delta^2 \sigma_{RWFM}^2 \frac{\tau_0}{\sqrt{12}} \quad . \qquad (4.3)$$

The optimal PEV is only reached when we use infinitely small data spacing. This is due to the frequency offset parameter changing during the interval between the two measurements.

4.4 Single integrated Markov noise process

To study the performance of both the Kalman filter and simple linear predictor in the presence of FFM it is useful to determine the performance of these predictors in the presence of a single integrated Markov noise process. The relative performances of the two predictors are shown in figure 1. The linear predictor is optimised at each prediction length δ.

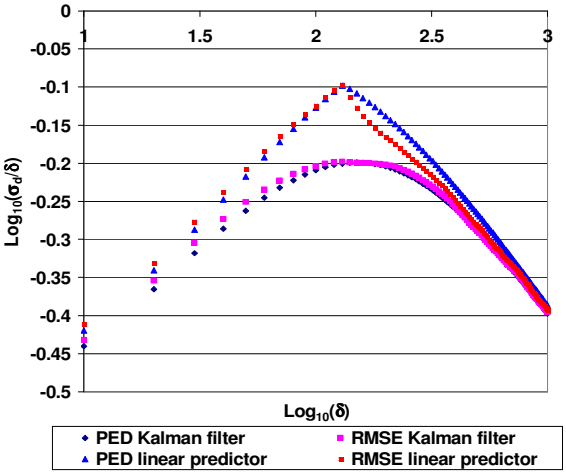


Figure 1 shows plots of $Log_{10}(\sigma_\delta / \delta)$, against $Log_{10}(\delta)$, obtained from the Kalman filter predictor (black and pink), and the simple linear predictor (blue and red), obtained both from theory (PED) and from simulations (RMSE) (2000 estimates).

Graphs of $Log_{10}(\sigma_\delta / \delta)$ against $Log_{10}(\delta)$ are plotted in figure 1, where σ_δ is the prediction error deviation. The results are presented in this form so that the gradients of the curves would be similar to, and hence easily compared with, graphs of $Log_{10}(\sigma_y(\delta))$ against $Log_{10}(\delta)$. The Prediction Error Deviation (PED) values obtained from theory and Room Mean Square Error (RSME) obtained from simulation agreed well.

A single Markov process is considered where R = 1.18×10^{-2} s^{-1}, $\tau_0 = 10$ s, and the data length is ten times the maximum prediction length. The PEV was found to be lower in the case of the Kalman filter predictor at the "centre" of the Markov process. At long and short prediction lengths where the noise process resembles WFM and RWFM, respectively, the PEV values obtained from the two predictors were found to be similar.

5. PERFORMANCE IN THE PRESENCE OF MULTIPLE NOISE PROCESS

5.1 PEV inequality

The following inequality applies to the PEV obtained for a fixed when multiple noise processes are present:

$$V_{A+B}^{M} \geq V_A^M + V_B^M \qquad (5.1)$$

where V_{A+B}^M is the minimum PEV for the sum $x_A(t)+x_B(t)$ of two independent noise processes, and V_A^M and V_B^M are the minimum PEV values obtained individually.

In previous literature [1, 7] the above inequality has been assumed to be an equality. While this may in some situations provide a reasonable approximation it may lead in many examples to significant errors in the PEV estimation.

To understand the origin of the above inequality we examine further equation (3.11). The second term on the right hand side of equation (3.11)

$$V^S = HQ(\delta)H^{\mathrm{T}} \qquad (5.2)$$

is the stochastic component of the PEV, its origin is due to noise occurring within the clock *after* the last measurement has occurred. When two or more noise processes are present this component of the PEV will add linearly.

The first term on the right hand side of equation (3.11)

$$V^D = H\Phi(\delta)P_c^+\Phi(\delta)^{\mathrm{T}}H^{\mathrm{T}} \qquad (5.3)$$

gives the contribution to the Kalman filter PEV from uncertainties in the extrapolation of the drift parameters. The PEV inequality occurs within this "parameter extrapolation" term. When only the individual noise processes are present, the individual PEV values V_A and V_B are obtained using optimal values of the Kalman gain, K_A and K_B respectively. When we compute the PEV V_{A+B} from the combined noise processes we are using a different value K_{A+B} for the Kalman gain. In general the K_{A+B} value of the Kalman gain will not be the optimal Kalman gain when computing the PEV for either component noise process. Hence the resulting PEV V_{A+B} will be greater than the sum of the individual V_A and V_B PEV values and so leads to the inequality. If we used the same Kalman gain in all three cases the inequality would not occur, however the PEV values would then in most situations be sub-optimal.

5.2 Combination of WFM and RWFM noise processes

Consider the following example where there is a mixture of WFM and RWFM noise processes present with magnitudes σ_{WFM}^2 and σ_{RWFM}^2. We assume an infinite data set so that the Kalman filter has reached a steady state, and that the linear frequency drift term is either known or absent.

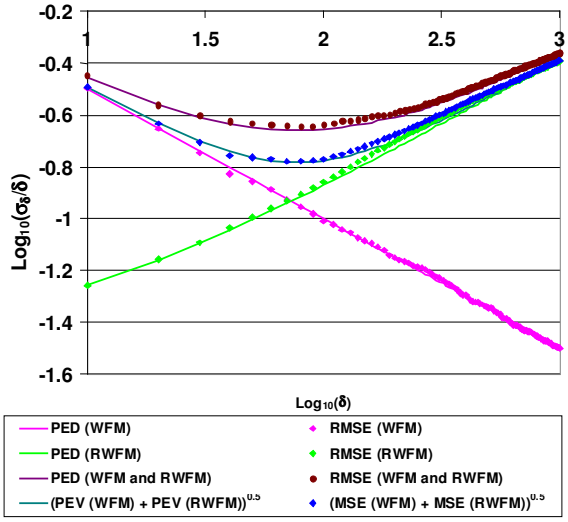


Figure 2 slowing plots of $Log_{10}(\sigma_\delta/\delta)$ against $Log_{10}(\delta)$, obtained from the Kalman filter parameter covariance matrix (PED) and from simulation (RMSE) (500 estimates), in the case of WFM (pink), RWFM (green), a combination of WFM and RWFM (brown) and the sum of the individual noise process PEV values (blue).

The parameter covariance matrix then propagates according to the steady state Kalman filter equations

$$P^- = \Phi P^+ \Phi^{\mathrm{T}} + Q \qquad (5.4a)$$

$$K = P^- H^{\mathrm{T}}(HP^- H^{\mathrm{T}})^{-1} \qquad (5.4b)$$

$$P^+ = (I - KH)P^- \qquad (5.4c)$$

where

$$\Phi = \begin{pmatrix} 1 & \tau_0 \\ 0 & 1 \end{pmatrix}, \qquad (5.5)$$

and [8]

$$Q = \begin{pmatrix} \tau_0\sigma_{WFM}^2 + \dfrac{\tau_0^3\sigma_{RWFM}^2}{3} & \dfrac{\tau_0^2\sigma_{RWFM}^2}{2} \\ \dfrac{\tau_0^2\sigma_{RWFM}^2}{2} & \tau_0\sigma_{RWFM}^2 \end{pmatrix}, \qquad (5.6)$$

$H = (1 \quad 0)$, and because the measurements are noiseless

$$P^+ = \begin{pmatrix} 0 & 0 \\ 0 & p \end{pmatrix}. \qquad (5.7)$$

Solving the above simultaneous equations (5.4) for p and then substituting into equation (3.11) gives

$$V_\delta^K = \delta\sigma_{WFM}^2 + \frac{\delta^3}{3}\sigma_{RWFM}^2 + \delta^2\sigma_{RWFM}(\sigma_{WFM}^2 + \frac{\tau_0^2}{12}\sigma_{RWFM}^2)^{1/2}. \quad (5.8)$$

The right hand term of (5.8) contains terms that include the multiplication of σ_{WFM} and σ_{RWFM}, the inequality given in equation (5.1) is clearly evident.

Comparing the brown and blue curves of figure 2 we observe that the PEV equality is clearly evident with the PEV values obtained from the combined noise processes (brown) possessing the higher PEV values.

5.3 Performance in the presence of FFM

FFM noise is present in active hydrogen masers. The optimal PEV obtained in the presence of a FFM noise process has been previously calculated [1, 7] as

$$V = \frac{\sigma_y^2\delta^2}{\ln(2)}. \quad (5.9)$$

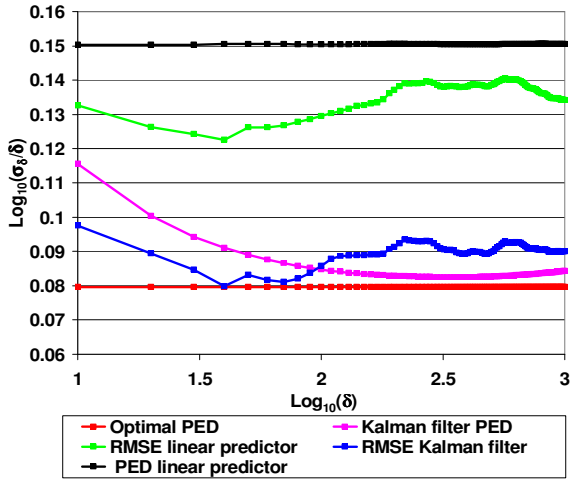


Figure 3 slowing plots of $Log_{10}(\sigma_\delta/\delta)$ against $Log_{10}(\delta)$ obtained from the Kalman filter predictor (blue and pink), simple linear predictor (green and black), and from an optimal predictor (red); obtained both from theory (PED) and from simulations (RMSE) (2000 estimates).

The performance of both the Kalman filter and simple linear clock predictors in the presence of FFM and in the absence of linear frequency drift is shown in figure 3. These plots were obtained using simulated data and a data set that is ten times the maximum prediction length. The performance of the Kalman filter clock predictor is close to that of an optimal predictor, except at prediction lengths that are only a small multiple of the minimum data spacing τ_0. In contrast, the simple linear predictor does not provide an optimal prediction. The best linear prediction was found to occur where the predictor point spacing was equivalent to the prediction length.

6. PREDICTOR EVALUATION USING REAL CLOCK AND TIME TRANSFER DATA

The performances of both the Kalman filter and simple linear predictors have been examined using 150 days of clock measurements made between NPL's three active hydrogen masers, with minimal spacing τ_0 of 600 s. The three pairs of clock differences were examined separately. The magnitude of the noise parameters were estimated as outlined in section 6 assuming that the clock noise may be described as a linear combination of integrated Markov noise processes. Clock predictions were made throughout the length of the data sets, and these were then compared against the actual clock data. The method was repeated using a simple linear predictor with optimal "point spacing".

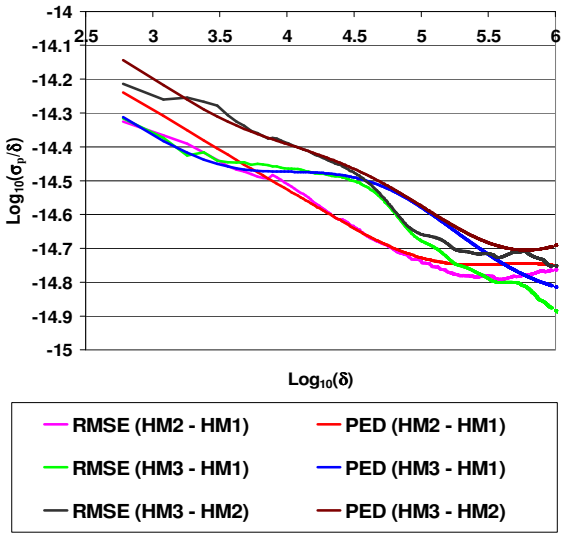


Figure 4 showing plots of $Log_{10}(\sigma_d/\delta)$ against $Log_{10}(\delta)$ for all three pairs of maser differences obtained from theory (PED) and simulation (RMSE) using the Kalman filter predictor.

Figure 4 plots of $Log_{10}(\sigma_d/\delta)$ against $Log_{10}(\delta)$ show that the RMSE obtained from the clock data and the PED estimates obtained from the Kalman filter agreed well at all prediction lengths. This helps confirm the validity of our model.

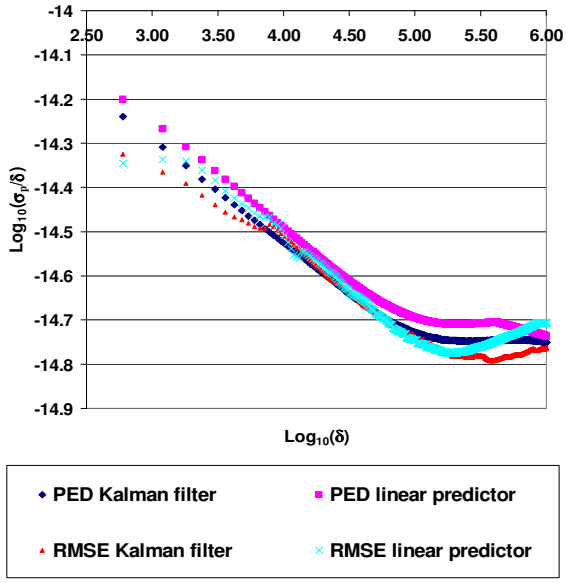


Figure 5 showing plots of $Log_{10}(\sigma_y/\delta)$ against $Log_{10}(\delta)$ for the (HM2 –HM1) maser comparisons obtained from theory (PED) and simulation (RMSE) using both a Kalman filter and simple linear predictor.

Figure 5 shows the variation of PEV against prediction length for both the Kalman filter and simple linear predictor. The PEV and the root mean square prediction error agree well for both the Kalman filter and simple linear clock predictors. At all prediction lengths the theoretical PEV values are lower in the case of the Kalman filter clock predictor. The root mean square prediction errors obtained from the Kalman filter predictor are lower than those obtained from the linear predictor at most prediction lengths. At prediction lengths close to one day we observe that the root mean square prediction error obtained from the simple linear predictor is very slightly lower. One possible cause may be diurnal changes occurring within NPL's clock rooms, this may result in un-modelled periodic like effects perturbing the Kalman filter predictor.

8. CONCLUSIONS

A Kalman filter clock predictor has been developed. The performance of this predictor was found to be very close to optimal when operating on noise processes where the optimal PEV was known. The Kalman filter clock predictor performs significantly better than a simple linear predictor, in the presence of a single integrated Markov noise process, FFM noise and when used on NPL's hydrogen maser clock data.

The relationship between the PEV obtained in the presence of a single noise process and multiple noise process has been examined, and the magnitude of the PEV "inequality" examined.

9. ACKNOWLEDGEMENTS

The authors thank P W Stacey, R Hlavac and P B Whibberley for the provision of NPL's time scale data. This work has been supported by the Time and Frequency programme of the National Measurement System Policy Unit of the Department of Trade and Industry, UK. Part of the work was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

REFERENCES

[1] L-G Bernier, "Use of the Allan Deviation and Linear Prediction for the Determination of the Uncertainty on Time Calibrations Against Predicted Timescales". IEEE Trans. Instrumentation & Measurement, **52**, no.2, 2003, p.483-486.

[2] F Vernotte, Delporte J, M Brunet, and T Tournier, "Uncertainties of drift coefficients and extrapolation errors: Application to clock error prediction". Metrologia 2001, **38**, no 4, 325-342

[3] R K Mehra, "On the identification of variances and adaptive Kalman filtering". IEEE transactions on automatic control, **AC-15**, no 2, April 1970.

[4] Handbook: selection and use of precise frequency and time systems, Chapter 3, International Telecommunication Union, 1997.

[5] J A Davis, C A Greenhall, and P W Stacey, "A Kalman filter clock algorithm for use in the presence of flicker frequency modulation noise", Metrologia **42,** 1-10, 2005.

[6] R G Brown and Y C Hwang, "Introduction to random signals and applied Kalman filtering" third edition, John Wiley and Sons, 1997, chapter 5.

[7] P Tavella and M Gotta, "Uncertainty and prediction of clock errors in space and ground applications". Proceedings of the 14th EFTF, 2000, pp 77-81 Turin.

[8] S T Hutsell, "Relating the Hadamard variance to MSC Kalman filter clock estimation". Proceedings of the 14th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 1995, pp 291-301.

UTC(SU) STEERING TIME SCALE STRATEGY DATA STIMULATION AND FIRST RESULTS

N. Koshelyaevsky, E. Zagirova

Institute of Metrology for Time and Space FGUP «VNIIFTRI», Russia, Mendeleevo, 141570,
nkoshelyaevsky@imvp.ru

Keywords: Time scale, Time algorithm

1 Introduction

During the whole history of atomic timing in Russia the National Universal Coordinated Time scale UTC(SU) have been generated quite autonomously, except for leap second. It means that no continuos steering correction were introduced in it. Naturally as far as time links became more accurate and requirements on synchronisation of UTC(k) relative to UTC become more and more insistent [1, 2] some once-only correction have been introduced into TA(SU) and UTC(SU) [3, 4].

Statistical analysis of the clocks contributing to the UTC(SU) and data published in the BIPM Circular T for the period of last years shows that UTC(SU) performances can be considerably improved for sample interval starting from few months in case of properly introduced steering corrections.

This paper presents proposals on possible realization of steering time scale, based on Phase-Lock Loop ideas, steering algorithm, estimations of time scale steering ability based on data stimulation, and first experimental results.

2 Current status and necessity of steering time scale

It seems obvious that sooner or later any autonomous time scale will gain any preassigned time error compare to reference one. The problem is elapsed time only. Statistical analysis of the data for UTC – UTC(SU) published in the BIPM Circular T for the period of last few years shows, Fig .1, that existing UTC(SU) gains about 100 ns for one year. This value is within CCDS Recommendation S5 (1993) [2]. On other hand time link between UTC and UTC(SU) and stability performances of UTC(SU) itself estimated basing on official publications of the BIPM Time Section [5] enables maintain |UTC-UTC(SU)| difference within much more less limits in case of corrective actions introduced into UTC(SU). The achievable level of time uncertainty will mainly depend on UTC(SU) stability in autonomous mode and a time elapsed between steering corrections. Because of UTC itself is a postponed time scale, delayed about half a month regarding the end of previous interval the minimal interval to apply steering corrections will be next month compare to current. More over, because of one need to have a proper time interval to estimate a value of steering corrections and this interval have to be at least not shorter than prediction period – one month – the total time step will be at least about three months. Referring to Fig.1 that corresponds to theoretical limit of gained time error about 10 ns.

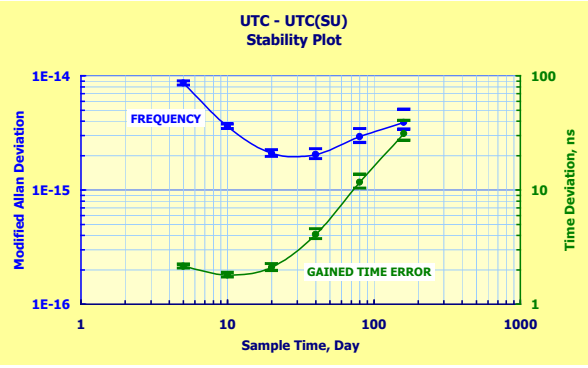


Fig .1

3 Possible realization and proposed algorithm

Possible realization of steering UTC(SU) system is considerably blown by Phase-Lock Loop ideas [6]. In this case one get phase raw difference information (or time difference) from BIPM Time Section publication (PD – Phase Detector), process it (Filter) and apply a steering correction (frequency step) to the UTC(SU) (VCO – Voltage Controlled Oscillator in PLL terminology) in a way to minimise UTC – UTC(SU) time difference. In contrast to usual analogous PLL the problem is that control correction (steering correction) is considerably delayed (is applied every other step).

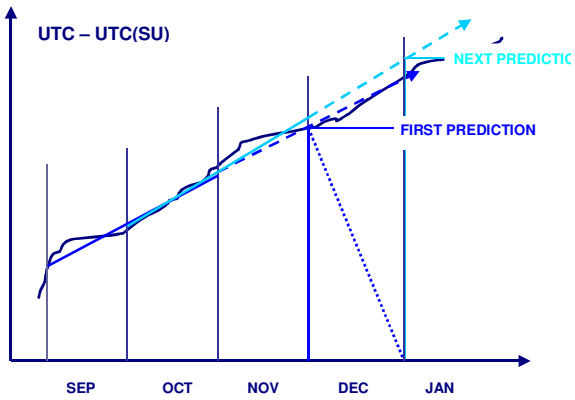


Fig .2

Let us suppose that we are living in the middle of November, Fig.2. At that period we get UTC-UTC(SU) data from BIPM ending the October. Basing on two months data block one may predict UTC-UTC(SU) for the December and foresee steering correction of such a value that compensate the expected time difference at the end of December. There will be no problem with next steering correction which have to be estimated in the middle of November and be applied in January, because of data for November have not been affected by corrections. In contrast to these two steps each next evaluation have to take into account previously applied steering correction. The example of steering actions is demonstrated for time slew dependence, Fig. 3.

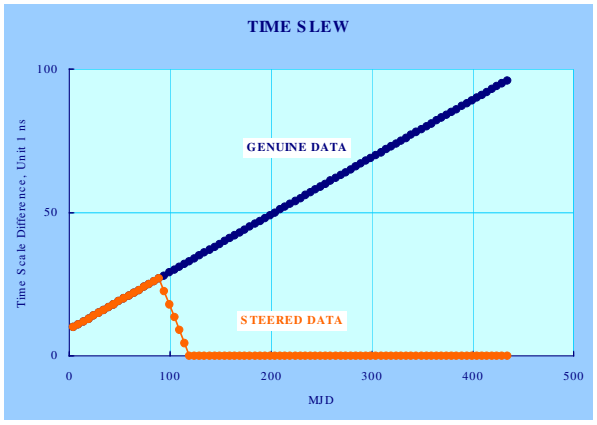


Fig .3

4 Data stimulation and preliminary accuracy estimations

First of all we tried to estimate possible uncertainty level for steering time scale basing on data stimulation. We have tested two stimulated data sequences. The instability of fist may be expressed as the quadratic sum of three components with τ in 5 days:

a white frequency noise of $1\times10^{-15}\times\tau^{-1/2}$

a flicker frequency noise of 1×10^{-15}

and a random walk frequency noise of $1\times10^{-15}\times\tau^{1/2}$

The other data stream apart from first sequence included term corresponding to possible frequency drift $15\times10^{-17}\times\tau$. It seems interesting because of our clock ensemble consists of H-masers.

The data stimulation results for source and steering time scales and corresponding estimations are presented on Fig.4 and 5. Some commentaries. On Fig.4 there are two realizations of the steering time scales. The upper one corresponds to the mode when steering correction is applied to the source scale without any analysis – one gets correction value and apply it immediately. The lower, less dense curve corresponds to the mode when along with steering correction value one estimates the uncertainty of it and introduce correction when value of estimated corrections exceed uncertainty at least twice – approximately 95% probability. Despite that in latter case corrections have been applied not at every step integral steering result looks at least twice more

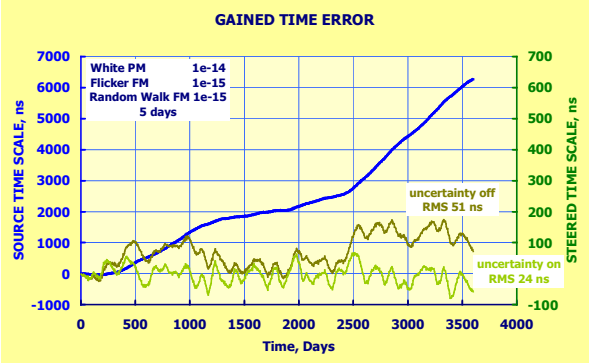


Fig .4

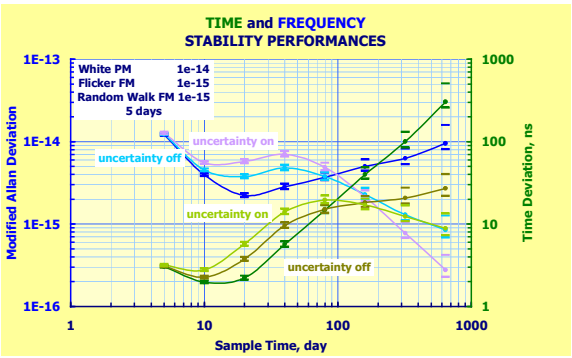


Fig .5

better. On other hand looking at Fig.5 one may see that first mode of operation does not solve steering problem completely – steering scale continues gain uncertainty - strictly speaking with much more lower rate compare to the source.

The results for other data stream which includes frequency drift are presented at two next pictures, Fig. 6 and 7.

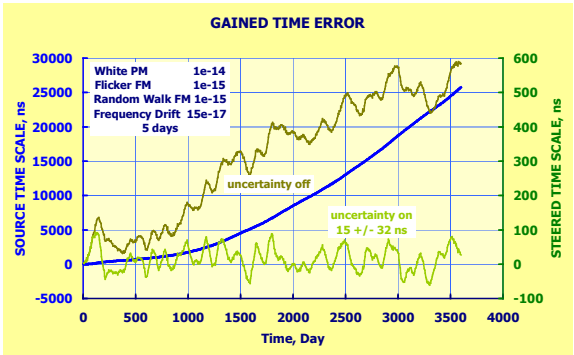


Fig. 6

There is a qualitative difference between steering behaviour in presence and absence of frequency drift, Fig.4 and Fig.6. And this difference became more obvious when one doesn't control application of frequency correction steps.

Even in case of controlled frequency correction steps the steering time scale is somewhat biased (15±32 ns) and its integral uncertainty is more then without drift. The problem to our understanding originates from second order PLL, which can't track frequency drift. So, our further steering filter stimulation have to include into prediction model possible frequency drift.

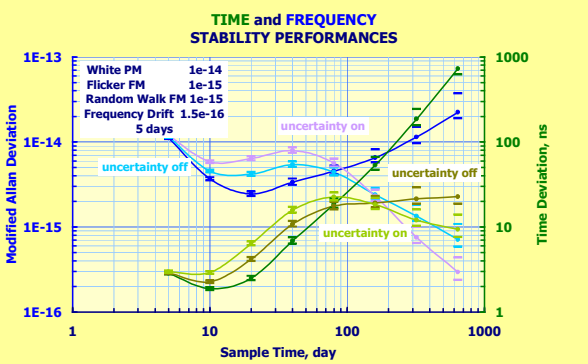


Fig. 7

As a next step of evaluation of possible improvement in steering time scale we tried to estimate what would one get if applied the proposed algorithm to real UTC-UTC(SU) scale difference starting beginning of 1999, MJD ~ 51200. Fig. 8 depicts the obvious improvement more two times from more than 50 ns RMS to 21 ns.

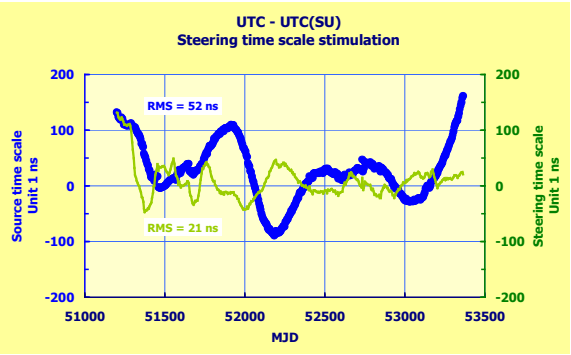


Fig. 8

5 First step in practical realization of steering UTC(SU) time scale

According to notification issued and distributed between registered users of the UTC(SU) time scale [7] realization of UTC(SU) have been changed on MJD 53369. These changes in realization have been accompanied by 160 ns time step and 1×10^{-14} frequency step. These steps have been introduced to UTC(SU) only, TA(SU) realization is still previous. In February the first steering frequency step have been introduced to the UTC(SU) and all this history is depicted on Fig. 9. Due to some internal limitations we could not zero in frequency difference, that's why we expect to wait up to April

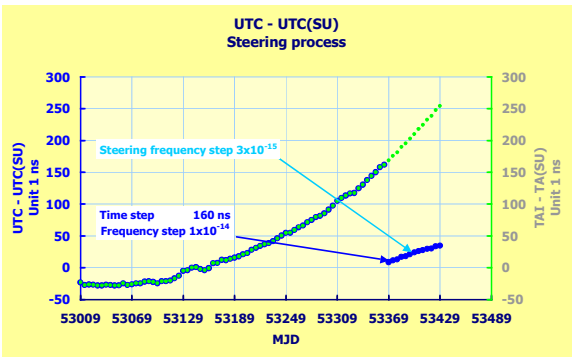


Fig. 9

to cancel frequency difference between UTC and UTC(SU) and start to decrease the corresponding time difference.

6 Conclusion and further improvements

The proposed algorithm seems to work and considerably decreases time deviation of steered time scale regarding to reference one. The rest part of year will show us in practice the achievable uncertainty. On other hand in a full agreement to control theory frequency stability of the steered time scale is somewhat worth compare to original one, at least for sample time comparable to prediction and feedback time constant.

So it is obvious that as a next step of steering time scale behavior study we are going to investigate some ideas to decrease steering time scale stability degradation and to improve tracking ability.

7 Acknowledgement

Authors would like to expess their gratitude to the EFTF Scientific Committee for support to participate in the Forum and deliver this paper.

8 References

[1] Synchronization of UTC(k) and UTC, Recommendation S1 (1989), Comité Consultatif pour la Définition de la Seconde, Report of the 11th Meeting, 1989.

[2] Need to improve worldwide time coordination to UTC Recommendation S5 (1993), Comité Consultatif pour la Définition de la Seconde, Report of the 12th Meeting, 1993.

[3] Annual Report of the BIPM Time Section, Volume 8, 1995, p. 54, Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92312 Sevres Cedex.

[4] Annual Report of the BIPM Time Section, Volume 9, 1996, p. 61, Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92312 Sevres Cedex.
[5] Circular T ISSN 1143 - 1393, 2004 March 11, Bureau International des Poids et Mesures, Organisation Intergovernementale de la Convention du Metre, Pavillon de Breteuil F- 92312 Sevres CEDEX, tai@ bipm. org
[6] Gardner, F. M. (1979) Phaselock Techniques (Wiley, New York).
[7]Федеральное агентство по техническому регулированию и метрологии, Институт метрологии времени и пространства ФГУП «ВНИИФТРИ», ISSN 0135-2415, БЮЛЛЕТЕНЬ Г- 78, 20 ДЕКАБРЯ 2004 г., И З В Е Щ Е Н И Е, Об изменении реализации шкалы времени UTC(SU)

Session 5C / Time and frequency transfers

TIME AND FREQUENCY COMPARISONS BETWEEN FOUR EUROPEAN TIMING INSTITUTES AND NIST USING MULTIPLE TECHNIQUES

A. Bauch[(1)], J. Achkar[(2)], R. Dach[(3)], R. Hlavac[(4)], L. Lorini[(5)], T. Parker[(6)], G. Petit[(7)], and P. Uhrich[(2)]
(1) Physikalisch-Technische Bundesanstalt Braunschweig, Germany
(2) Laboratoire National de Métrologie et d'Essais – Observatoire de Paris / Systèmes de Référence Temps Espace, Paris, France
(3) Astronomical Institute, University of Bern, Switzerland
(4) National Physical Laboratory, Teddington, UK
(5) Istituto Elettrotecnico Nazionale Galileo Ferraris, Torino, Italy
(6) National Institute of Standards and Technology, Boulder, Colorado, USA
(7) Bureau International des Poids et Mesures, Sèvres, France
Corresponding Author: A. Bauch, PTB, Bundesallee 100, 38116 Braunschweig, Germany, e-mail: andreas.bauch@ptb.de

ABSTRACT

Istituto Elettrotecnico Nazionale Galileo Ferraris (IEN), National Institute of Standards and Technology (NIST), National Physical Laboratory (NPL), Laboratoire National de Métrologie et d'Essais – Observatoire de Paris / Systèmes de Référence Temps Espace (OP), and Physikalisch-Technische Bundesanstalt (PTB) performed an intense comparison campaign of selected frequency references maintained in their laboratories during about 25 days in October/November 2004. The active hydrogen maser reference standards used at all laboratories served at the same time as frequency references for the institutes' fountain frequency standards. The fountains of IEN, NPL, and OP were operated for most of the time during the same period, and results of the fountain comparisons are given in an accompanying paper. Three techniques of frequency (and time) comparisons were employed. Going beyond the standard procedures, two-way satellite time and frequency transfer was performed in an intensified measurement schedule of 12 equally spaced measurements per day. The data of dual frequency geodetic GPS receivers were processed to yield an ionosphere-free linear combination of the code observations from both GPS frequencies, typically referred to as TAI P3 analysis. Last but not least, the same GPS raw data were separately processed allowing GPS carrier-phase based frequency comparisons to be made. These showed the lowest instability at short averaging times of all of the methods. The instability was at the level of 1 part in 10^{15} at one-day averaging time, similar to what was achieved with TWSTFT. The GPS TAI P3 analysis is capable of giving a similar quality of data provided that averaging over two days or longer is acceptable. The analysis presented here shows that all techniques provide the same mean frequency difference between the standards involved within the measurement uncertainty of a few parts in 10^{16}.

1. INTRODUCTION

The study of time and frequency transfer is an important sector of time and frequency metrology in general since it is essential for the wide application of state-of-the-art frequency standards. In recent years, research into primary frequency standards has led to several devices whose uncertainty to realize the SI second is at the level of one part in 10^{15} or below and whose frequency instability is good enough to verify this accuracy during averaging times of one day or even below [1-5]. Data from these clocks have been regularly used by the Bureau International des Poids et Mesures (BIPM) as an input for the realization of International Atomic Time (TAI) and of Coordinated Universal Time (UTC). Generally speaking, two different methods of time transfer have traditionally been used by BIPM for maintaining the network of participating laboratories, Global Positioning System (GPS) common view (CV) and two-way satellite time and frequency transfer (TWSTFT) via geostationary satellites [6, 7]. Classical GPS C/A code analysis is not discussed in this article since it does not provide the required measurement resolution. A step forward has been the use of geodetic GPS receivers providing dual frequency code observables which allow the so-called TAI P3 analysis to be made [8]. The use of GPS phase observables for frequency transfer has been proposed for similar purposes [9, 10] and has already been used to compare the fountain clocks CSF1 of the Physikalisch-Technische Bundesanstalt, Germany (PTB), and F1 of the National Institute of Standards and Technology, USA (NIST) [11].

Three further institutes, the UK National Physical Laboratory (NPL), the Italian Istituto Elettrotecnico Nazionale Galileo Ferraris (IEN), and the French Laboratoire National de Métrologie et d'Essais – Observatoire de Paris / Systèmes de Référence Temps Espace (abbreviated as OP throughout the paper), agreed with NIST and PTB on a campaign of comparisons among their fountain type primary frequency standards, to be effected during 20 to 25

Tab.1 Institutes involved and equipment in use

Institute	Clock or frequency reference	TWSTFT modem	IGS station code	GPS receiver antenna
IEN	HM: Symmetricom, UTC(IEN): Agilent 5071	MITREX	IENG	ASHTECH Z-XII3T, ASH701945C_M
NPL	HM: Symmetricom, representing UTC(NPL)	SATRE	NPLD	ASHTECH Z-XII3T, AOAD/M_T
NIST	HM: Symmetricom, representing UTC(NIST)	SATRE	NISU	NOV EURO4-1.00-222, NOV600
OP	HM: Symmetricom	SATRE	OPMT	ASHTECH Z-XII3T, 3S-02-TSADM
PTB	HM: VREMYA-CH UTC(PTB): primary clock CS2	SATRE	PTBB	ASHTECH Z-XII3T, ASH700936E

Tab. 2 TWSTFT measurement schedule and Mitrex codes assigned to the stations. The receive (RX)and transmit (TX) frequency in Europe differs for the inner-European and the US-Europe link.

Measurements during each second hour of a day in UTC First	Last	Duration	OP		NPL		PTB		IEN		NIST	
hh:mm:ss	Hh:mm:ss	(s)	TX	RX	TX	RX	TX	RX	TX	RX	TX	RX
00:10:00	00:11:59	120	0	1	1	0						
00:13:00	00:14:59	120					4	6	6	4		
00:19:00	00:20:59	120	0	4			4	0				
00:22:00	00:23:59	120			1	4	4	1				
00:25:00	00:26:59	120	0	6					6	0		
00:28:00	00:29:59	120			1	6			6	1		
00:37:00	00:38:59	120	0	6							6	0
00:40:00	00:41:59	120			1	6					6	1
00:49:00	00:50:29	120					4	6			6	4
00:52:00	00:53:59	120							6	6	6	6

days in October and November 2004, Modified Julian Days (MJD) between 53304 and 53329. The five institutes are part of a regularly operated TWSTFT network, and for the duration of the campaign an intensified measurement schedule was agreed upon. The institutes also operate geodetic GPS receivers and their stations belong to the network of sites of the International GPS Service (IGS) [12]. The institutes obtained the support of the Astronomical Institute of the University of Bern (AIUB), Switzerland, which processes the GPS data with an innovative software, providing long-term monotonous series of clock solutions for the stations involved with extremely high resolution, based entirely on the GPS carrier phase (GPS CP) [13]. Last but not least, the BIPM provided the TAI P3 analysis of the GPS data [8].

It finally turned out that the fountains of NIST and PTB were not ready to provide data during the campaign, so that fountain comparisons were only carried out between the IEN CsF1, the NPL CsF1, and the OP-FO2. The results are reported in a separate paper [14]. Also in this paper, particular attention is given to the properties of the time links between the three relevant institutes. But the analysis included all available links and techniques, which has not been done before in such depth. Clearly, only a selection of the results could be presented in this paper, and we aimed at showing the achievable performance as well as pointing to some features that were not understood. Results beyond those presented here will be made available elsewhere.

In the following sections we will detail the equipment and time transfer techniques involved before presenting results for each technique and for inter-comparisons among techniques.

2. DESCRIPTION OF HARDWARE AND TIME TRANSFER TECHNIQUES

2.1 Station equipment

In Tab. 1, we describe the equipment involved in the participating stations. In all cases, the hydrogen masers served as the fly-wheel oscillators for comparisons with the fountain frequency standards, and the maser signals were connected to the time transfer equipment. This most straightforward configuration is depicted in Fig. 1, indicating the three different time transfer techniques connecting the masers at the remote sites. In the cases of NIST and NPL, the maser output represented the local UTC scale, UTC(NIST) [15], and UTC(NPL), respectively. In the case of IEN and PTB, however, the GPS receiver was connected to UTC(IEN) and UTC(PTB), respectively, which are derived from caesium clocks, a

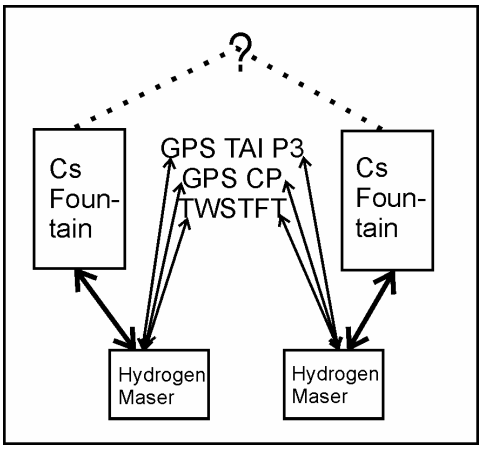


Fig. 1 Configuration of frequency comparisons during the campaign, the frequency standards at left and at right, respectively, are operated in different institutes.

commercial model at IEN, and the primary clock CS2 at PTB [16]. As a result, further steps in the data analysis were required to reference all techniques to the same frequency standard. The time differences UTC(PTB) – HM(PTB) were available only on an hourly basis and this dictated the minimum averaging time for GPS CP and GPS TAI P3 data presented below. At IEN the corresponding measurements were taken more frequently so that a more complete data analysis could have been performed in principle. This has, however, not been included in this paper.
In the sections dealing with results links are often only denoted by the station acronyms, e. g. by PTB-NPL for simplicity. In the graphs, connecting lines serve the purpose to guide the eye.

2.2 TWSTFT analysis

The stations involved performed TWSTFT, co-ordinated by the CCTF Working Group on TWSTFT, following standard procedures: At each pair of stations, named A and B for now, a characteristic pseudo-random noise (PRN) spread spectrum signal which is designated by a certain Mitrex code (see Tab. 2), synchronous with the local 1 PPS time references $T(\mathrm{Ref_A})$ and $T(\mathrm{Ref_B})$, respectively, is generated on a 70 MHz intermediate frequency. It is up-converted to the rf region (Ku-band, up link at about 14 GHz) and transmitted to a geostationary telecommunication satellite, in the present case the IS-903 of Intelsat Corp., located at 325.5° East. At the satellite, the signal is translated to the downlink frequency (about 12 GHz) and sent back to earth. The time of arrival (TOA) of the PRN coded signal from station A (B) is determined at station B (A) by cross-correlation of the received and down-converted PRN signal with a local replica of the same PRN code, synchronous with the local reference clock. TOA measurements are made each second during sessions of two minutes per station pair. The time differences exhibited a 1 σ - standard deviation of below 400 ps in favourable cases, but going up to almost 1.5 ns for some links. This fact is discussed further below. At the end of a session, a quadratic function is fitted to the data and the TOA midpoint values derived from the fit are exchanged via file transfer through the internet. This allows the quantity of interest, $T(\mathrm{Ref_A}) - T(\mathrm{Ref_B})$, to be calculated, as explained by Kirchner [7] and in an ITU-R Recommendation [17]. For the purpose of an intense comparison, TWSTFT sessions were performed nominally once every two hours. The data sets contained very few gaps, on some links about 315 out of 320 possible data points were recorded, while in a few cases about 20 data points were missing. The TWSTFT measurement schedule including the assigned Mitrex codes is reproduced in Tab. 2. One can recognize that measurements were not taken perfectly simultaneously.

2.3 GPS carrier- phase analysis

All five participating institutes contribute to the IGS network. The observations are provided in RINEX format and contain pseudorange (code) and carrier-phase measurements from the two GPS frequencies for all tracked satellites with a sampling interval of 30 seconds. The midnight epoch is missing in the observation files from the stations OPMT and NPLD, but the receiver tracked continuously. Therefore, this has no effect on the GPS CP solution (apart from the missing results for this particular epoch).
The GPS CP analysis was performed at AIUB using the Bernese GPS Software, Version 5.1 [18]. The solutions are based on the contributions to the IGS of the Analysis Center CODE (Center for Orbit Determination in Europe) hosted at AIUB. CODE is a collaboration between the Astronomical Institute of the University of Bern (AIUB), the Federal Office of Topography (Swisstopo, Switzerland), the Federal Agency for Cartography and Geodesy (BKG, Germany), and the Institut Géographique National (IGN, France). The satellite orbits, Earth orientation parameters, station coordinates, and troposphere delays are introduced from the (geodetic) three-day, double-difference solution calculated for the IGS at CODE. In the clock estimation procedure the receiver and satellite clocks are solved for along with other parameters (e.g. phase ambiguities). The result is a network solution with a set of consistent receiver and satellite clock corrections for each epoch. All receiver clock comparisons (baselines) are extracted from this network solution, hence all loops of extracted baselines give zero. The GPS CP method imposes no a priori information on the clocks by any filtering. The clock corrections are freely estimated for each epoch. The GPS CP solution used in the current campaign contains no day boundary discontinuities because the phase ambiguities from the daily data processing are reconnected before estimating the clock parameters. As in addition frequency transfer has been the main interest, the code observations have been discarded.

Such an approach, using only phase information and neglecting code information, has never been made before for a several-weeks long campaign by the AIUB and, as far as we know, has also not been published by any other group. A detailed discussion is given in [13].

2.4 GPS TAI P3 analysis

The GPS TAI P3 technique has been used since 2002 by the BIPM, with about 15 participating laboratories equipped with geodetic GPS receivers. The dual-frequency P1 and P2 observables delivered by the receivers are linearly combined to form the ionosphere-free P3 observable. They are transferred to the CGGTTS format used in standard GPS time transfer [8, 19] using locally recorded broadcast GPS parameters. This format dictates in particular the sampling rhythm for the time differences of 16 minutes, and that the tracking schedule shifts in time from day to day. The CGGTTS data files are gathered by BIPM and used to compute time links after applying different corrections: precise satellite orbits and clocks obtained from the IGS, and station displacement due to solid Earth tides. The six time links (no data from NIST were processed) were computed using the common-view technique. For each 16-minute interval, all available common-view differences were formed and averaged with a weighting scheme based on the satellite elevation, after a first screening for big outliers. BIPM provided unsmoothed and Vondrak smoothed data, the results shown below were derived from the former.

3. RESULTS

3.1 General observations

To give an example, data obtained on the link NPL – OP during one day are depicted in Fig. 2. With a small offset in the epoch of data taking, the TWSTFT measurements once every two hours (red dots, one data point missing) follow the GPS CP data (open blue circles) obtained at the full hour. As it is important to eliminate as far as possible the effect of clock instabilities when comparing the two techniques (see Section 4.1), in such cases the GPS CP measurement taken closest in time to the TWSTFT measurement (see Tab. 2) was used. The TAI P3 data at 16-minute intervals (black dots) are offset by 2 ns for clarity.

A few data points were in general missing for all links, and even complete GPS TAI P3 data would not represent a continuos time series. In all stability plots, however, the minimum τ was chosen as the nominal value irrespective of the actual mean separation in time between measurements as the availability was quite high and the impact of missing data on the plots was considered as minor in the current study.

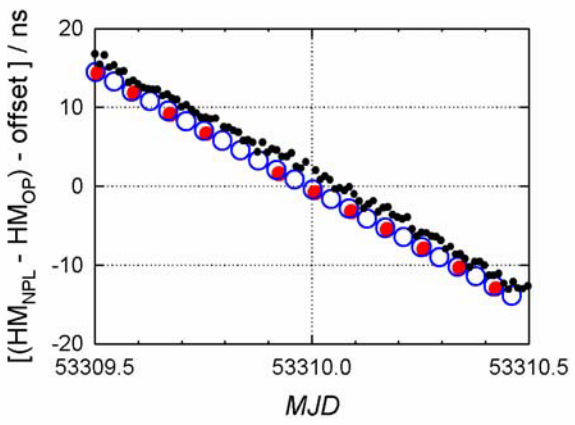


Fig. 2 Example of data collected during one day on the NPL – OP link; GPS CP (open circles), TWSTFT (red dots), and GPS TAI P3 (black dots), offset by 2 ns for clarity.

In Fig. 3 the general performance of the hydrogen masers involved is illustrated. Here UTC(NIST) was chosen as the common reference, and GPS CP time differences to all other masers at an hourly separation were evaluated, linear trends were subtracted, and the residuals to the linear fitted functions were plotted. It can be ssen that the NPL maser performed very well except for a kink towards the end of the campaign, and the OP maser had a significant non-linear frequency change during the period.

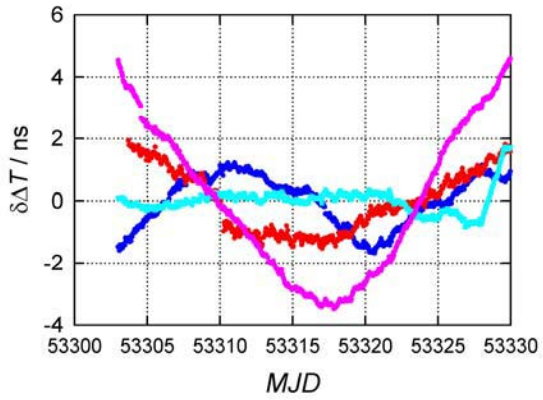


Fig. 3 General characteristics of the frequency standards at NIST (common reference), IEN (red), NPL (cyan), OP (magenta), and PTB (blue) during the study period, MJD 53303 – 53330; residuals to linear fits to time differences UTC(NIST) – local references obtained via GPS CP.

3.2 TWSTFT results

The extended TWSTFT schedule provided valuable insight in the stability of TWSTFT links. At an averaging time of about one day the measurement noise (white phase noise) tends to become insignificant compared to the instability of the masers involved. It turned out that links to NIST, which utilize a different transponder on the satellite than the intra-European links, are in all cases more noisy. Fig. 4 showing PTB-NPL and PTB-NIST data may serve as an illustration, and Fig. 6 below gives further proof.

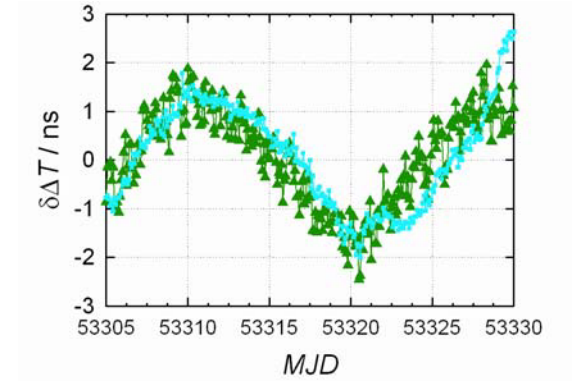


Fig. 4 Time differences PTB – NPL (cyan), and PTB – NIST (green) obtained via TWSTFT, mean time offset and mean rate removed.

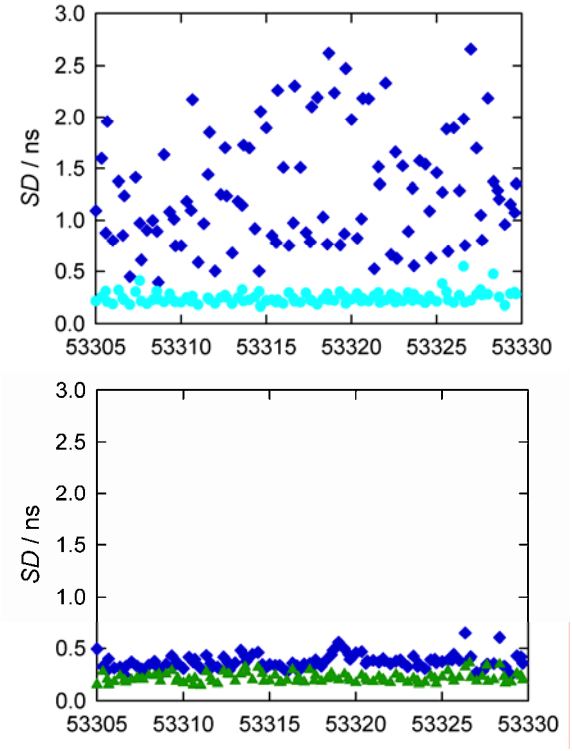


Fig. 5 Observed instability of the TWSTFT data for the links PTB-NPL (upper) and PTB-NIST (lower); standard deviation in ns of the individual one second measurement data recorded at PTB (blue diamonds), NPL (cyan circles) and NIST (green triangles), respectively, around the quadratic regression function during the comparison campaign.

Here we find $\sigma_X(\tau = 2\text{ h, PTB-NPL}) = 0.12$ ns, and $\sigma_X(\tau = 2\text{ h, PTB-NIST}) = 0.28$ ns. These findings are seemingly in contradiction to the observed standard deviation of individual one second measurement results around the quadratic regression function obtained at the respective sites which are reported in Fig. 5. Apparently, the short-term measurement noise whose origin is not yet understood is efficiently averaged, and longer term variations dominate the links to the US. Presumably, these are in general due to the longer baseline which may have an impact due to unaccounted satellite motion (respectively incorrect time tagging of the measurements) or due to uncorrelated environmental effects. Errors due to unaccounted satellite motion are distinctively sinusoidal in shape (with a one day period), and this was not observed. The temperature sensitivity of the NIST ground station as well as the unaccounted ionospheric delay (small at Ku-band anyway) can be ruled out as causes. What remains as a potential cause is the instability of the transponder on the satellite, which was also noted by the United States Naval Observatory (which is routinely part of the TWSTFT network, but which did not participate in the extended measurement schedule).

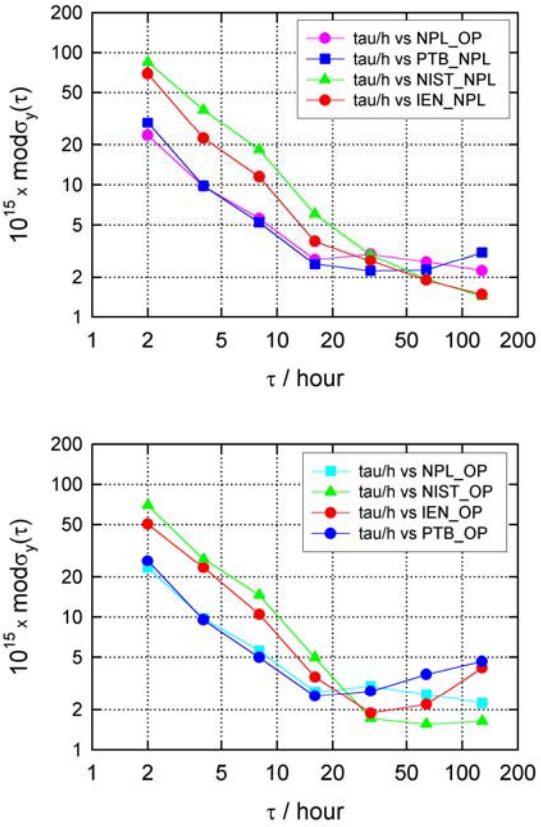


Fig. 6 Relative frequency instability, expressed by $\text{mod}\sigma_y(\tau)$, in the TWSTFT comparison of masers at NPL (upper graph) and OP (lower graph) to the other participating stations explained in the insets to the figures.

The frequency instability obtained from the TWSTFT comparisons of all institutes with respect to the masers at OP and NPL is depicted in Fig. 6. The lowest measurement noise is observed in the links between NPL, OP, and PTB. The slightly excessive noise in the links to IEN is probably due to the older TWSTFT modem (of Mitrex type) still used at IEN. When forming the so-called "closure" (OP – NPL) + (PTB – OP) + (NPL – PTB) a 0.34 ns deviation from zero is observed, which can be partially explained by the mean clock rate of the OP maser of about –28 ns/day and the fact that the two measurements involving OP

are separated by about 9 minutes (see Tab. 2). A real closure error would point to the fact that the signal delays in the receive path of at least one of the stations depend on the received PRN code. The closure data exhibited a standard deviation around the mean value of 0.14 ns, which represents the combined measurement noise of three TWSTFT measurements plus the time stability of the frequency standards over 9 minutes. The closure results are depicted in Fig. 7.

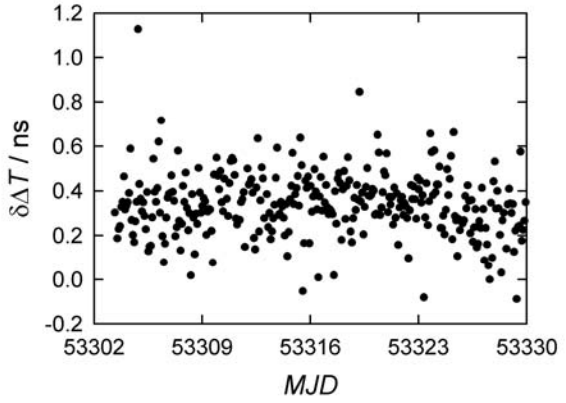


Fig. 7 TWSTFT "closure" results – in total 312 data points - combining the three individual measurements OP – NPL, PTB – OP, and NPL – PTB collected during the campaign according to the schedule represented in Tab. 2

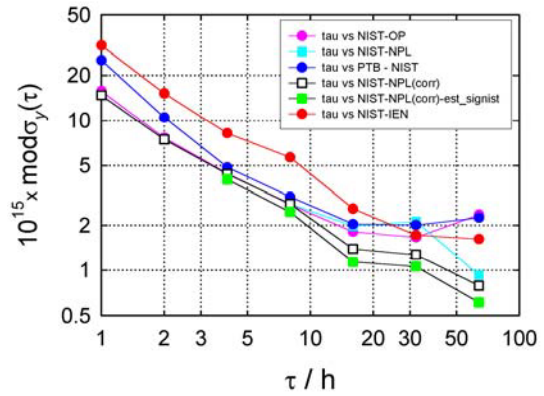


Fig. 8 Relative frequency instability, expressed by mod$\sigma_y(\tau)$, in the GPS CP comparison of UTC(NIST) with masers at OP, NPL, PTB, and IEN, see coding of symbols in the inset. Additional plots: instability after removal of the frequency step of the NPL maser (open squares), and after subtracting thereof the estimated instability of UTC(NIST) (green squares).

3.3 GPS CP results

The GPS CP data analysis revealed a very low level of measurement noise and thus an assessment of the instability of the frequency standards involved at the shortest averaging time was possible. As explained before, the original data of AIUB had UTC(PTB) and UTC(IEN) as the references. In order to get a meaningful comparison of the three frequency transfer techniques, local measurements were used to provide the link to the masers at PTB and IEN, and only hourly data were dealt with. The results of GPS CP frequency comparisons were previously used as illustration (see Fig. 3). In Fig. 8 we provide the frequency instability derived from these data. Two plots were added which give further proof of the merits of this technique. When the frequency step of the NPL maser (see Fig. 2-3) after MJD 53328 is removed from the data, the NPL-NIST instability is further lowered (open squares). The frequency instability of UTC(NIST) has been estimated based on an analysis with respect to the group of active hydrogen masers in operation at NIST [15]. Subtracting this noise contribution from the previously mentioned data should allow the combined instability of the frequency transfer and the NPL maser (green squares) to be assessed.

3.4 TAI P3 results

In the TAI P3 analysis the same raw data provided by the geodetic GPS receivers are employed as in the GPS CP analysis discussed before, but the code measurements rather than the phase data are employed. Thereby the analysis is simplified, and in fact BIPM does the analysis on several links routinely as part of the monthly work of generating TAI. To fully assess the quality of the TAI P3 data it would have been necessary to resample the GPS CP measurements according to the epochs of the midpoints of the TAI P3 16-minute average values. This was postponed for a further refinement of this study. In Fig. 9, the frequency instability results obtained in the three links connecting to NPL are depicted. Only the link to OP connects two hydrogen masers and is suited to assess fully the merits of the TAI P3 analysis. The IEN (red) and PTB (blue) data are characterized by white frequency noise (slope −1/2) at averaging times around one day, and the observed instability is very close to the expectation value for a commercial caesium clock in the case of IEN and for the primary clock CS2 of PTB [16].

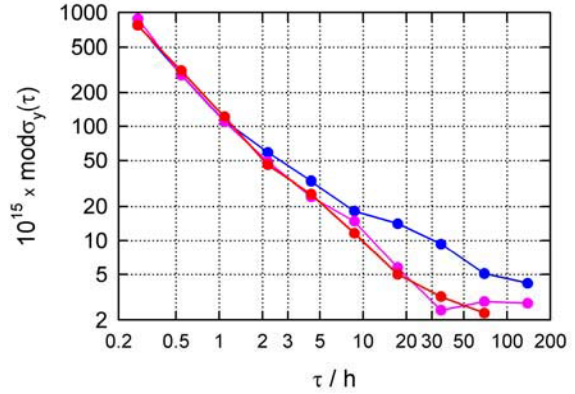


Fig 9 Relative frequency instability, expressed by mod$\sigma_y(\tau)$, in the GPS TAI P3 comparison of UTC(NPL) with the maser at OP (magenta), and UTC(IEN) (red), and UTC(PTB) (blue).

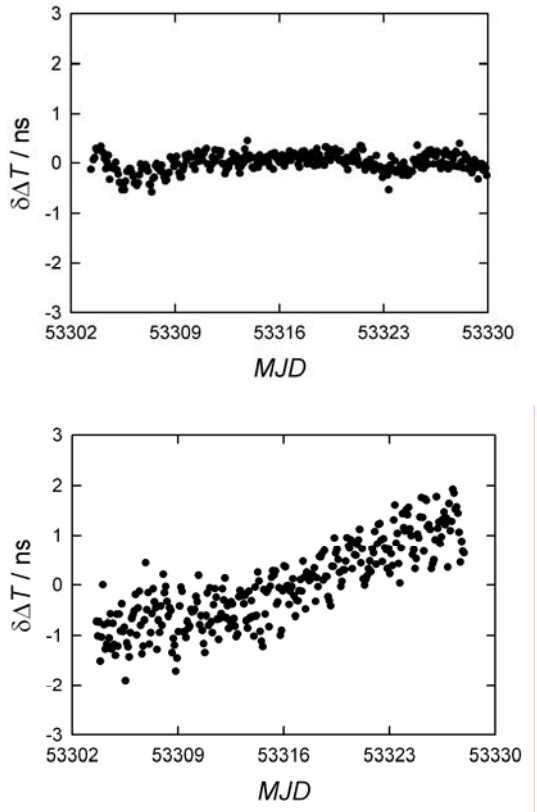


Fig. 10 Double differences of TWSTFT – GPS CP results, mean offset subtracted, upper: NPL – OP, lower: NIST – OP.

4. COMPARISONS

4.1 Discussion TW – GPS CP

During this campaign, the signal delays in the time transfer equipment were not of relevance since frequency comparisons, not true time comparisons, were the major concern. In this case only the variations of such signal delays matter since they could cause an erroneous frequency measurement result, which would in turn be detrimental for the success of the fountain comparisons [14]. TWSTFT data and GPS CP data obtained for the same epoch - within the 30 s separation of the GPS CP data - were compared. In Fig 10 two examples are depicted. In both cases, the mean offset in the data was subtracted. Whereas in the NPL-OP data all points are scattered almost within a 1-ns interval, a significant drift of about 2 ns is obvious in the NIST-OP data. If the latter was attributed to one of the techniques it would constitute a frequency error generated by the varying delays of only slightly less than one part in 10^{15} which is at the margin of being acceptable. A drift of the same magnitude is found when the link NPL-NIST is analyzed, but only half the magnitude is observed in case of NIST-PTB. No significant drift is visible in the links among NPL, OP, and PTB. When the stability of the data between OP and all four stations are analyzed, the results depicted in Fig. 11 are obtained. Since the data represent the combined uncertainty of two techniques, the PTB-NPL-OP data prove the capability of frequency comparisons with a measurement uncertainty of close to 1 part in 10^{15} at averaging times of 1 day, and the potential to do much better once a longer averaging time is accepted.

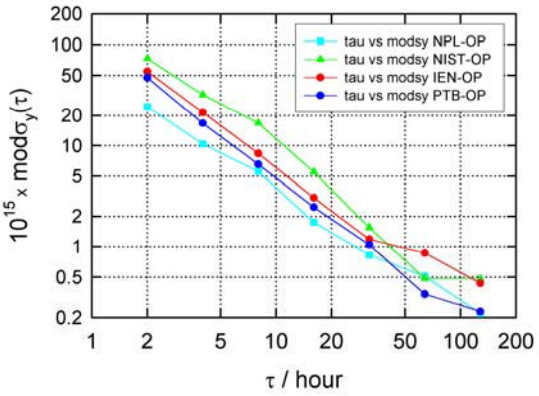


Fig. 11 Stability analysis of double differences TWSTFT – GPS CP for all links referenced to OP, see inset for the coding of the symbols.

4.2 Link comparisons and discussion of the frequency transfer uncertainty

In the graphs of Fig. 12 we compare the instability of a frequency comparison for the links NPL-OP, NPL-IEN, and IEN-OP using the three techniques. In all cases the same frequency references were used. In addition, the graphs contain the calculated instability for the double differences TWSTFT minus GPS CP. These graphs represent in a certain way the summary of the finding presented in different graphs before. We notice that the instability of GPS CP frequency transfer reaches the clocks' instability at about half a day; in the case of TWSTFT a full day is needed, and about 2 days are needed in the case of TAI P3. In case of longer averaging times there is no significant distinction between the three techniques visible from the graphs because of the clocks' instability. The instability of double differences reflects the long-term delay variations of the equipment involved. Since the equipment is the same for GPS CP and TAI P3 we cannot expect a difference in the long-term characteristics of the two techniques. This could of course be proven from a long-term comparison of TAI P3 and TWSTFT links which are routinely evaluated independent of a specific campaign.

In conclusion we compare the mean frequency differences obtained for selected links using the different techniques. The mean values were calculated in a straightforward manner, (δT(end) – δT(start)) / measurement interval. We chose the triplets of stations IEN-NPL-OP and NPL-OP-PTB. TAI P3 data with reference to the PTB maser were calculated based on its average frequency with respect to UTC(PTB) during the full period. The results are compiled in

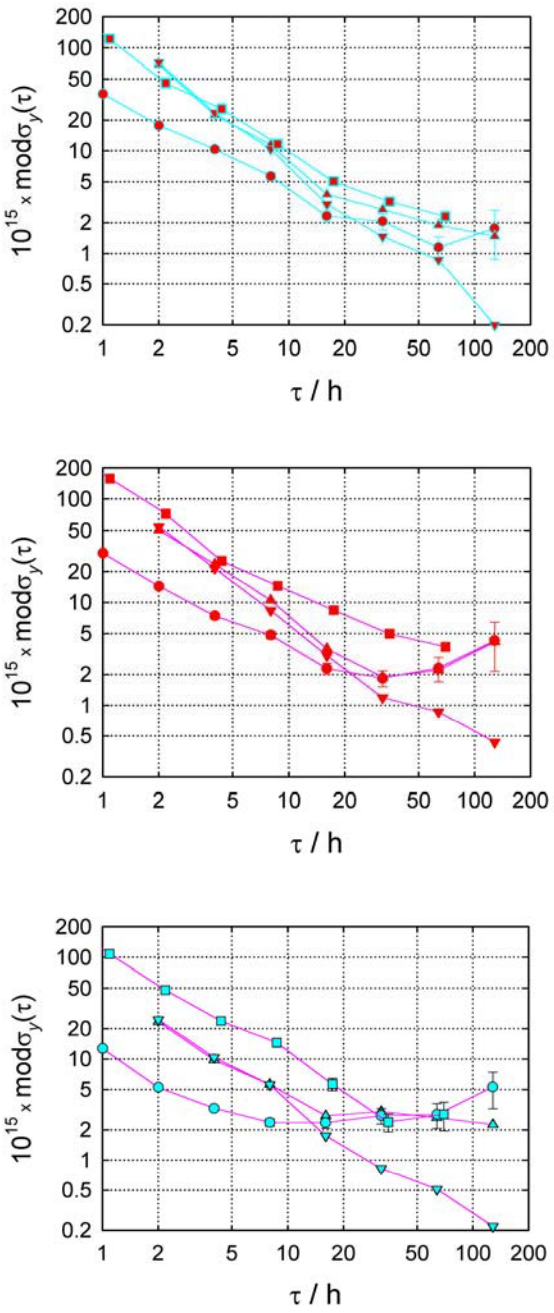


Fig. 12 Relative frequency instability, expressed by mod$\sigma_y(\tau)$, obtained for the three techniques, TAI P3 (squares), GPS CP (circles), and TWSTFT (symbol triangles up). In addition, the instability of double-differences TWSTFT – GPS CP (triangles down) has been illustrated for each link; upper graph: link IEN-NPL: middle graph: link IEN-OP; lower graph: link NPL-OP.

Tab. 3 Mean relative frequency differences in parts in 10^{15} during the interval 53304 to 53323 (inclusive, 20 days) between hydrogen masers at IEN, NPL, and OP, calculated using the three frequency comparison techniques, and the closure results

	IEN – OP	IEN – NPL	NPL – OP	Closure
TWSTFT	-430.8	-95.3	-336.1	0.6
GPS CP	-430.8	-94.5	-336.3	
TAI P3	-431.6	-95.1	-336.5	0.0

Tab. 4 Mean relative frequency differences in parts in 10^{15} during the interval 53304 to 53324 (inclusive, 21 days) between hydrogen masers at NPL, OP, and PTB and the closure results.

	PTB – NPL	PTB – OP	NPL – OP	Closure
TWSTFT	-74.24	-410.25	-336.08	0.07
GPS CP	-73.88	-410.13	-336.25	
TAI P3	-73.67	-410.07	-336.5	0.1

Tables 3 and 4. As said before, in the GPS CP analysis all loops comparing stations give zero. The maximum deviation within one column is 1×10^{-15}, but is typically much smaller.

5 CONCLUSIONS

We have reported on the results of a campaign originally stimulated by the intention to compare caesium fountain frequency standards in five institutes. GPS CP based analysis was used for the first time in such a campaign comprising a network of institutes. The campaign has provided much data not all of which could be presented in detail here. The following results are very encouraging:

1. TWSTFT in an intensified schedule and GPS CP analysis have allowed a frequency comparison between remote standards with a statistical uncertainty of 1×10^{-15} at averaging times of one day.
2. All studied techniques seem to be equally suited for the comparison of fountains without contributing significantly to the combined uncertainty of such comparisons, provided that a few days of averaging are allowed.
3. Systematic errors in frequency comparisons using any of the techniques did not exceed 1×10^{-15} during the period under study.

Further studies should address the following subjects

1. the nature and magnitude of drifts in the comparison to NIST;
2. the cause of the excessive noise observed in some TWSTFT links;
3. the optimum use of TAI P3 data, in continuation of previous work done by Petit and Jiang [20].

The low measurement noise provided by GPS CP at averaging times of less than one day makes it the first choice for use in comparisons among remote optical frequency standards since it is currently very challenging to operate such devices for extended periods. In the existing network such a comparison could be organized between NIST, NPL, and PTB, whose standards would be almost immediately ready for such purposes, and the experiences provided by the current study could be helpful in organizing such a comparison.

ACKNOWLEDGEMENT

The Corresponding Author would like to express his gratitude to his co-authors for providing guidance and rapid support. This work would not have been possible without the help of several individuals in the participating stations who took care of the operation of equipment, data retrieval, and analysis: Jürgen Becker, Davide Calonico, John Davis, Philippe Merck, Diego Orgiazzi, Valerio Pettiti, Dirk Piester, Jean-Yves Richard, David Valat. Thanks to Peter Whibberley for guidance in preparing the manuscript. The transponder on the IS-903 was provided free of charge by Intelsat Corp.

REFERENCES

[1] S. Bize et al., Advances in atomic fountains, *C. R. Physique*, **5** (2004), pp. 829-843.
[2] S. R. Jefferts et al., Accuracy evaluation of NIST-F1, *Metrologia* **39** (2002), pp. 321-336.
[3] S. Weyers, A. Bauch, R. Schröder, Chr. Tamm, The Atomic Caesium Fountain CSF1 of PTB, in Proc. 6th Symposium on Frequency Standards and Metrology (St. Andrews, Scotland, Sept. 2001), pp. 64-7
[4] K. Szymaniec, W. Chalupczak, P. B. Whibberley, S. N. Lea, D. Henderson, Evaluation of the primary frequency standard NPL-CsF1, *Metrologia* **42** (2005), pp. 49-57.
[5] F. Levi, L. Lorini, D. Calonico, A. Godone, IEN-CsF1 accuracy evaluation and two-way frequency comparison, *IEEE Trans. Ultrason. Ferroelec., and Freq. Control* **51** (2004), pp. 1216 – 1224.
[6] J. Levine, Time and frequency distribution using satellites, *Rep. Progr. Phys.* **65** (2002), pp. 1119 – 1164.
[7] D. Kirchner, Two-way Satellite Time and Frequency Transfer (TWSTFT): Principle, implementation, and current performance, Review of Radio Sciences 1996-1999, Oxford University Press, 1999, pp. 27-44.
[8] P. Defraigne, G. Petit, Time transfer to TAI using geodetic receivers, *Metrologia* **40** (2003), pp. 184-188.
[9] T. Schildknecht, G. Beutler, W. Gurtner, M. Rothacher, Towards subnanosecond GPS time transfer using geodetic processing techniques, in Proc. 4th European Frequency and Time Forum (Neuchâtel, Switzerland, 1990), pp. 335-346.
[10] K. M. Larson, J. Levine, Carrier-phase time transfer, *IEEE Trans. Ultrason. Ferroelec., and Freq. Control* **46** (1999), pp. 1001-1012.
[11] T. Parker, P. Hetzel, S. Jefferts, S. Weyers, L. Nelson, A. Bauch, J. Levine, First comparison of remote cesium fountains, in Proc. 2001 IEEE Freq. Control Symp. (Seattle, Washington, 2001), pp. 63-68.
[12] J. Ray, K. Senior, IGS/BIPM pilot project: GPS carrier phase for time/frequency transfer and timescale formation, *Metrologia* **40** (2003), pp. 270-288.
[13] R. Dach, L. G. Bernier, G. Dudle, T. Schildknecht, U. Hugentobler, Geodetic frequency transfer – actual developments in the AIUB-METAS collaboration, 19th EFTF (Besançon, France, 2005), in these Proceedings.
[14] D. Calonico et al., Comparison between remote Cs fountain primary frequency standards, 19th EFTF (Besançon, France, 2005), in these Proceedings.
[15] T. Parker, Hydrogen maser ensemble performance and characterization of frequency standards, in Proc. of the Joint meeting of the European Frequency and Time Forum and the IEEE International Freq. Control Symp. (Besançon, France, 1999), pp. 173-176.
[16] A. Bauch, The PTB primary clocks CS1 and CS2, *Metrologia* (2005) in press.
[17] Recommendation ITU-R TF.1153-1 "The operational use of two-way satellite time and frequency transfer employing PN time codes", ITU, Radiocommunication Study Group, Geneva, last update 2003.
[18] U. Hugentobler, S. Schaer, P. Fridez, eds., The Bernese GPS Software 4.2., University of Berne: Astronomical Institute, February 2001.
[19] D. W. Allan, C. Thomas, Technical directives for standardization of GPS time receiver software, *Metrologia* **31** (1994), pp. 69-79.
[20] G. Petit, Z. Jiang, Stability of geodetic time links and their comparison to two-way time transfer, in Proceedings of the 36th PTTI (Washington DC, 2004), in press.

RESULTS OF A FIELD TRIAL CAMPAIGN FOR THE PERFORMANCE ASSESSMENT OF AN INNOVATIVE TIME & FREQUENCY DISSEMINATION SYSTEM BASED ON ESA PATENT 407

D. GIUNTA [(1)], G. BUSCA [(2)], A. DELLA TORRE [(3)], M. VITTA [(3)], V. DE PERINI [(3)], R. GRIMOLDI [(3)], X. STEHLIN [(4)], P. ROCHAT [(4)]

[(1)] ESA-ESTEC, P.O. Box 299, 2200AG Noordwijk, (NL), email: domenico.giunta@esa.int; [(2)] KYTIME Sàrl, Ecluse 11, 2022 BEVAIX, (CH), email: tec.kytime@net2000.ch;
[(3)]Carlo Gavazzi Space S.p.A., v. Gallarate, 150 – 20151 Milano (I), email:adellatorre@cgspace.it; [(4)] TEMEX Neuchatel Time, 2000 Neuchatel (CH), email:stehlin@temex.ch

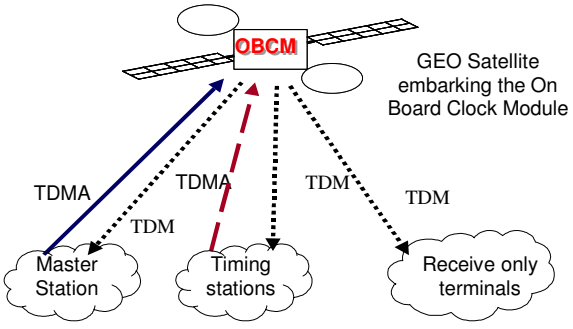
Keywords: Time & Frequency Dissemination, TDMA, PLL, Atomic Clock, On board clock module.

Abstract

The paper addresses the main achievements of a Field Trial Campaign successfully concluded on July '04 dedicated to the assessment of the performances achievable by a novel Time & Frequency Dissemination System based on ESA Patent 407.
The Field Trial campaign was performed in the frame of the ESA ARTES 5 program " Use of satellite Time Signal for Communication Systems".
The method allows transferring the accuracy and stability features of a state-of-the-art ground clock to a lower performances clock module on board a satellite .
Following the patent methodology the on board clock module inherits the accuracy and stability of the ground clock reference, becoming the master Time and Frequency dissemination source to all ground terminals located in the footprint of the satellite hosting the module.
The field trial tests were performed at SES ASTRA premises in Betzdorf, exploiting ASTRA 1A satellite.
The Field Trial experiment topology was based on satellite double hopping access using one On Board Clock module placed on ground, representative of a future flight Time & Frequency Dissemination Payload.
Three Terminals were actually deployed, performing respectively the roles of Master Clock Station (Ground Reference), On Board Clock Module (Time & Frequency Dissemination Payload) and Auxiliary / Timing Station (User terminal / Ground Reference Station backup).
The methodology on which the system is based, the theoretical analyses performed for the estimation of the expected performances and the previously performed system level activities, including the Field Trial campaign setup, were addressed in a precedent paper presented at EFTF 04 (Recent Developments in Time & Frequency Dissemination Systems, ref. [1]).
The present paper is the logical continuation and conclusion of the previous one and provides the description of the test campaign and of the main results achieved , as a first step for proceeding with the setup of a pre-operational system.
The test campaign confirmed the validity of the method, the expected performances and the design of the instrumentation. In particular, the test campaign measurements assessed a time stability of approximately 50 ps rms over an uninterrupted measurement interval of 8hrs and an extrapolated frequency stability over one day of the order of 10^{-15}, fully confirming the theoretical estimations.

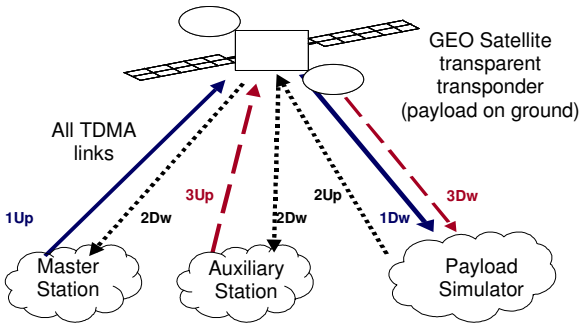
1. Program Background and Objectives

The Time & Frequency Distribution system concepts developed in the frame of the subject ESA ARTES 5 program are based on the innovative principle to distribute a very accurate and stable ground-based time reference by means of a satellite network, in agreement with ESA Patent 407.



Envisaged T&F Dissemination System Configuration

This is done by synchronizing a low performance On-board Oscillator (slave clock) with a very accurate Ground Time Reference (master clock) and then distributing the space clock time to ground users.
In this way the on board clock module actually becomes the master Time & Frequency Dissemination source for all the ground terminals located in the footprint of the satellite hosting the on board clock module.
The Field Trial (FT) campaign featured a system architecture where the space segment, i.e. the telecommunications satellite hosting the on board clock module, is replaced by a double-hop satellite access allowing to locate on ground a representative prototype of an on-board clock module, without the need of flying the on board clock module, as planned in the final target system.

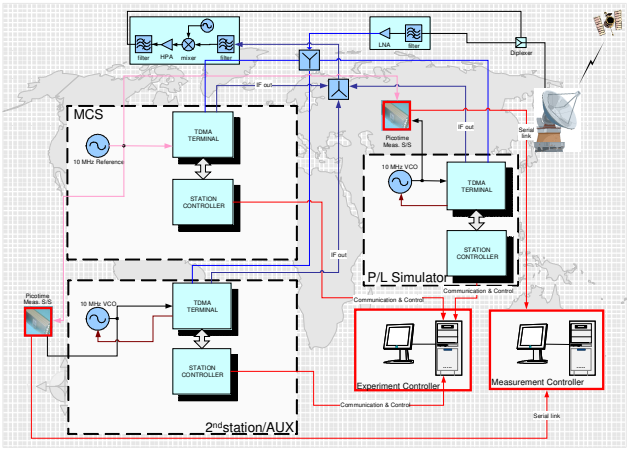


Field Trial Configuration Communications principle

This approach allowed to proof the concept with an easy deployment before embarking the "real space segment" on satellite.
The Field Trial Project team was leaded by Carlo Gavazzi Space and composed by TEMEX Neuchâtel Time, KYTIME and ALCATEL Space.
Carlo Gavazzi Space, as Prime Contractor, has been responsible for project management and technical co-ordination, system definition, definition of design, development and AIT plan, system integration and testing, management of the links with the satellite operator and field trial campaign results evaluation.
Moreover, Carlo Gavazzi Space developed the TDMA Terminal and the Station Controller which are the core elements of the Stations exploited during the experimental campaign, namely Master Clock Station, Satellite Payload Simulator and Auxiliary/Timing station.
TEMEX Neuchâtel Time was responsible of the specification and setup of the instruments to be exploited for the test campaign, including Atomic Clocks (RAFS, LPFRS), Measurement Subsystem (Picotime) and Micro Phase Stepper.
KYTIME supported Carlo Gavazzi Space in the implementation of the concepts defined in the ESA Patent 407, in the theoretical analyses, in the definition of the measurement methods employed during the test campaign and in the interpretation of the test results.
Alcatel Space supported Carlo Gavazzi Space on system definition aspects.

2. System Description

The Field Trial configuration general block diagram is shown in the following figure.
The system is composed of three stations, called MCS (Master Clock Station), P/L Simulator (PS) and AUX/2nd station.
For reasons of system configuration optimization this last station is capable of a double function, depending on the software loaded.



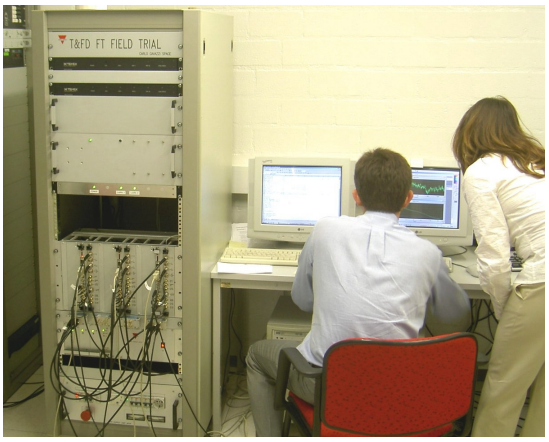
Field Trial Configuration General Block Diagram

The Timing Station is slaved to the P/L Simulator Clock, to perform performance measurements of the Time & Frequency Distribution function, or behaves as auxiliary/backup station, capable to take the role of MCS in case of MCS failure.
Each of the three stations is equipped with a couple of 70MHz IF input/output links to the SES ASTRA Ground Station exploited for the experiment and eventually to the transparent transponder hosted on board the ASTRA 1A satellite via Ku band links.
Any communication among the three stations is routed through the satellite.
Data exchanged between the stations is inserted into a dedicated slot of a P/L simulator generated TDMA frame. This implies that each station (MCS and AUX) must perform a TDMA slot acquisition when starting to transmit.
The M&C section is divided on its turn into two parts: the System Monitoring and Control block and the Measurement block.
The System Monitoring and Control block is based on a commercial environment (PC, commercial operating system and graphical MMI tool), connected to the three stations via serial links, running a MMI program; this block performs the acquisition and visualization, also in a graphical form, of the system parameters for each station.
This allows showing the contents of each station internal register and the output of the running algorithms, like the one for Delay and Doppler correction, for diagnostic purposes.

From the control program it is possible to send specific commands to each station, such as the ones dedicated to the programming of the internal registers controlling the station operations.
The Measurement block is based on two Picotime (Picosecond Resolution Test Set) devices from TEMEX Neuchâtel Time.
In the field trial configuration these equipments measure in real time the stability of the two slave clocks (Payload Simulator and 2nd / Auxiliary Station) with respect to the MCS clock, computing the Allan variance. The measurements are stored into dedicated test files and results are shown in real time to the operator in a graphical form.
The Picotime software, running on a personal computer, is able to show the time stability results as well as the frequency or phase samples in real time.

3. Field trial System layout

The system part specifically provided for the experimental campaign is hosted in a 19" 33U rack, except the RF subsystem and the Experiment and Measurement controllers. The rack layout is shown in the following Figure.



FT Configuration Instrumentation Rack Layout

The main units inserted into the rack are:

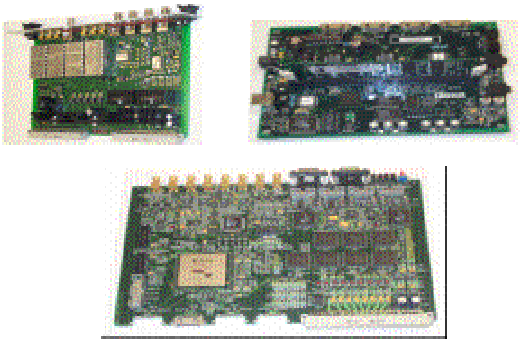
- **Dedicated Electronics Unit**
  This unit houses the electronics of the three stations (MCS, P/L Simulator and Aux.). The 3 boards composing each station (Station Controller, Digital and IF boards, developed by Carlo Gavazzi Space), have been placed side by side.
- **Power Electronics Unit**
  This unit contains the DC power supply modules feeding all the units of the rack
- **MCS Clock (RAFS)**
  This drawer houses the MCS reference clock.
- **P/L Simulator and Auxiliary Station Clocks**
  This drawer contains the two clocks (LPFRS) for the P/L Simulator ad the AUX stations
- **Measurement Subsystem Front End (Picotime)**
  The two stability measurement devices, Picotime ™, from TEMEX Neuchatel Time, are installed in one 3U drawer. Both units are connected to the Measurement Controller PC via a dedicated serial interface line.
  Both Picotime Units get as reference input the MCS clock while the two "Device Under Test" inputs are respectively connected to the P/L Simulator clock and to the Auxiliary Station clock, for measuring the related performances with respect to the MCS clock.
- **Microphase Steppers**
  Two Microphase stepper units were not used operationally in the FT system configuration under test but were used as laboratory instruments in case a precise phase step was necessary to adjust the clocks during the field trial campaign.



TEMEX Neuchatel Time Picotime and Phase stepper

- **Power Control & Distribution Unit**
  This 4 U drawer is dedicated to the power distribution, feeding all the units of the rack powered by a 220 V AC power.

In addition to the Units physically installed in the rack there are the two PCs running respectively the Experiment Controller and the Measurement SS SW.

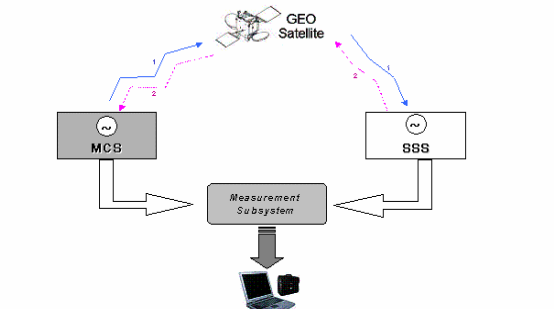


CGS TDMA Terminal: IF, Station Controller and Digital boards

4. Field Trial Campaign main achievements

Two measurement types were basically required in order to assess the ESA Patent 407 method effectiveness:
1. Measurement of the time difference between the Master clock and P/L SIMULATOR clock, locked to the master clock via the proposed technique. This

allows a direct verification of the synchronization error of the Space clock against the master clock, including the errors introduced by the Doppler compensation. This is a peculiar measurement not accessible in the real system.

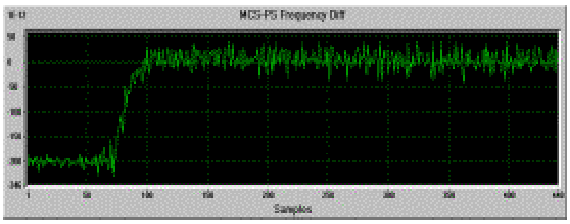


Test configuration 1

The following main test sessions were performed as part of the test campaign, as documented in the following:

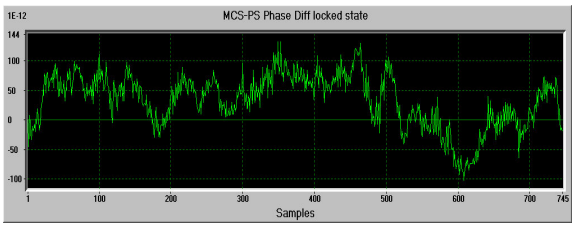
- Performance verification of the doppler compensation system by comparing the measured 2-way delay with the indipendent measurements performed by the ASTRA tracking station. The test was performed in conditions of important doppler (Df/f=0.5E-7 @ 16m/s).
- MCS – PS Synchronization test: during this phase only MCS and P/L Simulator are active, Aux Station is not yet part of the configuration. The tests results are spitted into two steps.
  a. The first one shows the behaviour of the MCS – PS frequency / phase difference during the transient phase, when PS starts locking to MCS.
  b. The second session shows the behaviour of the MCS- PS locking process at steady state, after the completion of the initial synchronization period

The following figure shows the behaviour of MCS–PS frequency difference during transient phase, when the PS starts tracking the signal received from MCS. The sample duration is 1 s.



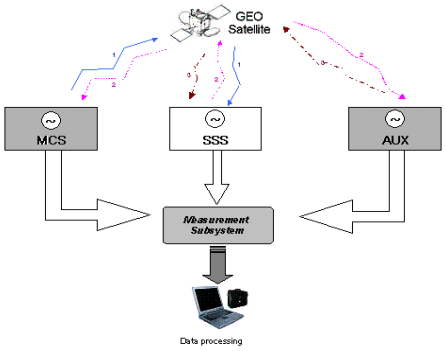
MCS-PS Frequency Difference during transient phase

The main transient phase lasts about 30 seconds. After this period the frequency difference is reduced and becomes about 0. After this transient, the frequency difference between MCS and Payload becomes less than 50e-12 Hz peak to peak with mean value equal 0.
The following figure shows the behaviour of the phase difference between the MCS and PS clocks during the Locked state. The deviation of the phase difference values around the mean value is about 100 ps peak to peak, as it is shown in the following figure.



MCS-PS phase difference during the locked state

2. Measurement of the time difference between the Master clock and the Auxiliary station clock, locked to the master clock via the proposed technique. This allows a direct verification of the User Terminals synchronization errors introduced by the technique, including the errors of double hop Doppler compensation. These measurements were performed by connecting the Master clock and slaved clocks (PS and Aux. Station) 10 MHz outputs to the two Picotime units part of the Measurement subsystem.



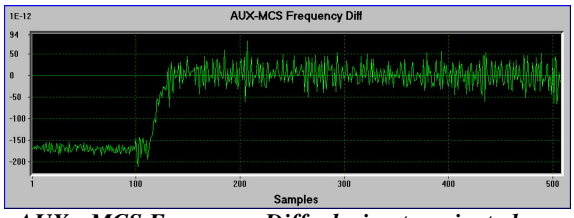
Test configuration 2

The following main test sessions were performed as part of the test campaign, as documented in the following:

- MCS – PS – Aux Station Synchronization test: After PS clock has been synchronized with the MCS clock, AUX Station starts the synchronization process with the MCS clock via the PS. During the TFD FT test campaign the AUX station behaves as a two way user of the Time & Frequency Dissemination System. Also during this test session it was tested first of all the behaviour during the transient phase and later on the behaviour during the steady state, when the Aux Station gets synchronized to the PS.
- MCS - Aux Station Hand-over Process performances test: The purpose of this test session was to verify the performances of the T & F Distribution process during the so called handover process, i.e. when in case of MCS failure the function of system master is passed from MCS to a backup master station, in this case simulated by the Aux Station.

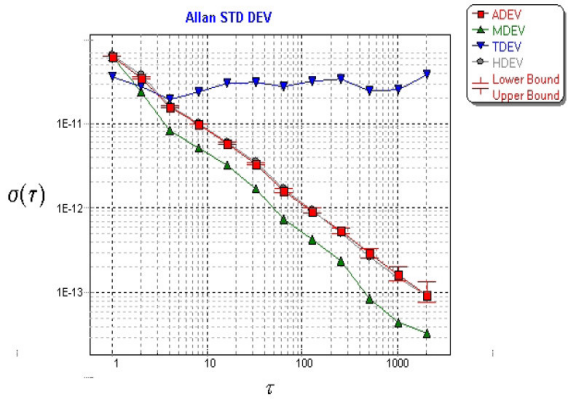
The following figure shows the transient phase during which the AUX station starts tracking the signal

received from PS. The diagram shows the Aux- MCS frequency difference.



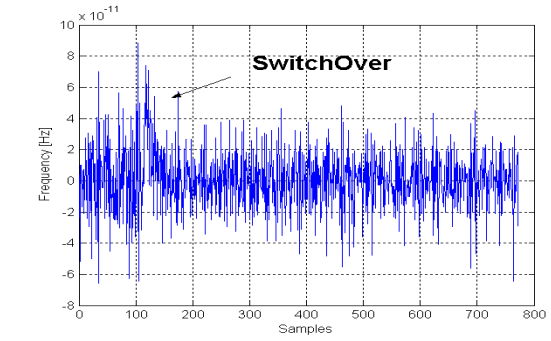
AUX - MCS Frequency Diff. during transient phase

As in the PS-MCS synchronization process, the main transient phase lasts about 30 seconds. During this period the frequency difference is reduced and becomes about 0.
After this transient, the frequency difference between MCS and Payload becomes less than 50e-12 Hz peak to peak with mean value equal 0.
The following figure shows the AUX vs. MCS Allan DEViation, Modified Allan DEViation and Time DEViation. These results show that the dominant noise process is a white phase noise corresponding to an ADEV improving as tau-1and a MDEV improving as tau-3/2 .



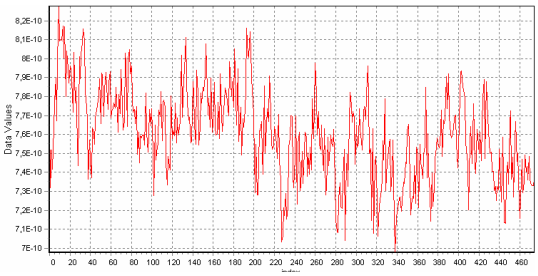
AUX Station vs. MCS Allan Deviation Plot

In the following figure the frequency difference between MCS and AUX during the switch over is shown.



MCS-AUX frequency difference behavior during switchover process

The hand-over process lasts few seconds (about 5s).
The Standard Deviation of the processed data (PS - AUX phase difference) is equal to 2.37e-11, as shown in the following figure. The x axis (index) is expressed in seconds.



PS - AUX phase difference during switchover

5. Experimental Campaign Conclusion

The following table provides a comparison of the Time and Frequency Distribution System performances estimated during the Field Trial study theoretical analysis phase against the ones actually achieved during the experimental campaign.

	Expected Performances (Simulations)	Field Trial Campaign Measurements
Time Stability (PL)	Standard Deviation 100 ps	Standard Deviation 49 ps
Time Stability (User)	Standard Deviation 250 ps	Standard Deviation 71 ps
Freq. Stability 1 day (PL)	Allan Deviation 10-15	Allan Deviation 1.23*10-15
Freq. Stability 1 day (User)	Allan Deviation 2*10-15	Allan Deviation 1.76*10-15

Comparison of the Time and Frequency Distribution System performances

The one day Allan deviation is extrapolated from the result obtained with the measurement over 8 hrs.
From the above data it can be noted that some measured performance figures are even better than the ones evaluated during the theoretical analysis phase. This is explained by the fact that the field trial configuration initially considered at theoretical analysis time was simplified and improved during the program evolution, reducing therefore the potential noise sources impacting the system performances.
Moreover, the following table provides the performances expected by the ESA patent 407 for the Time & Frequency Dissemination System.
The expected performances are given as function of the Symbol Rate used by the system and of the phase error featured by the Unique Word Detection Circuit part of the TDMA terminal. The Field Trial configuration setup is evidenced by the arrow while after the table are reported the performance figures actually measured during the experiment.

Phase error	Symbol rate	Rms jitter	Allan dev. over 1 day
0.1°	2 MHz	250 ps	2.86E-15
1°	2 MHz	2.4 ns	2.78E-14
0.1°	4 MHz	133 ps	1.54E-15
1°	4 MHz	1.2 ns	1.39E-14
FT Config. 0.1°	8 MHz	84 ps	9.72E-16
1°	8 MHz	604 ps	6.99E-15
0.1°	10 MHz	76 ps	8.78E-16
1°	10 MHz	484 ps	5.60E-15

ESA Patent 407 Expected Performances

The system loop time constant is of about 100 s. So, the loop is not acting on the short term but it is used to reduce the frequency drift of the slaved oscillator. The phase noise mask of the locked clock remains unchanged, but the long term behaviour of the clock can be improved.

6. Addressable market

The market for timing applications is characterized by its heterogeneous structure. It can be separated into the following major segments:

- the **scientific / institutional market**: it is the most demanding in terms of performance, but only very few users exist. This segment can be covered only by (inter)national entities who should take care of the creation and operation of such a "network". Such technique could of course be used for precise time transfer and experiments over much longer baseline will need to conduct to get this technique qualified for such application.
- the **mobile communication market**: it is driven by requested performances, overall costs, but also by quality of service. Same technique with fully integrated digital solution can be used in GSM /TDMA like transmissions in order to do precise frequency lock loop between base stations. Such solution could be interesting if costs can be reduced within few tens of € per base station.
- **Emerging request** for mobile location in relation with **rescue** (911) requirements is requiring precise timing for the computation of the mobile position. Such technique is a good alternative to GPS receiver in each mobile phone.
- The **military market**: secure networks using TDMA transmissions could be also used for precise synchronisation in relation to fast pseudo-random code acquisition used in positioning or code acquisition. Bi-static radar imaging techniques could also benefit from such very precise synchronisation mean over short base line.
- **Heterogeneous sub-segments** exist within each of these segments with specific needs in respect of the relevant performance parameters.

The markets show varying price sensitivities, ranging from high in the scientific / institutional market, over high/medium in the industrial sector, to medium, with a trend towards low, in the commercial sector.
The timing market has considerable potential for growth. It has been estimated that in 2005 the market size will be of the order of 35 Millions euros Worldwide (10M euros in Europe) and it is expected that it will grow up to 10% per year.
The main part of the timing applications market is within the highly industrialized areas of the world (advanced economy status). Even though the concentration of potential appliances within the third world countries and especially in the emerging nations will increase, the relevance cannot yet be compared with the potential of the highly industrialized nations.
On the other hand, especially those regions of Europe and the neighbouring countries not covered yet by long-wave time transmitters (such as DCF-77) might gain importance, too. These countries offer a considerable potential as TFD users due to non-established standards.

7. General conclusions and suggestions about future activities

The results of the experimental campaign have fully proven the validity of the Time & Frequency Dissemination Methodology described in ESA patent 407 and the relevant achievable performances.
Moreover, the development of the equipments required for the performance of the experiment allowed the project team to acquire a very deep experience which will be very valuable for the auspicated follow-on phases.
This is particularly true in the case of the TDMA terminals, which were fully developed in the frame of the program due to the unavailability on the market of equipments with the required characteristics.
The knowledge acquired will be of great help not only in the development of the ground terminals but also in the development of the flight terminals required by system configurations with the on board clock module located on board a satellite.
Moreover, the program due to its experimental nature allowed to acquire a good experience on all the involved topics, ranging from algorithms for Clocks locking, Doppler effect compensation and technology aspects related to the implementation of the system most critical components.
New design ideas related, for example, to the PLL architecture, have been generated which could allow to increase further the performances achieved with the current field trial configuration,.
The test campaign has also demonstrated that it is possible to achieve valuable Time & Frequency dissemination performances even in the Field Trial configuration with the payload located on ground.
This is of major importance for short-medium term follow-on activities because it was demonstrated that

with this configuration, fully deployed on ground without the need to fly an on-board clock module, it is already possible to satisfy the requirements of a wide range of potential users.
The field trial configuration might therefore be upgraded in order to support a pre-operational Time & Frequency Dissemination service without the need at this time to build a flight version of the on-board clock module.
The full validation of the system accuracy for metrological applications will require experiments involving independent and remotely located ground stations.
This exercise could be performed in a fairly limited timeframe and with an optimized budget due to the expertise acquired during the Field Trail Program and thanks to the predeveloped products.
The final goal of the program at medium-long term is however to build a flight version of the on board clock module and to fly it on a satellite in order to implement the final configuration foreseen in the ESA 407 patent.
The early exploitation study has also shown that the proposed methodology and the developed instrumentation can be effectively applied in fields other then the pure time and frequency dissemination, like for example satellite precise ranging, and that it will possible to apply this methodology also in applications far beyond the earth environment, like in the case of space exploration missions.

Acknowledgements

The authors would like to thank ESA for the support given to the initial studies and early definition of the Time and Frequency Dissemination System and the Italian, Swiss and French Delegations which, with their support, made possible the implementation of the Field Trial campaign in the frame of ARTES 5 program.

References

[1] D. Giunta, G. Busca, A. Della Torre, M. Vitta , V. De Perini, R. Grimoldi , X. Stehlin, P. Rochat "Recent developments in Time & Frequency Dissemination Systems"; Proc of 18th EFTF, 14-16 Mar 2004, University of Surrey, Guildford (UK)

[2] G. Garofalo , G. Busca, "Novel High Performance Satellite Time & Frequency Distribution and Synchronization System"; Proc of 14th EFTF, 14-16 Mar 2000, Torino (I)

[3] D. Giunta, P. Erhard, A. Della Torre, R. Grimoldi and M. Vitta, D. Levavasseur, O. de la Burgade "Signal Processing Techniques for Time and Frequency References Distribution by Satellite"; DSP 2001 Conf. Proc., Sesimbra (Portugal)

[4] D. Giunta, G. Busca, M. Vitta, V. De Perini, R. Grimoldi, X. Stehlin "Time Frequency And Distribution"; SPSC 2003 Conf. Proc., Catania (I)

[5] C.Mosquera, S.Scalise, G.Taricco, G.Garofalo, D.Giunta; *"Time-Transfer Performance in Burst-Mode Communication Systems"*; IEEE JSAC, Vol. 19, No. 12, December 2001

[6] G. Busca - "*Advanced Clock Generation and Distribution for Science Missions*" - ESA Contract no 12463/97/NL/MV Final Report

[7] Carlo Gavazzi Space - "*New Methods and Systems for Time & Frequency Distribution via Satellite*" - ESA Contract no 14427/00/NL/PB Final Report

[8] Alcatel Space- "New Methods and Systems for Time & Frequency Distribution via Satellite" - ESA Contract no 14428/00/NL/PB Final ReportM

GEODETIC FREQUENCY TRANSFER
Recent developments in the AIUB-METAS collaboration

R. Dach*, L.-G. Bernier†, G. Dudle†, T. Schildknecht*, U. Hugentobler*

*Astronomical Institute, University of Bern, Sidlerstrasse 5. CH-3012 Bern, Switzerland
e-mail: rolf.dach@aiub.unibe.ch, phone: ++41–31–631 3802, fax: ++41–31–631 3869
†Swiss Federal Office of Metrology and Accreditation, Switzerland

Keywords: GPS carrier phase, Frequency transfer, Continuous time series.

Abstract

The use of GPS Carrier Phase (CP) techniques is now widely accepted as an inexpensive and practical method to compare time and frequencies over large distances. A collaboration between the Astronomical Institute of the University of Bern (AIUB) and the Swiss Federal Office of Metrology and Accreditation (METAS) has produced a dedicated Geodetic Time Transfer Terminal (GeTT-Terminal) which was the subject of earlier communications.

Whereas in the past these terminals have been used as mobile units and deployed on various locations, a permanent station has now been set up at METAS and shall become a site of the International GPS Service (IGS) network. In parallel AIUB, responsible for the development of the processing software and the processing of the data, has improved the time and frequency transfer tools of its Bernese GPS Software package.

Small discontinuities in the resulting time series occur if the data are analysed in daily independent computation batches. Depending on the mean noise behaviour of the code observations a magnitude of up to one nanosecond may result for these artificial "clock jumps". Implementations in the Bernese GPS Software have been carried out to overcome the day boundary discontinuities:

- the hand–over of the receiver and satellite clock parameter at midnight, and
- the stacking of the ambiguity parameters at the day boundaries.

For the geodetic time and frequency transfer these implementations allow to generate a solution for an infinite time interval.

The paper describes the algorithms and focuses on an additional benefit from these developments: A geodetic frequency transfer solution can be computed based on the phase measurements only without using code data at all. In this case inconsistencies between the code and phase data caused by internal receiver clock events will not affect the solution. The stacking of ambiguities at the boundaries of independently analysed time intervals allows to generate a frequency transfer solution for a long time interval that is only limited by a loss of lock to all satellites.

1 Introduction

The Astronomical Institute of the University of Bern (AIUB) and the Swiss Federal Office of Metrology and Accreditation (METAS) have been collaborating for a long time in the field of geodetic time transfer, i.e. the use of the GPS carrier phase observations for time transfer (GPS CP). The development of the hardware was carried out at METAS some years ago resulting in the so called Geodetic Time Transfer Terminals (GeTT) (a description may be found, e.g., in [1]). One of these devices has been running since October 2003 (GPS station ID: WAB1) and a second one since August 2004 (ID: WAB2) at METAS in Bern, Wabern. It is planned for the near future to submit the data from WAB2 to the International GPS Service (IGS).

At AIUB the capability of the clock estimation was implemented in the Bernese GPS Software package [2]. With the Version 5.0 released last year to the user community a high standard in the analysis of the GPS carrier phase data for time and frequency transfer has been reached. Since that time the development for a continuous time and frequency transfer solution without discontinuities at the day boundaries has been enforced. The algorithms are summarized in Section 2.

One of the motivations for the collaboration between AIUB and METAS is to compare the primary frequency standard at METAS (FOCS-1, [3]) via GPS CP with similar devices of other metrology institutes. For this purpose a frequency transfer is sufficient. The software developments for the continuous geodetic time transfer offer the opportunity to compute a geodetic frequency transfer solution for several days without using the pseudorange measurements at all. In a conventional geodetic time or frequency transfer solution the pseudorange data are analized together with the carrier phase measurements to get access to the time information and to "connect" the independent daily solutions. This new approach is discussed in Section 3.

To demonstrate the successful application of the geodetic frequency transfer using only the phase data, four weeks in autumn 2004 (MJD 53303 to 53330) are selected. A number of IGS stations mainly located at time laborato-

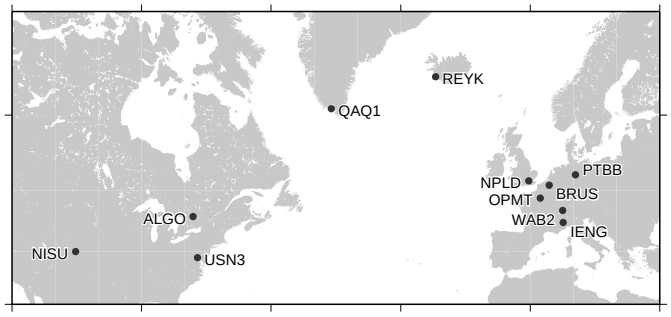


Figure 1: Location of the IGS stations included in the network analysis

ries are analysed in a network solution together with the GeTT data at METAS (see Figure 1). The GPS receivers at BRUS (ORB, Brussels, Belgium), NISU (NIST, Boulder, Co., USA), NPLD (NPL, Teddington, UK), OPMT (OP, Paris, France), USN3 (USNO, Washington, DC, USA), and WAB2 (CH, Wabern, Switzerland) are driven by H-Maser clocks whereas the GPS receivers at PTBB (PTB, Braunschweig, Germany) and IENG (IEN, Torino, Italy) are connected to a Cesium standard. All these sites are equipped with ASHTECH Z-XII3T receivers except NISU.

2 Continuous Geodetic Time Transfer

Using the GPS carrier phase measurements (GPS CP) very accurate results for time and frequency transfer can be achieved because of the low noise of the phase observations. The draw-back is that the estimated receiver and satellite clock corrections are correlated one-by-one with the unknown phase ambiguity parameter. This means that only the change of the clock corrections from epoch-to-epoch can be derived from the phase data with high accuracy. If the phase ambiguity parameters to all satellites in view have to be reinitialized at one epoch independent pieces of high accuracy epoch-to-epoch solution for the clock correction arise. This is, e.g., the case if the receiver loses the lock to all satellites or if the data are processed in independent processing flows (e.g., daily computing batches).

To get access to the absolute value of the clock corrections the GPS pseudorange measurements (code data) are included in the clock estimation with a proper weighting according to their higher noise level. Consequently the accuracy of the geodetic time transfer depends on the mean noise behaviour of the code measurements in the time interval of an independent piece of high accuracy epoch-to-epoch solution mainly obtained from the phase data. Independent pieces of the epoch-to-epoch solution from phase data are connected by the estimated absolute values for clock corrections from the code data for each piece. Therefore, small discontinuities in the time series of the clock corrections may result depending on the noise behaviour of the code measurements. Assuming a daily independent processing scheme discontinuities at the day boundaries have usually a magnitude of up to 1 ns [4].

In fact, the day boundary discontinuities induced by the daily independent processing strategies are one of the main issues limiting the geodetic time transfer method. Several strategies to overcome these discontinuities have been discussed by different groups within the last years: solutions for multiple days in one analysis step, baseline-wise concatenation of daily solutions by applying offsets estimated from overlapping analysed time intervals (e.g., [5]). In both cases important advantages of the geodetic time transfer are lost: a consistent network solution for all contributing receiver and satellite clocks can in fact be computed for a long time interval with a high sampling without any assuptions to the epoch-to-epoch clock characteristics.

At AIUB a number of developments in the Bernese GPS Software have been carried out within the last year to overcome the day boundary discontinuities already in the data analysis. The advantage is that the correlations between the parameters can correctly be considered and that no further steps after the processing are necessary to generate a continuous time series of clock corrections. The algorithms can be applied for a data analysis over a long time span.

Parameter hand-over

If an epoch (e.g., the midnight epoch) is included in the processing of the data from consecutive days the estimates for the parameters at midnight (including the clock corrections) from one day can be introduced directly for the parameter estimation of the following day. This implementation is easy to understand because for each epoch there is exactly one set of clock corrections results (an extended description of the algorithm is given in [6]). There are two disadvantages in this strategy: the continuity depends on the availability of the observations of a single epoch (e.g., the midnight epoch is missing in the datasets from a number of IGS stations) and no data after the midnight epoch contribute to the solution for this particular epoch (an extension of the algorithm that fixes this problem is reported in [7]).

Ambiguity stacking

The main problem when dividing the phase data into independent time intervals for processing is that the continuous phase ambiguity parameters are interrupted at the day boundaries. To reconcatenate these parameters in the analysis is simple in principle but needs a very involved bookkeeping in the analysis software. In addition it is crucial to check that the ambiguity parameters can be connected (no cycle slip event and no internal receiver initialization at midnight). The number of parameters that have to be handled to generate a continuous solution is

much higher than in the parameter hand-over procedure but the ambiguity stacking is more robust with respect to any (non-detected) data problem. A detailed description of the algorithm is given in [7].

3 Geodetic Frequency Transfer

The new developments in the Bernese GPS Software allow to generate a continuous solution across day boundaries without limitation for the processed time interval. If the absolute time information is not required (frequency transfer) the code observations are not necessary at all. Since the ambiguity parameters can be connected at the day boundaries a geodetic frequency transfer solution covering several weeks or months can be generated using phase observations only. The only limitation is the loss of the ambiguities for all satellites in view at one epoch due to a receiver reset or due to the loss of lock to all satellites. The benefits from a phase-only solution are obvious:

- As in the geodetic applications only the phase data of the geodetic receivers are used; depending on the receiver type the quality of the code data varies considerably. The noisy code observations are not introduced into the analysis.
- Inconsistencies between the internal receiver clock for code and phase measurements do not affect the solution; no additional ambiguities for all satellites have to be introduced to restore the consistency in a common analysis.
- In contrast to code observations the impact of multipath and other related effects on the solution is much smaller for phase data. Because of the application of cycle slip detection algorithms the impact on the final solution cannot exceed the half of a wavelength in any case [8].

For most stations in Figure 1 providing the test dataset a continuous geodetic frequency transfer solution without disconnected ambiguity parameters could be generated for the four weeks period analysed here for demonstration (see Figure 2). The interrupts at the IGS station NISU are associated with tracking problems of the receiver.

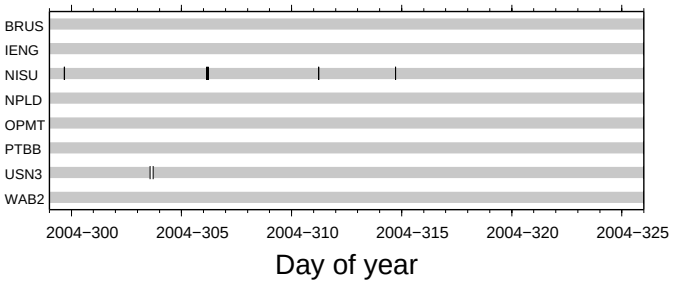


Figure 2: Time intervals continuisly connected by phase ambiguity parameters for the geodetic frequency transfer solution

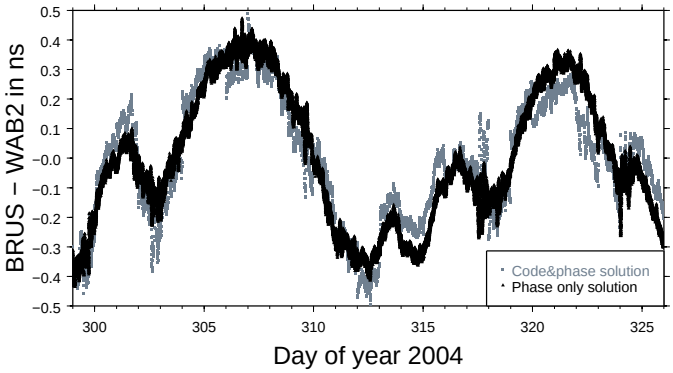


Figure 3: Comparison of different analysis methods for the baseline WAB2 to BRUS. A polynomial of second degree was subtracted for generating the plot.
Gray line: Geodetic time transfer results obtained from daily independent solutions (classical GPS CP analysis)
Black line: Geodetic frequency transfer results obtained from a solution with continuously connected ambiguity parameters using only the phase measurements (new analysis strategy)

Figure 3 shows the geodetic frequency transfer solution (black line) of the baseline from WAB2 to BRUS (distance 487 km) where the ambiguity parameter at the midnight epochs are connected and where only the phase observations are used. For comparison a conventional GPS CP solution (gray line) is superposed. Code and phase data are analysed together in daily independent solutions. The original clock difference time series mainly shows a parabolic behaviour. Because both GPS receivers are driven by H-masers this can be interpreted as a frequency drift of the external clocks. Because the main interest in our study is the behaviour of the frequency transfer analysis methods a polynomial of second degree was subtracted from the computed clock difference time series.

An expected consequence from connecting the ambiguity parameters at the midnight epoch is that the day boundary discontinuities of the daily independent solutions vanish. Whereas the discontinuities for the daily independent geodetic time transfer solution reach values up to 0.3 ns (for other baselines up to 1 ns) they are smaller than 0.01 ns for the geodetic frequency transfer solution using only the phase data. The remaining small discontinuities are caused by the introduced orbits and station positions which stem from independently computed three-day solutions.

In a next step the differences between the time transfer solution from the daily independent, combined analysis of code and phase measurements (classical setup for GPS CP) and the frequency transfer solution using only phase data (new GPS CP analysis) can be computed for each epoch (see Figure 4). Individual mean differences for each day are caused by the mean noise behaviour of the

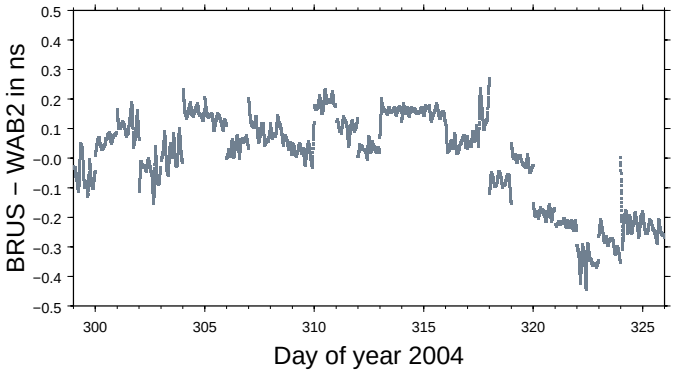


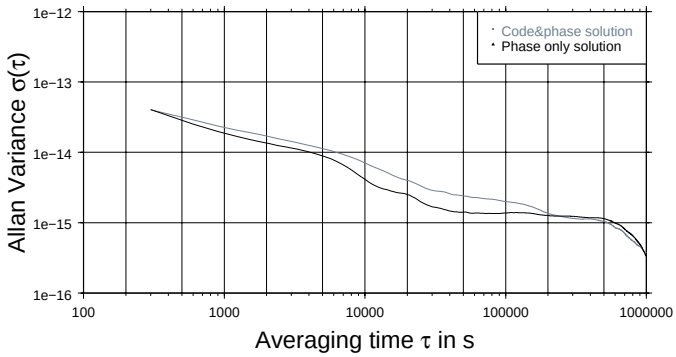
Figure 4: Differences between the daily independent geodetic time transfer solutions from a common analysis of the code and phase data and the geodetic frequency transfer solution with continuously connected ambiguity parameters obtained from phase measurements only for the baseline WAB2 to BRUS.

code measurements of the day in the classical GPS CP analysis. They result in the day boundary discontinuities at the midnigth epochs. In addition a pattern within each day can be seen. The hardware configuration as well as the data are identical for both solutions except that code measurements are used only in one of both solutions. The code data are very noisy compared to the phase measurements. That's why they are introduced with a lower weight into the common data analysis. Nevertheless, the code measurements seem to cause an additional noise of the combined solution with a magnitude of about 50 ps to 0.1 ns. One reason is that code data is much more affected by multipath effects than phase observations.

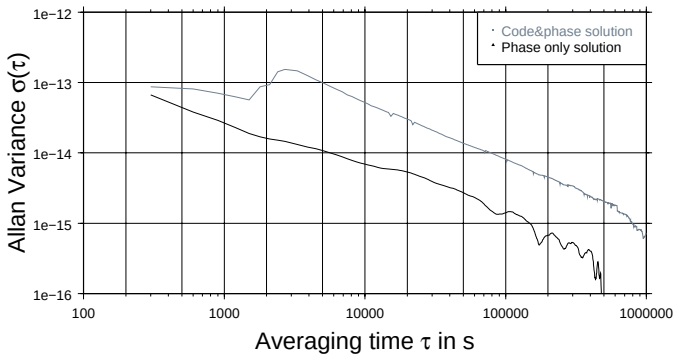
To compare the noise behaviour of the daily independent time transfer solutions based on code and phase data with that of the frequency transfer results obtained with continuously connected ambiguity parameters using only phase measurements the Allan variance for both solution types is computed and shown in Figure 5. Before computing the Allan variance a polynomial of second degree was subtracted.

In Figure 5(a) the Allan variance for the baseline from WAB2 to BRUS is displayed. $\sigma(\tau)$ reaches 10^{-15} after an averaging time of about 30 hours for the daily independent solutions using code and phase data. With the new analysis strategy with connected ambiguity parameters using only phase observations $\sigma(\tau)$ reaches 10^{-15} after an averaging time of about 12 hours. The limit of 10^{-15} for $\sigma(\tau)$ for this baseline can be attributed to the GPS receiver clock driven by H-masers at both sites (see the time series in Figure 3).

In the case of the baseline from WAB2 to NISU (see Figure 5(b), distance 7743 km) $\sigma(\tau)$ reaches 10^{-15} after 100000 seconds but it continues to decrease for longer averaging time intervals. Because of interrupts in the continuity of the connected phase ambiguity parameters for station NISU (see Figure 2) the Allan variance for long time intervals (several days) is not well determined. The



(a) Baseline WAB2 to BRUS



(b) Baseline WAB2 to NISU

Figure 5: Allan variances for the geodetic time transfer solution analysing the code and phase data together (gray lines) and the geodetic frequency transfer solution obtained only from phase measurements (black lines). A polynomial of second degree was subtracted from the time series before computing the Allan variances.

Allan variance for the daily independent, combined code and phase solution is affected by about one hour of corrupted code observations that clock jumps – likely of the internal receiver clock because nothing is known about events at the H-maser in NIST for that epoch – of about minus and plus 10 ns at day of year 306.

4 Code and Phase Consistency

As reported in [6] inconsistencies between code and phase measurement of a geodetic GPS receiver may occasionally occur. As a consequence for geodetic time and frequency transfer, code and phase data refer to different clocks after internal receiver clock events. To recover consistency the ambiguity parameters for all satellites have to be interrupted at the epoch of the internal receiver clock event when analysing code and phase data together.

Geodetic frequency transfer – as described in this paper – is not affected by such events because only the phase measurements are used. On the other hand independent

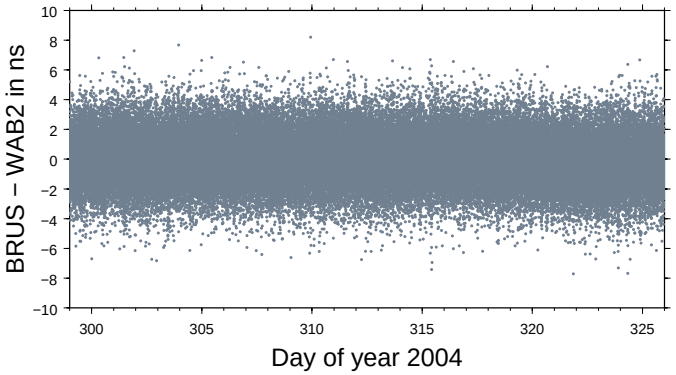
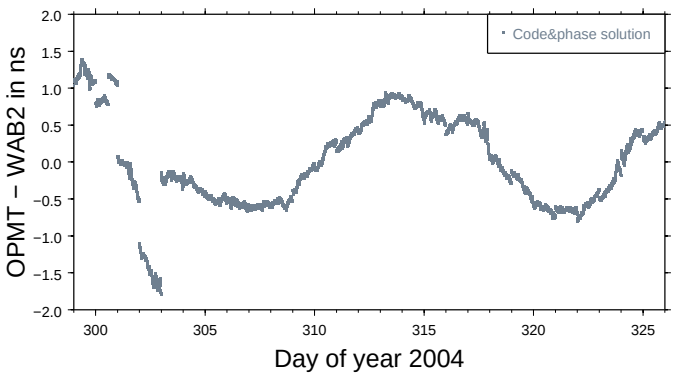


Figure 6: Differences between a phase-only frequency transfer and a code-only time transfer solution for the baseline WAB2 to BRUS. Both solutions are computed with identical geometrical information (satellite orbits, station positions, and troposphere parameters).

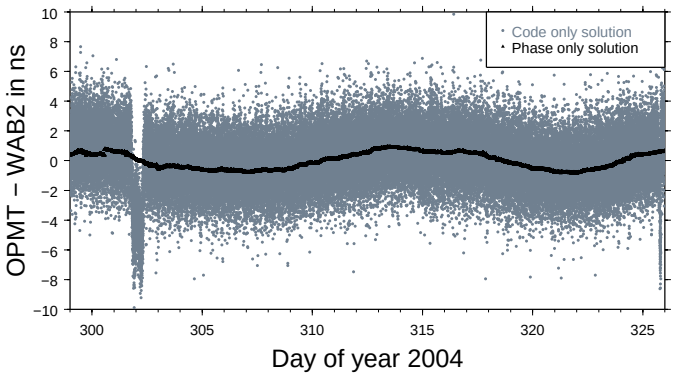
solutions using only the code and only the phase data can be generated. If all geometry information from the phase-only solution is introduced into a code-only solution that only solves for the receiver and satellite clocks. A comparison of both the estimates clock corrections from both solutions offers the possibility to compare the internal receiver clock as it is seen by the code and the phase data. The difference is expected to be a constant offset for all time intervals connected by phase ambiguity parameters. The reason is that the phase-only solution refers to any undefined clock value because of the one-to-one correlation of the phase ambiguity and the clock parameters. The noise in the difference originates from the code-only solution.

In Figure 6 the differences between a phase-only frequency transfer and a code-only time transfer solution are displayed for the baseline WAB2 to BRUS. As expected the graph shows a wide scatter around a constant offset. For both stations no interrupts occured in the phase ambiguity parameters for all satellites in the analyzed time interval (see Figure 2). That's why no discontinuity in the differences can be found. In addition, no significant trend in the offset can be detected which confirms the consistency of the code and phase measurements during the four weeks for this baseline.

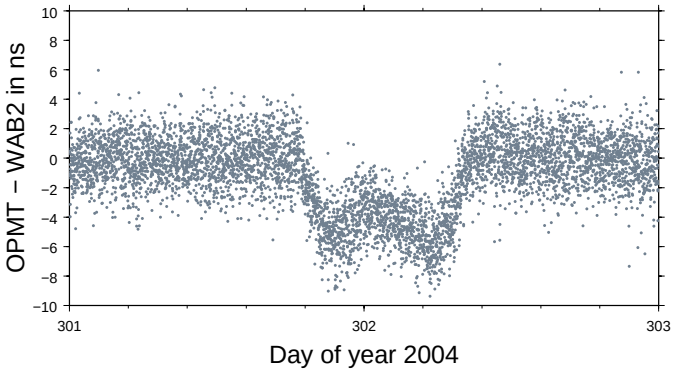
The results of the common analysis of the code and phase measurements in daily independent solutions for the baseline from WAB2 to OPMT (distance 438 km) displayed in Figure 7(a) show day boundary discontinuities – as expected – below 0.5 ns for most of the midnight epochs. Three large jumps can, however, be observed: −1.0 ns at the midnight epoch for day of year 301, −0.6 ns for 302, and +1.6 ns for 303. The phase ambiguity parameters for these midnight epochs are connected and thus the phase-only frequency transfer solution is continuous for this time interval (black line in Figure 7(b)). On the other hand the code-only time transfer solution shows two discontinuities with a magnitude of about ±5 ns at the epochs 301.8



(a) Geodetic time transfer solution using the code and phase measurements in a common data analysis. A polynomial of second degree was subtracted for plotting.



(b) Geodetic frequency transfer solution using only the phase measurements with connected ambiguity parameters (black line) and a corresponding code-only time transfer solution (gray dots). A polynomial of second degree was subtracted for plotting.



(c) Differences between a phase-only frequency transfer and a code-only time transfer solution. Both solutions are computed with identical geometrical information (satellite orbits, station positions, and troposphere parameters).

Figure 7: Comparison of different solution types for the baseline WAB2 to OPMT

and 302.3 which can clearly be seen in the differences between the code-only and the phase-only solution shown in

Figure 7(c). No event in the hardware configuration at OP is known to have occured at these epochs. Because of the continuous change in the difference between the code-only and the phase-only solution environmental effects influencing multipath effects are a likely cause In fact, code measurements are much more sensitive to multipath than phase observations. Figure 7(a) shows that also the results from a combined analysis of the code and phase measurements may be affected. It is therefore preferable to use only the phase measurements for the geodetic frequency transfer whenever possible.
Receiver clock events leading to an inconsistency between code and phase data were not detected in the analysed four weeks of data. A number of millisecond jumps (see [6] for more details) are detected in the measurements of stations NPLD (day of year 300 and 317) and OPMT (day of year 300) that are corrected in the phase observations.

5 Summary

With the latest developments of the Bernese GPS Software concerning the estimation of receiver and satellite clock corrections it is possible to connect the phase ambiguity parameters at the boundaries of independently analysed daily data pieces. As long as the receiver does not reinitialize the phase observations counter and no interruption in the observations to the satellites for any other reasons occur, time series of clock corrections with continuous ambiguity parameters can be generated for several weeks or even months. If absolute time information is not needed a geodetic frequency transfer solution can be computed for such a time interval without using the code measurements at all.
In contrast to the classical geodetic solution based on a common analysis of code and phase data the geodetic frequency transfer solution using only the phase observations has no day boundary discontinuities, a lower noise level in the estimated time series of receiver clock differences, and a lower sensibility to multipath effects. It is therefore preferable to use only the phase measurements for the geodetic frequency transfer whenever possible.
For the remote comparison between Cesium Fountain Primary Frequency Standards between four European National Metrological Institutes and NIST in November 2004 the geodetic frequency transfer solution was computed using only the phase measurements applying the analysis strategy described in this paper. In [9] the solution is compared with the results obtained from other independent time transfer techniques (in particular TWSTFT and GPS P3).

Acknowledgements

The AIUB hosts the Center for Orbit Determination in Europe (CODE), which is a collaboration between the Astronomical Institute at the University of Bern (AIUB), the Federal Office of Topography (Swisstopo, Switzerland), the Federal Agency for Cartography and Geodesy (BKG, Germany), and the Institut Géographique National (IGN, France). The results presented in this paper benefit from the work of the IGS analysis centre CODE.

References

[1] G. Dudle, F. Overney, L. Prost, T. Schildknecht, and T. A. Springer, "First Results on a Transatlantic Time and Frequency Transfer by GPS Carrier Phase," in *Proceedings of the 1998 PTTI meeting*, (Reston, Virginia, U.S.A), 1998.

[2] U. Hugentobler, S. Schaer, and P. Fridez, eds., *The Bernese GPS Software Version 4.2*. University of Berne: Astronomical Institute, February 2001.

[3] A. Joyet, G. Mileti, P. Thomann, and G. Dudle, "Continuous Fountain Cs Standard: Stability and Accuracy Issues," in *Proceedings of the 6th Symposium on Frequency Standards and Metrology* (P. Gill, ed.), pp. 273–280, 2002.

[4] J. Ray and K. Senior, "IGS/BIPM pilot project: GPS carrier phase for time/frequency transfer and timescale formation," *Metrologica*, vol. 40, pp. 270–288, 2003.

[5] C. Bruyninx and P. Defraigne, "Frequency Transfer Using GPS Code and Phases: Short- and Long-Term Stability," in *Proceedings of the* 31st *Precise Time and Time Interval Systems and Application Meeting PTTI 99*, (Dana Point, California), December 7–9 1999.

[6] R. Dach, G. Dudle, T. Schildknecht, and U. Hugentobler, "Status Report of the AIUB–METAS Geodetic Time Transfer," in *Proceedings of the* 18th *European Frequency and Time Forum EFTF 04*, (Guildford, UK), 2004.

[7] R. Dach, T. Schildknecht, U. Hugentobler, L. guy Bernier, and G. Dudle, "Continuous Geodetic Time Transfer Analysis Method," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. submitted, 2004.

[8] R. B. Langley, "Propagation of the GPS Signals," in *Lecture Notes International School "GPS for Geodesy"*, Delft, The Netherlands: Springer-Verlag, 1996.

[9] A. Bauch, R. Dach, J. Davis, L. Lorini, T. Perker, G. Petit, and J. Richerd, "Time and frequency comparisons between four European timing institutes and NIST using mutiple techniques," in *Proceedings of the* 19th *European Frequency and Time Forum EFTF 05*, (Besançon, France), March 21–24 2005.

Session 3C / Cryogenic oscillators

Simple Model for the Mode Splitting Effect in Whispering Gallery Mode Resonators

Vincent Giordano, Pierre-Yves Bourgeois, Nicolas Bazin and Yann Kersalé

Institut FEMTO–ST, Dpt. LPMO, UMR 6174 CNRS-Université de Franche-Comté
32 av. de l'Observatoire, 25044 Besançon Cedex, France

***Abstract*— In this paper we present a model describing the mode splitting effect in a whispering gallery mode resonator. The whispering gallery mode is represented by a equivalent transmission line ring resonator in which a localized perturbation is introduced. This perturbation induces a coupling between the two counter propagating waves which in turn lifts the degeneracy of the whispering gallery modes. This simple model explains almost all the line shapes experimentaly observed. It shows in a straightforward way how the coupling on the two twin modes are affected by changing the coupling probe and the perturbation angular positions. In that sense it is usefull to make easier the resonator adjustment. Moreover the same approach can be used to simulate any microstrip ring resonator with localized perturbation.**

I. Introduction

Mode degeneracy is common to any cylindrical ring resonator structure. Any geometrical imperfection affecting the resonator cylindrical symmetry lifts this degeneracy and each resonance splits in two modes. This problem is particularely prejudicial to high–Q Whispering Gallery (WG) mode sapphire resonators. Indeed at cryogenic temperature these microwave resonators present Q–factors in excess of 10^7 and a resonnance doublet is observed for every WG mode [1]. In high quality sapphire resonator the mode splitting ranges from few Hz to few kHz changing from a mode to an other. If the distance between the two twin modes is of the order of their bandwidth their line shapes are distorded which can affect the effective Q–factor. Moreover the coupling factors of the two modes obtained at low temperature are hardly pretictible. As even a 10kHz splitting is not observable at room temperature, it is impossible to favour only one mode by adjusting the coupling before cooling. The coupling factors depend on the angular position of the coupling probes and of the dominant perturbation causing the mode splitting. Since the latter is not known, the experimentalist is confronted to a three varying angular parameters problem which can be hardly solved with several iterations.

Although based on very basic assumptions, the model presented in this paper correctly explains the mode splitting and the line shapes experimentally observed. It then contributes to understand the changes in coupling and in line shapes during the resonator adjustment step.

It should be noted that one elegant solution to solve the problem of mode splitting has been proposed in [2]. A localized dielectric perturbation was added to the sapphire resonator. By rotating this perturbation with respect to the coupling structures, it is possible to provide controlled coupling to only mode. Our model provides another representation of this mode selection principle.

Eventually, this model can be applied to microstrip ring resonators which are commonly used to stabilize microwave oscillators, design filters [3], measure dielectric constant of materials or to characterized high–T_C superconducting films [4], [5]

II. Mode splitting effect

A. Model for the WG Mode Resonator

A microwave high–Q WG mode resonator is composed by a sapphire rod machined in a low defect Al_2O_3 single crystal. A resonator diameter R of 30–50mm and a thickness h of 15–30mm are typical for X–band operation. The C-axis of the crystal is parallel to rod axis z within less than 1°. Quasi–TE ($WGE_{m,l,n}$) or quasi–TM ($WGH_{m,n,l}$) can be excited in the rod depending on the coupling structure. The subscripts m,n and l refer respectively to the azimuthal, radial and axial variation of the electromagnetic field components. The usefull modes are characterized by $m \geq 5$ and $l = n = 0$. In this case the mode energy density is confined between the cylindrical dielectric–air boundary $\rho = R$ and the inner caustic surface $\rho = R_C$ (see Fig. 1). Elsewhere the waves are evanescent and the Q–factor is essentially limited by the sapphire dielectric loss. The sapphire rod is generally placed in the center of a metallic cavity which has a low influence on the frequency and on the Q-factor of the high azimutal order WG modes.

In a standing waves WG mode resonator, two counter–propagating waves are allowed to travel in the sapphire rod as guided in a bent waveguide forming a ring. The height of this waveguide is h whereas its width is $R - R_c$ which depends on the caustic radius and then on the considered WG mode. This waveguide ring can be linken to an equivalent transmission line ring (TLR) (see Fig. 2). The length and the characteristic impedance Z_C of the transmission line depends on the WG mode we consider. If necessary these parameters could be

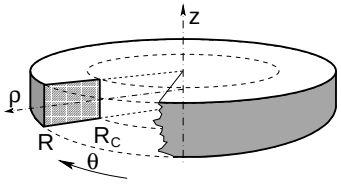


Fig. 1. The Whispering Gallery Mode Resonator seen as a dielectric waveguide ring resonator.

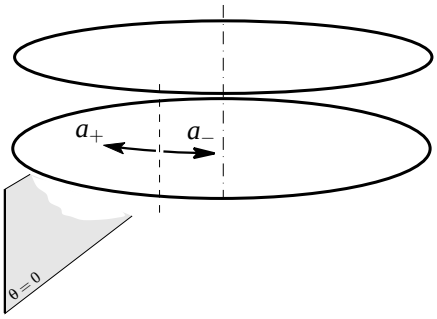


Fig. 2. The Whispering Gallery Mode Resonator seen as a propagation line ring resonator

determined from the waveguide model of the WG resonator [6].

To describe the waves in the resonator the azimutal angle θ is the relevant parameter. The plane $\theta = 0$ corresponds to the location of the resonator coupling element. In this plane two counter–propagating waves are generated with the same amplitude and the same phase. Using the classical scattered waves formalism we define in any plane θ the incident $a_+(\theta), a_-(\theta)$ and the reflected waves $b_+(\theta), b_-(\theta)$. At resonance the phase shift Φ accumulated by the waves during one revolution along the ring should be equal to $2m\pi$. Indeed the incident waves should verify:

$$\begin{cases} a_-(\theta) = a_-(\theta)e^{-j\Phi} \\ a_+(\theta) = a_+(\theta)e^{-j\Phi} \end{cases} \Longrightarrow \Phi = 2m\pi \qquad (1)$$

In a perfect sapphire crystal the two counter–propagating waves $a_-(\theta)$ and $a_+(\theta)$ are completely independant. The position of the nodes of the stationnary wave pattern is only set by the resonator coupling structure located at $\theta = 0$. Along the ring the voltage is :

$$V(\theta) = \frac{a_-(\theta) + a_+(\theta)}{\sqrt{Z_C}} = V(0)\cos m\theta \qquad (2)$$

B. Ring resonator with a localized perturbation

A discontinuity in the propagation line such as a localized change in the waveguide dimensions or in the dielectric constant can be represented by a reactance jX shunting the line (see Fig. 3). A wave approching the discontinuity from either side is then partially reflected and partially transmitted. The two counter propagating waves are now coupled.

Taking $x = X/Z_C$ the normalized reactance value of X, the transmission coefficient T and the reflexion coefficient Γ of

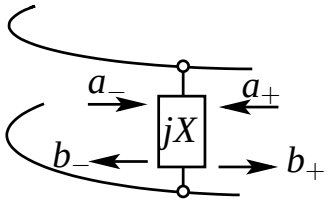


Fig. 3. Perturbation on the propagation line

the perturbation plane are given by:

$$T = \frac{2jx}{1+2jx} \quad ; \quad \Gamma = -\frac{1}{1+2jx} \qquad (3)$$

The equations (1) have to be modified to take into account the coupling between the two counter–propagating waves. The resonace condition is now given by:

$$\begin{cases} a_-(\theta) = Ta_-(\theta)e^{-j\Phi} + \Gamma a_+(\theta)e^{-j\Phi} \\ a_+(\theta) = \Gamma a_-(\theta)e^{-j\Phi} + Ta_+(\theta)e^{-j\Phi} \end{cases} \qquad (4)$$

The resolution of these coupled equations is straightforward. The equation system (4) has a non–trivial solution only if $e^{j\Phi} = T \pm \Gamma$. Then a resonance occurs for two values of the phase shift Φ. Assuming a small perturbation , i.e. $x \gg 1$, the phase shift can be approximated by $\Phi = 2m\pi + \delta\Phi$ where $\delta\Phi \ll 1$. The two resonant modes are then given by:

$$\begin{cases} e^{j\Phi} = T - \Gamma \Longrightarrow \Phi = 2m\pi & \Longrightarrow \delta\Phi = 0 \\ e^{j\Phi} = T + \Gamma \Longrightarrow \Phi = 2m\pi + \frac{1}{x} & \Longrightarrow \delta\Phi = \frac{1}{x} \end{cases} \qquad (5)$$

If ν_m is the $WG_{m,0,0}$ mode frequency for the unperturbed resonator mode, the value of mode splitting $\delta\nu$ due to the presence of a small perturbation is given by:

$$\frac{\delta\nu}{\nu_m} = \frac{\delta\Phi}{\Phi} = \frac{1}{2m\pi x} \qquad (6)$$

The degeneracy is lifted by the perturbation and two resonances separated in frequency by $\delta\nu$ can be observed. The frequencies of the two twin modes are:

$$\begin{aligned} &\text{Mode } \mathrm{WG}^{(1)}_{m,0,0}: \quad \nu_m^{(1)} = \nu_m \\ &\text{Mode } \mathrm{WG}^{(2)}_{m,0,0}: \quad \nu_m^{(2)} = \nu_m\left(1 + \frac{1}{2m\pi x}\right) \end{aligned}$$

A typical mode doublet for a 50 mm diameter, 30 mm thick sapphire resonator at 4.2K is presented in figure 4. The $WGH_{16,0,0}$ mode doublet is characterized by a frequency $\nu_{16} = 11.425\text{GHz}$ and a mode splitting $\delta\nu \approx 24\text{kHz}$. The Q–factor for the two twin modes are of the order of 450×10^6.

A perturbation such as $x = 5 \times 10^4$ is sufficient to induce the 24kHz mode splitting. The modulus of the corresponding reflexion coefficient is $|\Gamma| \approx 1 \times 10^{-4}$ (swr=1.0002) which is a very low value.

The voltage at the discontinuity can easily be calculated. For the first solution for which the resonator phase shift

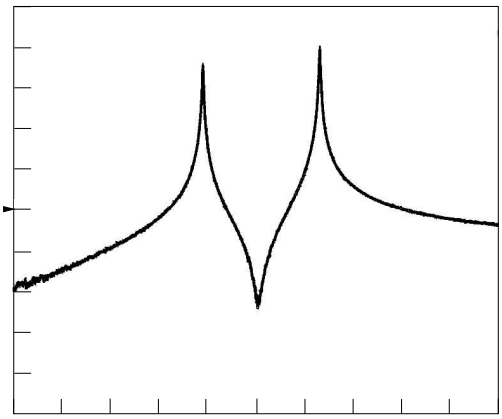


Fig. 4. $WGH_{16,0,0}$ mode doublet for a 50 mm diameter, 30 mm thick sapphire resonator at 4.2K. This plot represents the magnitude of the resonator S_{21} coeffi cient around 11,425 GHz (center frequency). The frequency span is 10kHz. The verticale scale is 10dB/div and the reference value (center of the screen) is -60dB

remains unchanged, i.e. $\Phi = 2m\pi$, there is a voltage node in the plan of the perturbation. Conversely for the second mode, i.e. $\Phi = 2m\pi + 1/x$, there is an voltage antinode at the discontinuity. In the perturbed resonator the position of the standing wave pattern is no longer determined by the coupling system.

III. Full resonator model

A. Resonator coupling and losses

To go further in the description of the WG resonator we have to take into account the energy losses in the resonator itself and in the external circuit and then to define unloaded and loaded Q–factors. Assuming the losses homogeneous in the sapphire rod, the wave attenuation for one revolution is noted α. After one revolution a wave will be then attenuated by a factor $e^{-\alpha}$. The resonator coupling is generally ensured by coaxial probes as small magnetic loops or a straight antennas. In the coupling plane two counter–propagating waves a_1 and a_2 are generated in the sapphire rod (Fig 5A). To be consistent with the TLR description of the WG resonator we replace the actual coupling structure by an equivalent adjacent transmission line feeded by a 180° hybrid coupler (Fig. 5B).

The scattering parameters describing the coupling between the adjacent line and the resonator are deduced from the theory of the directional couplers. The different waves in the coupling plane verify the following equation [7]:

$$\begin{pmatrix} b_e \\ b'_e \\ b_1 \\ b_2 \end{pmatrix} = \begin{bmatrix} 0 & \sqrt{1-k^2} & 0 & jk \\ \sqrt{1-k^2} & 0 & jk & 0 \\ 0 & jk & 0 & \sqrt{1-k^2} \\ jk & 0 & \sqrt{1-k^2} & 0 \end{bmatrix} \times \begin{pmatrix} a_e \\ a'_e \\ a_1 \\ a_2 \end{pmatrix}$$

The parameter k is the voltage coupling coefficient. It depends on the geometry and on the distance separating the dielectric rod and the coupling line. It is now straightforward to represent the resonator and its coupling structure by the signal flow graph given in Fig. 6:

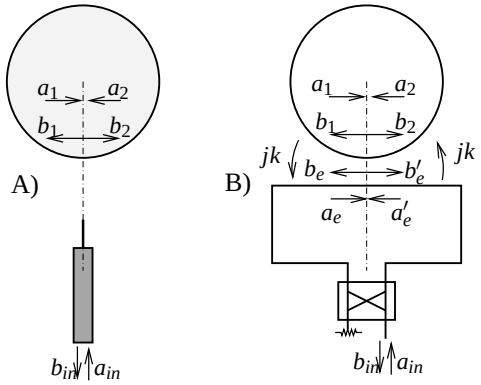


Fig. 5. Modelisation of the resonator coupling

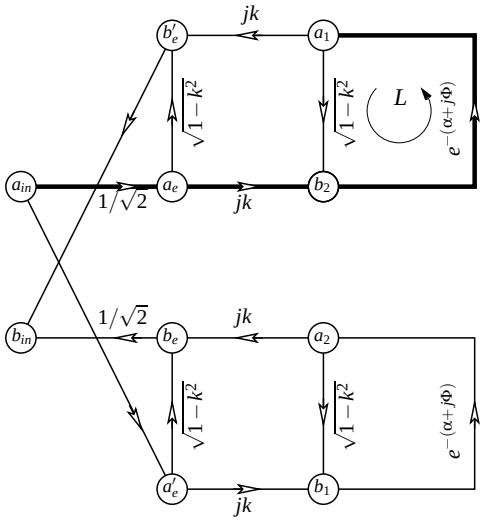


Fig. 6. Equivalent fbw graph of the one port whispering gallery mode resonator. As an example the path P linking the input wave a_{in} to the circulating wave a_1 is depicted in bold line. The loop L represents the recycling of the counter clockwise waves in the resonator.

In this graph the two indenpendant pathes correspond to the two counter–propagating waves travelling indenpendanlty in the rod. By appling the Mason's rule [8], any wave can be determined as a function of the input wave a_{in}. For example, the transmission function from a_{in} to a_1 may be written as :

$$\frac{a_1}{a_{in}} = \frac{P}{1-L} = \frac{jk}{\sqrt{2}} \frac{e^{-(\alpha+j\Phi)}}{1-\sqrt{1-k^2}e^{-(\alpha+j\Phi)}} \qquad (7)$$

where P defines the only path linking a_{in} to a_1. This path is depicted in bold in the figure 6. L is the signal loop describing the recycling of the clockwise travelling waves.

By taking the absolute value squared $\left|\frac{a_1}{a_{in}}\right|^2$ and calculating the half power points, we derived the loaded Q–factor:

$$Q_L = m\pi \frac{\sqrt{\sqrt{1-k^2}e^{-\alpha}}}{1-\sqrt{1-k^2}e^{-\alpha}} \qquad (8)$$

which is consistent with the travelling wave model presented in [9]. By assuming a low loss resonator, i.e α is small, the expressions given both the unloaded and loaded Q–factors

have been derived:

$$Q_0 = \frac{m\pi}{\alpha} \quad ; \quad Q_L = \frac{Q_0}{(1+k^2/2\alpha)} \tag{9}$$

The coupling coefficient to the resonator is then: $\beta = k^2/2\alpha$.

B. Two ports resonator with a localized perturbation

In high stability oscillator applications, the WG mode resonator is generally configured as a two ports resonator implemented in the positive feedback loop of a microwave amplifier. Setting as previously the input coupling probe in the plane $\theta = 0$, we note ψ the angular position of the output coupling probe and φ the position of the localized perturbation. The two port resonator is schematized as shown in figure 7.

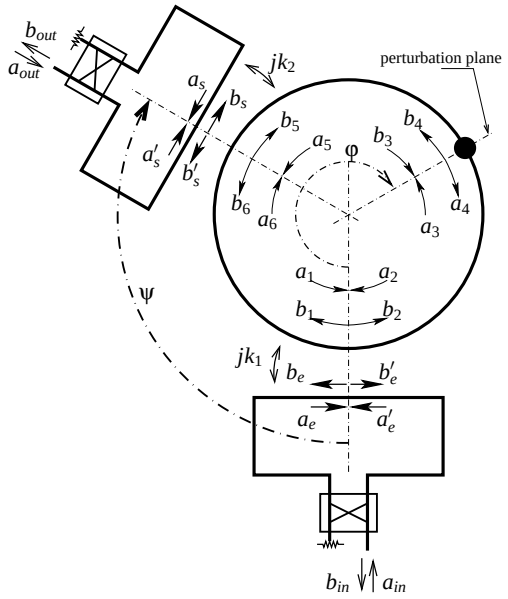


Fig. 7. Full TLR model for the Whispering Gallery Mode Resonator

The equivalent signal flow graph is given in figure 8.

In order to simplify the discussion we asssumed that the impedances connected at the resonator input and output ports are perfectly matched. Then $a_{out} = 0$.

To derive from this flow graph the resonator transmission coefficient $S_{21} = b_{out}/a_{in}$, we have to search for all the different paths linking the input wave a_{in} to the output one b_{out}. Five independant pathes can be identified:

$$P_1 = -\frac{k_1 k_2}{2}\, e^{-(\alpha+j\Phi)\frac{\psi}{2\pi}} \tag{10a}$$

$$P_2 = -\frac{k_1 k_2}{2}\, T\, e^{-(\alpha+j\Phi)(1-\frac{\psi}{2\pi})} \tag{10b}$$

$$P_3 = -\frac{k_1 k_2}{2}\sqrt{1-k_1^2}\,\Gamma\, e^{-(\alpha+j\Phi)(2-\frac{\varphi}{\pi}+\frac{\psi}{2\pi})} \tag{10c}$$

$$P_4 = -\frac{k_1 k_2}{2}\sqrt{1-k_2^2}\,\Gamma\, e^{-(\alpha+j\Phi)(\frac{\varphi}{\pi}-\frac{\psi}{2\pi})} \tag{10d}$$

$$P_5 = -\frac{k_1 k_2}{2}\sqrt{1-k_1^2}\sqrt{1-k_2^2}\,\Gamma^2\, e^{-(\alpha+j\Phi)(2-\frac{\psi}{2\pi})} \tag{10e}$$

Moreover waves recycling occurs in three different loops defined by:

$$L_1 = \sqrt{1-k_1^2}\sqrt{1-k_2^2}T\; e^{-(\alpha+j\Phi)} \tag{11a}$$

$$L_2 = \sqrt{1-k_1^2}\sqrt{1-k_2^2}T\; e^{-(\alpha+j\Phi)} \tag{11b}$$

$$L_3 = (1-k_1^2)(1-k_2^2)\Gamma^2 e^{-2(\alpha+j\Phi)} \tag{11c}$$

By applying the Mason's rules the resonator transmission coefficient S_{21} is calculated:

$$S_{21} = \frac{b_{out}}{a_{in}} = \frac{P_1(1-L_1)+P_2(1-L_2)+P_3+P_4+P_5}{1-L_1-L_2-L_3+L_1L_2} \tag{12}$$

This coefficient can be computed as a function of the phase shift Φ and for different values of the parameters x, α, k_1 and k_2.

IV. Experimental verification

To test the validity of our model we conducted experimental measurements at room temperature on the $WGH_{5,0,0}$ mode of a resonator constituted of a cylindrical sapphire crystal 50mm diameter and 20mm height. This resonator was placed in the center of a copper cavity 100mm diameter and 40mm height. Two coaxial antennas parallel to the cylindrical axis are anchored in each cavity lids to perform the resonator coupling. For the unperturbed resonator this mode is characterized by a frequency of 4.866GHz and a unloaded Q–factor of 290.000. As the actual non homogeneities of the crystal are not sufficient to give an observable mode splitting at room temperature we force a perturbation by gluing a small aluminia pie shapped piece on the sapphire rod as depicted in figure 9. The sapphire crystal and the output probe can be rotated. A dial bearing an angular mark every 5° has been attached to external face of the cavity bottow lid. A pointer attached to the piece mainting the sapphire allows the measure of φ. In this section the angle are given in degres(°) to be consistent with the experimental values.

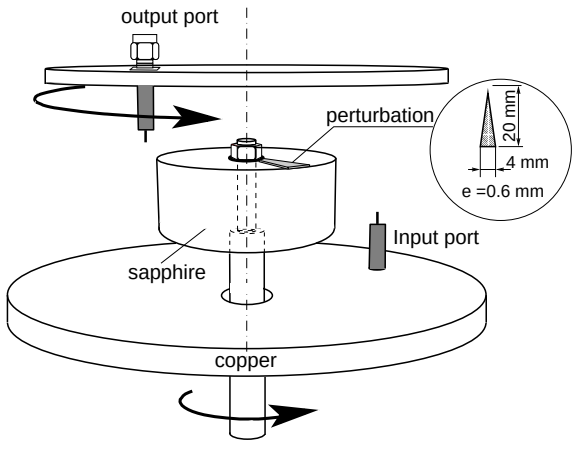


Fig. 9. Experimental set-up. The two angles Ψ and ϕ can be independantly adjusted

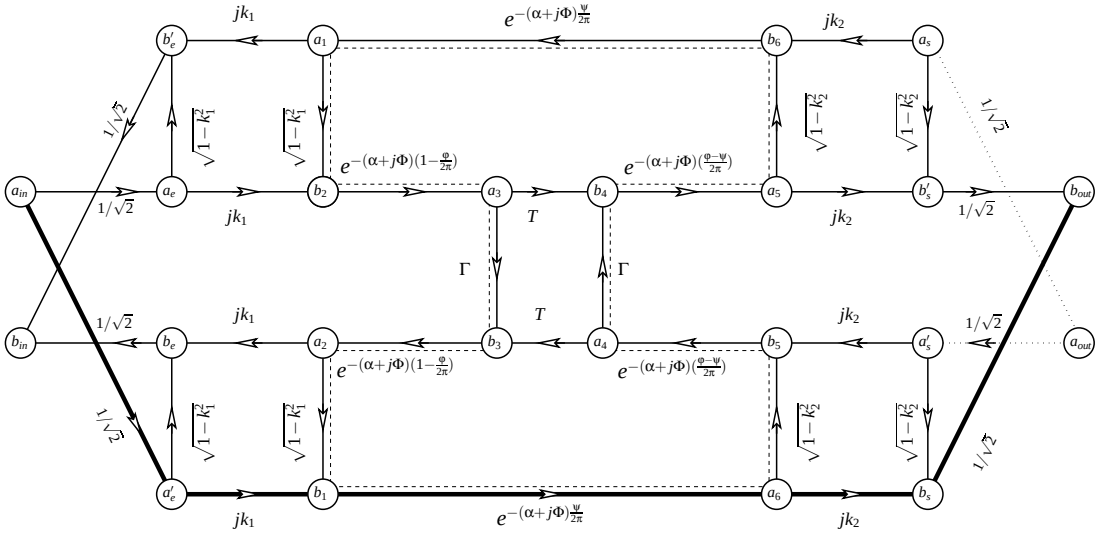


Fig. 8. Signal fbw graph for the full TLR model of the Whispering Gallery Mode Resonator. As an example the bold line represents the path P_1, the dashed line depicts the loop L_3. As we assumed a perfectly matched load at the output port ($a_{out} = 0$), the signal pathes emerging from a_{out} (dotted lines) are not taken into account in the calculation.

Firstly, the resonator without the perturbation was placed in the cavity. We roughly set the two antennas diametrally opposed; i.e. $\Psi \approx 180°$. By rotating the output probe, we adjusted the angle Ψ to minimize the resonator insertion losses for the $WGH_{5,0,0}$ mode. Indeed every WG mode presents an antinode diametrally opposed to the input coupling structure. For Ψ exactly equal to 180° the insertion losses should be minimized. The ouput probe position was marked out and the upper lid removed to install the dielectric perturbation. With the dielectric perturbation in place, the measured frequency separation between the two twin modes is $\delta\nu = 385$kHz. The bandwidth of the two twin modes is 35kHz which corresponds to a loaded Q–factor of 140,000. The resonator coupling coefficients have been set to $\beta_1 = \beta_2 = 0.5$. The unloaded Q–factor is then 280,000 which don't differ notably from the Q–factor of the unperturbed resonator. By applying the equations (6) and (9) we deduced the set of TLR model parameters:

$$m = 5; \qquad \nu_5 = 4.866GHz$$

$$x = 400, \qquad \alpha = 5.6 \times 10^{-5}; \qquad k_1 = k_2 = 0.0075$$

These parameters have been used to compute the scattering coefficients from equation (12) setting $\Psi = 180°$. The comparison between the experimental and theoritical transmission coefficient modules are given in figure 10.

As previously stated, the position of the nodes and antinodes for each mode are determined by the position of the perturbation plane, i.e. angle φ. The antinodes planes for each mode are represented in figure (11) for three positions of the perturbation plane. When $\varphi = 18°$, the two antennas assumed to be diametrally opposed are in the antinode plane of the mode $WGH^{(1)}_{5,0,0}$. In this case, we should only observe one resonance centered on the frequency ν_5. When the perturbation plane is rotated antennas couple the two twin modes. The exact

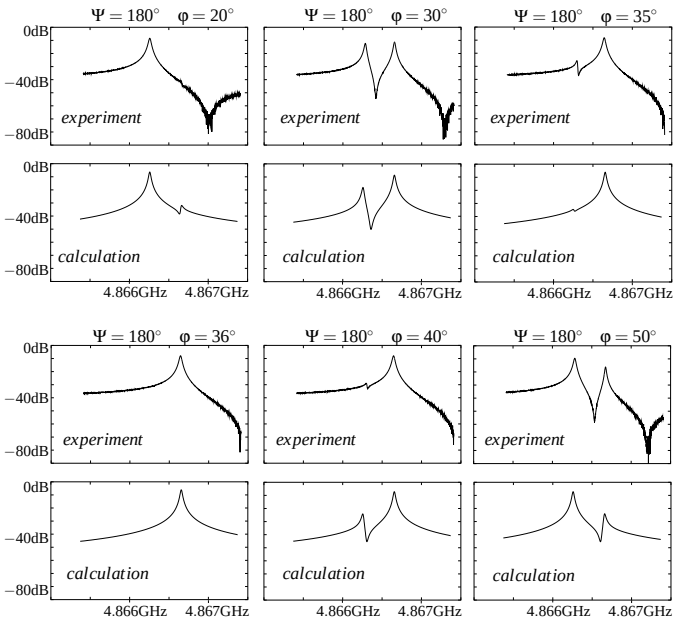


Fig. 10. Comparison between experimental and theoritical S_{21} coeffi cients assuming $\Psi = 180°$.

coupling coefficients for these modes depend on φ. When $\phi = 36°$ the two antennas couple only the perturbed mode at $\nu_5 + \delta\nu$.

That is exactly what we observed when φ was varied from 20° to 50°. We note that for 20° only one mode is strongly coupled. This mode corresponds to the $WGH^{(1)}_{5,0,0}$ mode which is at the same frequency than the $WGH_{5,0,0}$ mode of the unperturbed resonator. When φ is increased the coupling for this mode decreases whereas the $WGH^{(2)}_{5,0,0}$ mode appears. Eventually when 36°, the second mode only

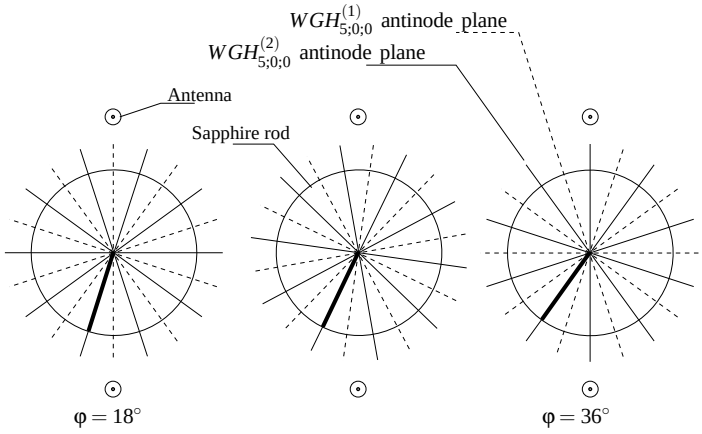


Fig. 11. Rotation of the perturbation plane. The two antennas are assumed to be diametrally opposed: $\Psi = 180°$. Dashed lines: antinode planes for the $WGH^{(1)}_{5,0,0}$ mode at ν_5. Solid lines: antinode planes for the $WGH^{(2)}_{5,0,0}$ mode at $\nu_5 + \delta\nu$.

remains observable. This behavior is well represented by the simulation although some discrepancies exist. Especially for $\varphi = 20°$, the small S–shaped resonance corresponding to the second mode and appearing in the right wing of the main resonance is not correctly represented by the model. Moreover for the other φ values the calcultated relative amplitudes of the two twin modes differ from the experimental data.

Afterwards, we realized that the adjustement method we used to set $\Psi \approx 180°$ was not very efficient. It is not very sensitive with respect to the angle Ψ and moreover the resonator insertion losses can be affected by the presence of spurious modes and by the geometrical defaults of the structure. By making more carefull angle measurements we deduced that the actual angle was $\Psi = 183° \pm 1°$. We then repeated the simulation with $\Psi = 183°$. The calculated spectra better fit with the experiments as demonstrated in figure (12).

V. Conclusion

A simple model for the mode splitting effect in a whispering galery mode resonator has been presented. The dielectric whispering galery mode resonator is represented by an equivalent transmission line ring resonator where a single perturbation is located. This perturbation is responsible for the coupling of the two counter propagating waves allowed to propagate in the ring. This coupling modifies the resonance condition and then each resonator mode splits in a doublet. In our model the resonator coupling structures have been represented by two adjacent transmission lines. The full resonator model has been analysed with the flow graph method to express the quadrupole scaterring parameters as a function of the resonator geometry. A dedicated experiment has been realized to verify the theoretical prediction. Measured and calculated scaterring parameters agree very well demonstrating thus the validity of our assumptions. The principle advantage of our model is its simplicity. The parameters of the model are simply deduced from frequency and Q-factors experimental

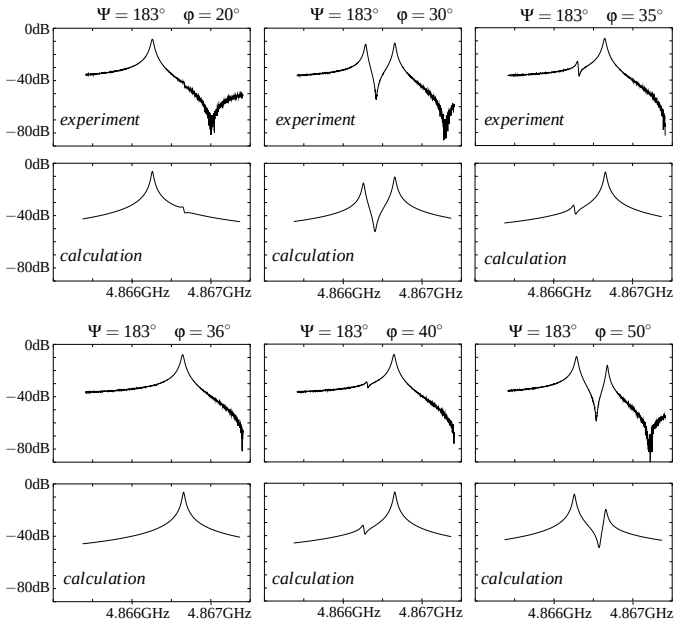


Fig. 12. Comparison between experimental and theoritical S_{21} coeffi cients assuming $\Psi = 183°$.

measurements. By comparing the experimental and theoretical scaterring parameters it is then possible to deduce the angular position of the main perturbation. The resonator can then be rotated to adjust the coupling on one mode of the doublet.

References

[1] P. Bourgeois, Y. Kersal´e, N. Bazin, M. Chaubet, and V. Giordano., "A cryogenic open-cavity sapphire reference oscillator with low spurious mode density.," *IEEE Transactions on Ultrasononics, Ferroelectrics and Frequency Control*, vol. 51, Oct. 2004.

[2] D. Santiago, G. Dick, and A. Prata, "Mode control of cryogenic whispering-gallery mode sapphire dielectric-ring resonator," *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, pp. 52–55, Jan. 1994.

[3] I. Wolff, "Microstrip bandpass fi lter using degenerate modes of a microstrip ring resonator," *Electronics Letters*, vol. 8, pp. 302–33, 15 1972.

[4] J. Bray and L. Roy, "Microwave characterization of a microstrip line using a two-port ring resonator with an improved lumped–element model," *IEEE Transactions on Microwave Theory and Techniques*, vol. 51, pp. 1540–1547, May 2003.

[5] F. C. S.C. Wu, C.Y. Yeh and T. Tseng, "Double resonant behavior of microstrip ring resonator in Tl–based high T_c superconducting fi lms.," *Applied Physics Letters*, vol. 65, pp. 3281–3283, Dec. 19 1994.

[6] E. Marcatili, "Dielectric rectangular waveguide and directional coupler for integrated optics.," *The Bell System Technical Journal*, vol. 48, no. 7, pp. 2071–2102, 1969.

[7] D. Cros and P. Guillon, "Whispering gallery dielectric resonator for w-band devices," *IEEE Transactions on Microwave Theory and Techniques*, vol. 38, pp. 1667–1674, Nov. 1990.

[8] C. Gentili, "Amplifi cateurs et oscillateurs micro-ondes," in *Collection Technologies - Série Electronique*, Paris France: Masson, 1984.

[9] M. Tobar, E. Ivanov, P. Blondy, D. Cros, and P. Guillon, "High–q whispering gallery traveling wave resonators for oscillator frequency stabilization," *IEEE Transactions on Ultrasononics, Ferroelectrics and Frequency Control*, vol. 47, pp. 421–426, Mar. 2000.

Ti^{3+} Doped Microwave Temperature Compensated Sapphire Resonator Oscillator at 34K

N. Boubekeur, J.G. Hartnett†, M.E. Tobar†, N. Bazin, Y. Kersalé and V. Giordano
e-mail: ykersale@lpmo.edu
Dept. LPMO, Institut FEMTO-ST, UMR 6174
32 Ave. de l'Observatoire
F-25044 Besançon cedex, France.
†University of Western Australia, School of Physics Crawley 6009, Dept. of Physics
Nedlands 6907 WA Australia

***Abstract*— Residual paramagnetic impurities are present in high quality sapphire crystal. This phenomenom was exploited as paramagnetic spin compensation at liquid helium temperature to build a high Q Cryogenic Sapphire Resonator Oscillator (CSRO). In this paper we present the characterization of an intentionally Ti^{3+} doped sapphire resonator oscillator. The 12.7 GHz mode operates at 34K and exhibits a frequency stability order 2×10^{-13} for $8s<\tau<20s$ with the use of a Gifford-MacMahon cryocooler as a cold source. This corresponds to state-of-the-art line splitting of the order of 10^{-7}. A second experiment with a Pulse Tube cryocooler shows a short term frequency instability improvement.**

I. Introduction

Over the last few years a lot of work has been done to develop ultra stable microwave sapphire-resonator oscillators at temperatures achievable with commercial cryocoolers or with a solid nitrogen bath. For temperatures above 10K the high frequency-temperature sensitivity of Whispering Gallery Mode Sapphire Resonators (WGMSR) means that the development of thermal compensation techniques are necessary to obtain high frequency stability. Various methods of temperature compensation are described in the literature: mechanical compensation [1], a composite resonator comprising monocrystalline sapphire and rutile [2], a dual mode compensation technique which implements two orthogonal Whispering Gallery (WG) modes [3] and a deposition of a thin film of rutile on sapphire single crystal [4].

In contrast, at liquid helium temperature the presence of residual paramagnetic impurities [5] (a few ppm concentration of $Ti^{3+}, Cr^{3+}, Fe^{3+}, Mo^{3+}, ...$) in high quality sapphire crystal provides temperature compensation of the mode frequencies in the range 5-13K, without the complication of imposing other conditions mentioned above.

A few years ago, Hartnett and Tobar [6] have shown that 0.03 - 0.1 % doping by weight of Ti^{3+} ions in pure monocrystalline sapphire gave frequency temperature compensation in the 20 - 77K range, with quality factors of a few million . The principal advantages of such resonator are the monolithic structure (less sensitive to vibrations) and the small thermal time constant between the sapphire and the perturbation (Ti^{3+} ions).

In this paper we present the characterization of a Ti^{3+} doped sapphire resonator oscillator (as opposed to one with residual impurities). Two configurations are compared : the use of Gifford-MacMahon (GM) and Pulse Tube (PT) cryocoolers as cold sources. The resonator parameters, electronic noise floor and relative frequency stability are also given.

II. Resonator

The resonator consists of a HEMEX Ti^{3+} doped sapphire single crystal (**0.1%** by weight) disk of 25 mm diameter and 20 mm high mounted in a silver plated OFHC copper cavity. Two magnetic loops in the sapphire equatorial plane are used to excite the H_φ component of the operating mode (the quasi–transverse magnetic field $WGH_{8,0,0}$) at 12.7093 GHz. A drawing of the resonator is shown in figure 1.

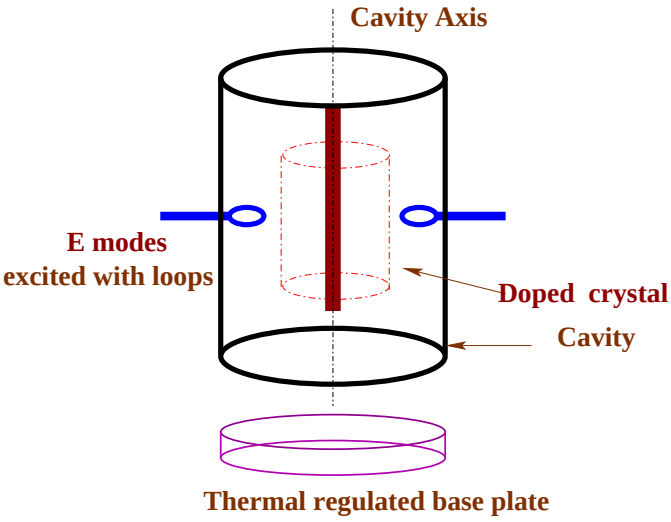


Fig. 1. Schematic diagram of the Ti^{3+} doped sapphire resonator.

Then the resonator was fixed to the cold finger of a GM cryocooler. Heaters and a "Cernox" sensor, with 1mK resolution at 34K, were anchored in the cold finger to enable temperature stabilization using a commercial Proportional Integral Derivate (PID) controller. The cavity was first cooled down to 12K, then its temperature was raised with the heater. Figure 2 shows the evolution of the $WGH_{8,0,0}$ frequency as a function of temperature around its temperature turning point.

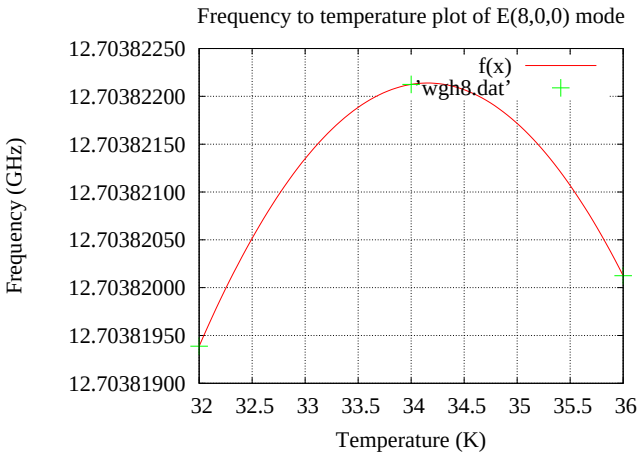


Fig. 2. Frequency to temperature characteristic of the $WGH_{8,0,0}$ around its temperature turning point.

The experimental data have been fit with a second order polynomial. Around the turnover temperature $T_0 = 34.286K$ the resonator frequency temperature sensitivity can be expressed as :

$$\frac{1}{\nu}\frac{\partial \nu}{\partial T} = A\Delta T \qquad (1)$$

where A, expressed in K^{-2}, characterizes the curvature of the frequency temperature plot. For the $WGH_{8,0,0}$ the curvature is estimate to be $A = 6.04 \times 10^{-8} K^{-2}$. Assuming both the ultimate temperature stability and the departure from the turning point are of the order of magnitude of the temperature resolution (i.e 1mK) we get from (1) the limit relative frequency instability achievable with this kind of temperature control which is : $\partial\nu/\nu = 6 \times 10^{-14}$.
At 34K the resonator loaded Q factor was 8.2×10^5, the input coupling coefficient $\beta_1 = 0.93$ and the output coupling coefficient $\beta_2 = 0.01$.

III. Oscillator characterization

A. Oscillator circuit

The temperature compensated resonator stabilized around 34K has been implemented in the oscillator loop shown in figure 3.

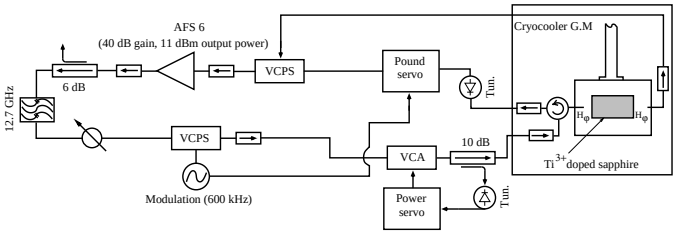


Fig. 3. Schematic diagram of the oscillator based on the Ti^{3+} doped sapphire resonator.

Oscillation on the $WGH_{8,0,0}$ is ensured by a commercial *GaAs* amplifier with 40 dB gain and a band pass filter. Two Voltage Controlled Phase Shifters (VCPS) and a Voltage Controlled Attenuator (VCA) have been inserted in the circuit to control the phase and the amplitude of the signal in the loop. We used two VCPS to realize independently the phase modulation and the electronic control of the oscillator frequency in order to minimize the residual AM modulation generated by non ideal VCPS. The resonator access planes are fixed by a set of three ferrite isolators. A circulator placed at the input resonator access enables extraction of the reflected signal which is detected by a highly-efficient backward diode placed outside the cryocooler.
The loop electrical length is controlled by a Pound frequency stabilization servo by locking the oscillation frequency to the center of the resonance [7]. The carrier is phase modulated by one VCPS at $f_m = 600\ kHz$. The backward diode signal is then demodulated by synchronous detection. The resulting error signal is filtered and sent back to the second VCPS. The noise floor of the frequency stabilization servo has been evaluated considering the experimental slope of the error signal, the noise at the lock-in amplifier input and the backward diode power to voltage conversion efficiency. It leads to a noise floor of $\Delta\nu/\nu = 1 \times 10^{-13} \times \tau^{-1/2}$.
We have also implemented a direct current to prevent frequency instability due to power fluctuations [8]. The resonator relative frequency sensitivity to power has been measured and it is of the order of $8 \times 10^{-10}/mW$. The loop power fluctuations are detected by a backward diode, temperature controlled at 313K by a Peltier element, and corrected by the VCA. The diode and VCA are at room temperature and so power fluctuations in the microwave coaxial copper cables linking the resonator to the loop oscillator are not corrected.

B. Oscillator relative frequency stability

The oscillator signal is mixed with the signal of a microwave synthesizer (Agilent 8254A) referenced to an Hydrogen Maser. The resulting 20 kHz beat-note is sent to a high resolution frequency counter (Agilent 53132A) and then the fractional frequency Allan deviation is computed. The Allan deviation is shown in figure 4.

At one second integration the relative frequency instability is 2×10^{-12}. Between one and 10 seconds the slope seems to be τ^{-1}. For long integration times, $30s \leq \tau$, it clearly appears a frequency drift of the order of $5 \times 10^{-10}/day$. For $10s \leq \tau \leq 30s$ we can see the contribution of the frequency drift and the short term perturbation. The minimun relative frequency instability, 2×10^{-13}, is obtained for $8s \leq \tau \leq 16s$. When the drift is removed the relative frequency instability is of the order of 2×10^{-13} for $8s \leq \tau \leq 200s$. This result is comparable with the ones obtained by using different temperature-compensation techniques [9] and represents a line splitting factor of order 10^{-7}.
The data, with the drift removed, used to estimated the oscillator Allan variance has been computed with FFT software in order to determine the signal power spectral density. The

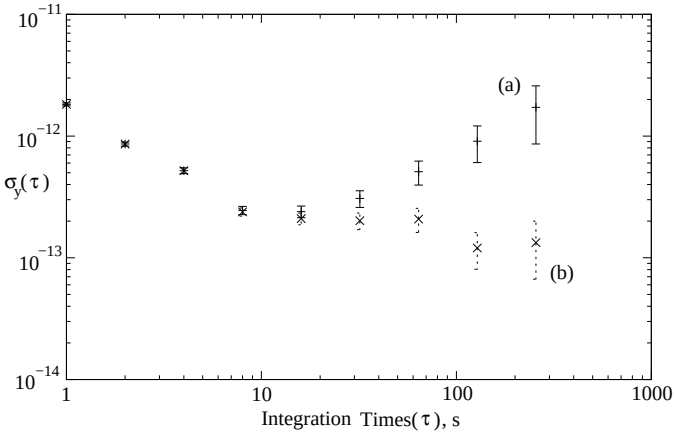


Fig. 4. Relative frequency instability of the Ti^{3+} doped sapphire resonator oscillator in Gifford-MacMahon cryocooler : (a) raw data; (b) drift removed.

result of this computation is represented on figure 5.

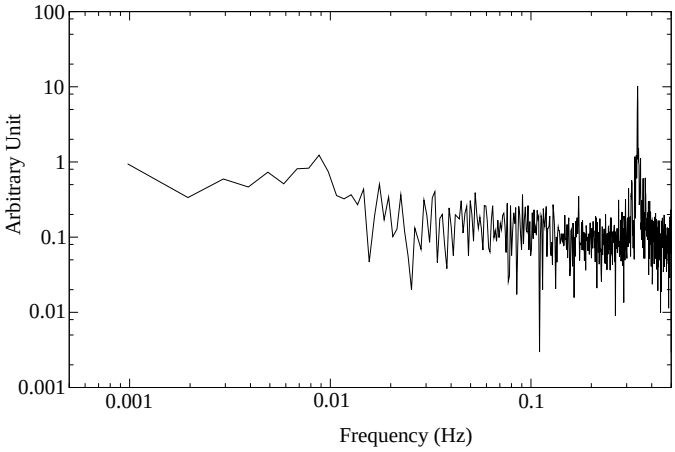


Fig. 5. Power spectral density of the Ti^{3+} doped sapphire resonator oscillator in Gifford-MacMahon cryocooler.

On this plot there clearly appears a big "spike" at 0.34 Hz. This spike, attributed to the GM cryocooler vibrations creates an excess noise for $0.1Hz \leq f \leq 1Hz$. We can conclude that the short term frequency instability of the oscillator is limited by the GM cryocooler vibrations.

C. Pulse Tube cryocooler

As mentioned in the previous section GM cryocoolers generate a lot of vibrations, which limit the short term frequency stability of our oscillator. At FEMTO-ST we recently got a Pulse Tube cryocooler well known to produce less vibrations than a GM cryocooler. But, in preliminary measurement, we find that the temperature of the thermal regulated base plate, fixed on the cryocooler cold finger, was not very stable. Indeed temperature fluctuations of the order of 4mK on the cernox sensor was measured when the PID controller was running. If we use such temperature fluctuations in (1) we get the minimum frequency instability achievable with this kind of temperature control, $\partial\nu/\nu = 6 \times 10^{-13}$, which is not sufficient for our application. We have designed a thermal filter with a 200 seconds of time constant to minimize the effect of the cold finger temperature fluctuations on the thermal regulated base plate. A 86 mm diameter and 11 mm high piece of *Nylon* was then fixed between the cryocooler cold finger and the thermal regulated base plate. The Ti^{3+} doped sapphire resonator was then fixed on the temperature regulated plate and cool down to 34K. The temperature compensated resonator has then been introduced in the same kind of oscillator loop as presented in figure 2. For this experiment coupling to the sapphire resonator and Pound servo parameters was changed a little. We had $\beta_1 = 0.8$, $\beta_2 = 0.001$ and the Pound servo noise floor was estimated to be of the order of $\Delta\nu/\nu = 6 \times 10^{-13} \times \tau^{-1/2}$. The $12.7035GHz$ signal oscillator was then compared to the microwave signal of the Agilent (8254A) referenced to an Hydrogen Maser. On the next plot we represent the standard Allan deviation ($\sigma_y(\tau)$) of this oscillator.

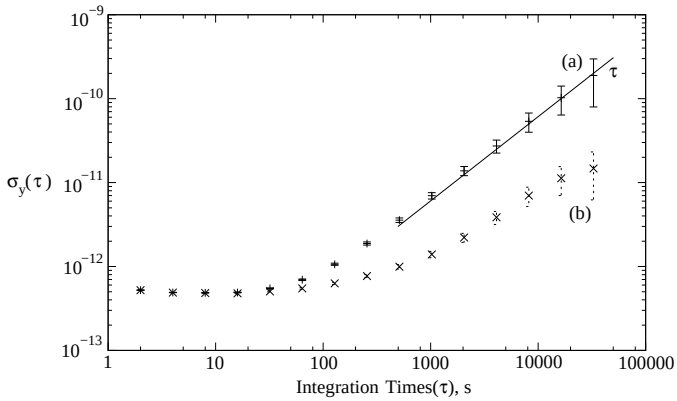


Fig. 6. Relative frequency instability of the Ti^{3+} doped sapphire resonator oscillator in Pulse Tube cryocooler : (a) raw data; (b) drift removed.

Compared to the GM cryocooler experiment this measurement with the PT cryocooler is not disturbed at short integration times. For $\tau \leq 32s$ we have reached a flicker floor of 5×10^{-13} representing a line splitting factor of 4×10^{-7}. The obtained flicker floor at short integration times can be attributed to the power servo control. This power servo has been used for a liquid helium temperature sapphire resonator oscillator. The relative frequency stability measurement of that oscillator showed a flicker floor of $\Delta\nu/\nu = 2 \times 10^{-14}$ for a relative frequency to injected power sensitivity of $4 \times 10^{-11}/mW$. It means that this power servo controls the power fluctuations of the loop in the $0.5 \times 10^{-6}W$ level. If we consider this level of power control with the relative frequency to injected power sensitivity of the Ti^{3+} doped sapphire resonator ($8 \times 10^{-10}/mW$) it leads to a flicker floor of $\Delta\nu/\nu = 4 \times 10^{-13}$ which is consistant with our measurement. This quite high noise floor of the power servo is due to the effects of the room temperature fluctuations on the DC level of the backward diode used in the power servo. To remove this effect the backward diode will be temperature controlled in the cryogenic environnement in the future.

At long integration times the oscillator frequency exhibits the same frequency drift as was observed in the GM cryocooler experiment (i.e $5 \times 10^{-10}/day$).
The drift removed data have been also computed in the FTT software to confirm that the short term relative frequency instability in the GM cryocooler was limited by the mechanical vibrations. The result of this computation is represented on the next plot.

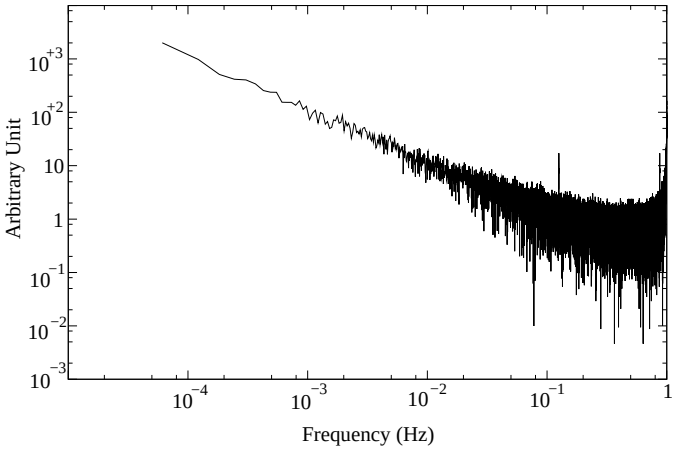


Fig. 7. Power spectral density of the Ti^{3+} doped sapphire resonator oscillator in Pulse Tube cryocooler.

As we can see on this plot, by using a pulse tube cryocooler, the spike has disappared on the spectrum and so the oscillator short term frequency instability is no more disturbed at short integration times. This confirm the hypothesis made on section 3.B.

IV. Conclusion

We have demonstrated in this paper the capability of Ti^{3+} temperature compensated sapphire resonator oscillator to achieve a frequency stability of 2×10^{-13} for $8s \leq \tau \leq 16s$ limited by the GM cryocooler vibrations at short integration times and 5×10^{-13} at short integration times with the PT cryocooler with an unloaded resonator Q factor of only 1.6 million. These results correspond respectively to a line splitting factor of 1.6×10^{-7} and 4×10^{-7}. In both case the oscillator long term frequency stability is limited by the resonator ageing leading to a frequency drift of $5 \times 10^{-10}/day$.
In 1999 J.G. Hartnett et al [10] have shown that quasi transverse electric mode (i.e WGE) excited in Ti^{4+} doped sapphire resonator exhibits unloaded Q factor better than 5 millions at temperature turnover above $30K$. Moreover WGE modes are less sensitive to temperature fluctuations than WGH modes [11]. Then if we are able to get the same line splitting factor with the Ti^{4+} doped sapphire resonator a relative frequency instability below 10^{-13} is achievable. This Ti^{4+} doped sapphire resonator will be tested in the future.

References

[1] R.T. Wang and G.J. Dick, “High stability 40 Kelvin cryo-cooled sapphire oscillator,” Proc. of the 2003 IEEE International Frequency Control Symp., Tampa Bay, Fl, USA, May 2003, pp. 371–375.

[2] M.E. Tobar, J. Krupka, J.G. Hartnett, E.N. Ivanov and R.A. Woode,“High Q sapphire-rutile frequency-temperature compensated microwave dielectric resonators,” *IEEE Trans. on UFFC*, 1998, **45**, (3), pp. 830–835.

[3] M.E. Tobar, G.L. Hamilton, E.N. Ivanov and J.G. Hartnett,“New method to build a high stability sapphire oscillator from the temperature compensation of the difference frequency between modes of orthogonal polarization,” *IEEE Trans. on UFFC*, 2003, **50**, (3), pp. 214–218.

[4] Y. Kersalé, S. Vives, C. Meunier and V. Giordano, “Cryogenic monolithic sapphire-rutile temperature compensated resonator oscillator,” *IEEE Trans. on UFFC*, 2003, **50**, (12), pp. 1662–1666.

[5] A.G. Mann, A.J. Giles, D.G. Blair and M.J. Buckingham, “Ultra-stable cryogenic sapphire dielectric microwave resonators: mode frequency-temperature compensation by residual paramagnetic impurities,” *J. Phys. D.: Appl. Phys.*, 1992, **25**, pp. 1105–1109.

[6] J.G. Hartnett, M.E. Tobar, A.G. Mann, J. Krupka and E.N. Ivanov, “Temperature dependence of Ti^{3+} doped sapphire whispering gallery mode resonator,” *Electronics Letters*, 1998, **34**, (2), pp. 195–196.

[7] A. Luiten, A.G. Mann and D.Blair, “Cryogenic saphire microwave resonator oscillator with exceptional stability,” *Electronics Letters*,1994, **30**, pp. 417–418.

[8] A.N Luiten, A.G. Mann, M. Costa and D.Blair, “Power stabilized cryogenic saphire oscillator,” *IEEE Transactions on Microwave Theory technics*,1995, **44**, pp. 132–135.

[9] J.G. Hartnett, P.Y. Bourgeois, J.D. Anstie, M.E. Tobar, N. Bazin, E.N. Ivanov, D. Cros, V. Giordano and Y. Kersalé , “High Q frequency-temperature compensated solid-nitrogen-cooled resonator-oscillators: first results,” *Electronics Letters*, 2004, **40**, (1), pp. 41–42.

[10] J.G. Hartenett, Michael E. Tobar, “Frequency-Temperature Compensation in Ti^{3+} and Ti^{4+} Doped Sapphire Whispering Galery Mode Resonators,” *IEEE Transactions on UFFC*, 1999, **46**, (4), pp. 993–1000.

[11] M.E. Tobar, E.N Ivanov, R.A. Woode, J.H Searls, “Low noise microwave oscillators based on high-Q temperature stabilized sapphire resonators,” Proc. of the 1994 IEEE International Frequency Control Symp., Boston, MA, USA, 1994, pp. 433–440.

OBSERVATION OF A BISTABILITY EFFECT IN A CRYOGENIC MICROWAVE WHISPERING GALLERY MODE RESONATOR

P.-Y. Bourgeois, N. Bazin, Y. Kersalé, V. Giordano,

and M. Oxborrow†

Institut FEMTO–ST, Dpt. LPMO, UMR 6174 CNRS-Université de Franche-Comté
32 av. de l'Observatoire, 25044 Besançon Cedex, France

†National Physical Laboratory, Queens Road, Teddington, Middlesex TW11 0LW, UK

***Abstract*— We report the first observation of maser oscillation on a whispering-gallery (WG) mode in an electromagnetic resonator. The quasi-transverse-magnetic WG mode, with a loaded (non-magnetic) Q of approx. 4.5×10^8, is defined by a cylinder of mono-crystalline sapphire, held at a temperature at or slightly above 4.2 K. The mode's frequency is approx. 12.038 GHz, which coincides with that of a paramagnetic resonance in Fe^{3+} ions, that are substitutively included within the sapphire's lattice at a concentration of a few parts per billion. By applying c.w. 'pump' power to the resonator at a frequency of approx. 31.339 GHz, a population inversion between the lowest two paramagnetic levels of the Fe^{3+} ion is obtained with no applied d.c. magnetic field. Despite the low concentration of Fe^{3+} ions within the sapphire, continuous maser self-oscillation is achieved on the whispering-gallery mode due to its high Q. Our first measurements demonstrate a frequency stability (Allan deviation) of a few times 10^{-13} for sampling intervals up to 1000 s.**

I. Introduction

Short-term fractional frequency stabilities better than 1×10^{-14} have only ever been achieved at microwave frequencies with oscillators incorporating cryogenic (sub 10 K) resonators with Q values in excess of 100 million [1]. Such oscillators have been used successfully as ultra-frequency-stable 'flywheels' for cold-atom frequency standards [2], as reference oscillators for (close-in) phase-noise measurements [3], or in fundamental-physics experiments [4]. In each realization reported to date, the resonator has functioned as an entirely passive, linear (either 1- or 2-port) device, as can be modelled with scattering parameters (e.g. S_{21}).

In recent years, the most actively studied sorts of resonators have been those that support whispering-gallery (WG) modes, defined by dielectric cylinders or rings of mono-crystalline sapphire [5]–[7]. With these, the resonator is maintained at a temperature lying conveniently above 4.2 K, where a frequency-versus-temperature turnover point is located. The turnover point can be precisely sat upon through a temperature-control system employing resistive heaters, with no need to pump on the helium bath. An oscillator in a Pound-stabilized-loop (PSL) configuration [8] is built around the resonator, with all of its active (i.e. power-consuming) components, including the amplifier(s), located outside of the cryostat at room temperature.

The PSL oscillator is a spatially extended system; at least two microwave lines, each typically > 1 m in length, are required to connect the cryogenic resonator to the room-temperature section of the oscillator's loop. In reality, each line comprises a series of semi-rigid microwave cables that are connected by feedthroughs between the cryostat's different chambers and sections. To achieve stabilities at the 1×10^{-14} level, several auxiliary cables and sensors, supporting the control of the resonator's temperature, the Pound servo, loop-power regulation [9], ... all need to be wired into the cryostat. Pound stabilization does suppress, within a finite bandwidth, the phase (thus frequency) shifts associated with mechanical vibrations and/or temperature fluctuations along the microwave lines. Residual amplitude modulation, causing offsets in the Pound servo's d.c. error signal remains a problem, however [10].

In this paper, we report the observation of maser [11] oscillation in a cryogenic sapphire resonator that is simply excited with c.w. microwave ('pump') power. Such a WG-mode maser oscillator, or 'WGMO', offers several immediate and significant advantages over the PSL configuration described above: (i) the system comprises far fewer critical components; (ii) there simply is no Pound stabilization; (iii) the oscillator's amplifier is located in a stable cryogenic environment; (iv) the number and/or complexity of the remaining auxiliary control circuits can be reduced (at least in less critical applications). Furthermore, the setting up of a PSL oscillator, where the cryogenic resonator needs to be appropriately coupled both in reflection and in transmission, is still not a wholly straightforward task. The couplings required for optimal

maser oscillation are, in contrast, far less critical.

Almost half a century ago, a number of seminal research works were undertaken in the field of electron paramagnetic (or spin) resonance (i.e. 'EPR', or 'ESR') in solids [12], and in particular towards the development of solid-state masers [11], [13]. Sapphire crystals doped with Cr^{3+} ions (i.e. ruby), and also Fe^{3+} ions [14], [15], were studied. A few maser amplifiers and oscillators were developed as research instruments/prototypes using the latter ion [16]–[18]. Since the technological goal for this research was the realization of wideband low-noise amplifiers for satellite communications and astronomy, as opposed to frequency standards *per se*, maser oscillators with crystals containing paramagnetic ions at extremely low (parts-per-billion) levels of concentration, within resonators exhibiting the extremely high ($> 1 \times 10^7$) Q values necessary to reach maser threshold were, to the best of our knowledge, never realized.

II. Description of Experimental Work

Our active electromagnetic resonator, shown in figure 1, comprises a monocrystalline sapphire (HEMEX) cylinder, 50 mm in diameter and 30 mm high, mounted coaxially within a copper cavity, whose interior cylindrical wall is silver-plated.

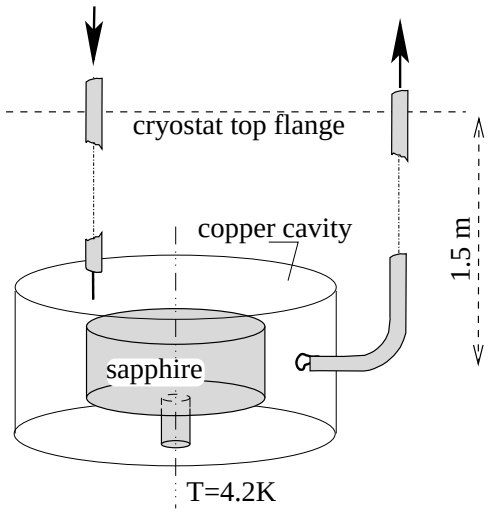


Fig. 1. *Experimental set-up*

Two coaxial probes couple to a quasi-transverse-magnetic mode of whispering-gallery type within the resonator. This 'WGH' mode is dielectrically guided by and predominantly confined within the sapphire cylinder; its azimuthal mode-order is 17 and it has no radial or axial nodes; let us refer to it henceforth as the $WGH_{17,0,0}$ mode. The copper cavity is mounted within a vacuum can on the end of the cryostat's insert, which is loaded into a large liquid-helium dewar. Microwave transmission lines comprising lengths of semi-rigid RG-405 cable connect the resonator's probes to connectors on the insert's top flange; each line is approx. 1.5 m long.

The resonator was first cooled to the base temperature of the liquid-helium cryostat (i.e. approx 4.2 K), with no active temperature control. We then observed the resonator's transmission spectrum around the $WGH_{17,0,0}$ mode with a network analyzer (N.A.), whose output reference signal was set to a power level of -20 dBm; see figure 2.

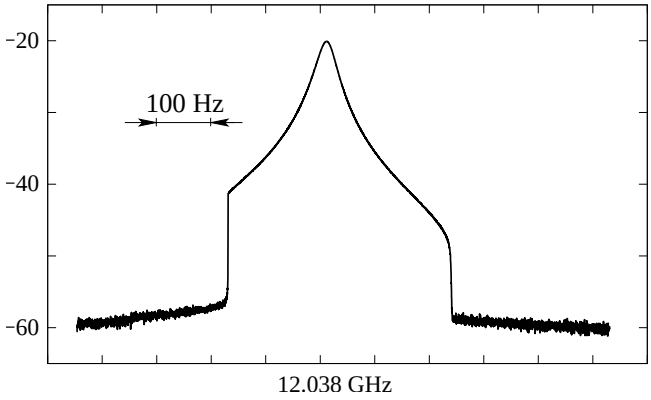


Fig. 2. *Transmission coefficient S_{21} of the $WGH_{17,0,0}$ mode at 12.038 GHz. The full span is 1kHz.*

Note that this spectrum departs dramatically from the Lorentzian lineshape expected for a 'linear' (single-pole) resonance. Next, by operating the N.A. in c.w. mode (with the interrogating frequency held fixed at the $WGH_{17,0,0}$ mode's line centre), the resonator's transmission coefficient S_{21} was recorded as a function of the applied reference signal's power level (with the resonator still at 4.2 K); see figure 3.

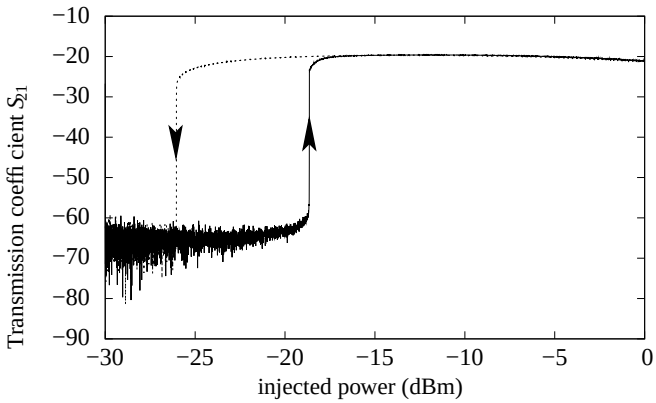


Fig. 3. *Anomalous behavior of the resonator's S_{21} coefficient as a function of the injected power. The full line has been recorded when the power was increased from -30 dBm to 0 dBm. For the dashed line, the power has been swept from 0 dBm to -30 dBm.*

Depending on the direction of the power sweep, one of two thresholds is observed. At low applied powers, the $WGH_{17,0,0}$ mode cannot be observed. It then suddenly appears when the power is raised above a threshold of -18.6 dBm. When the applied power is reduced from a high level, the mode remains observable down to a

second threshold at -26 dBm, upon falling below which it suddenly vanishes.

Though the sapphire monocrystal was not intentionally doped, one can reasonably expect that it contains paramagnetic Fe^{3+} ions as residual impurities at a low concentration, substituting for Al. These ions exhibit an ESR at a frequency of approx. 12.04 GHz, corresponding to the $|1/2\rangle \leftrightarrow |3/2\rangle$ transition shown in figure 4; at 4.2 K, this ESR has an expected linewidth of a few tens of MHz [**?**]. The observed phenomenon of bistability can only be explained by the saturation of the Fe^{3+} ion's ESR at 12.04 GHz in conjunction with the $WGH_{17,0,0}$ mode's extremely high (non-magnetic) *Q*. The phenomenon is wholly equivalent to the behaviour of an optical cavity loaded with a saturable absorber, for which bistability effects are well known [19]. To use the technical vernacular, our system's 'cooperativity' (a dimensionless, characterizing parameter) is sufficiently large to cause bistability [20]. By fitting the relevant so-called state equation to the hysteresis loop shown in figure 3, we estimate the concentration of the paramagnetic Fe^{3+} impurities to be approx. 2.4 parts per billion.

The Fe^{3+} ion's paramagnetic energy levels at zero d.c. magnetic field are represented in figure (4). Each of these three levels is in fact a degenerate Kramers doublet [12], [16]. At liquid-helium temperature, there are significant differences in the populations of these levels, in accordance with the Maxwell-Boltzmann law. Transitions between any of these three levels are allowed, though they vary significantly in strength [15]; their linewidths are all expected to be a few or several tens of MHz (again see ref. [?]).

Our whispering-gallery-mode maser oscillator (WGMO) exploits Bloembergen's classic three-level scheme [13], though note at zero applied magnetic field. $|1/2\rangle \leftrightarrow |3/2\rangle$ transitions at 12.04 GHz are stimulated by the resonator's $WGH_{17,0,0}$ mode, whose frequency lies well within the transitions' linewidth. With no pump applied to the resonator, these stimulated transitions extract net power from the mode –above and beyond the losses associated with its regular, 'non-magnetic', loaded *Q*. Application of c.w. pump power at approx. 31.29 GHz, however, causes a net transfer of Fe^{3+} ions from the $|1/2\rangle$ level to the $|5/2\rangle$ level. The population of the $|3/2\rangle$ level is in turn increased through non-radiative transitions. At sufficient pump power, a population inversion is obtained between the $|3/2\rangle$ and $|1/2\rangle$ levels. The $WGH_{17,0,0}$ mode then *receives* power from the $|1/2\rangle \leftrightarrow |3/2\rangle$ transitions that it stimulates. If this received power can match the losses associated with the mode's loaded *Q*, i.e. if the WGMO exceeds the threshold for active masing, sustained maser oscillation on the $WGH_{17,0,0}$ mode will result.

C.w. microwave pump power from a microwave synthesizer was injected into the resonator through one of its probes. The resonator exhibited many different

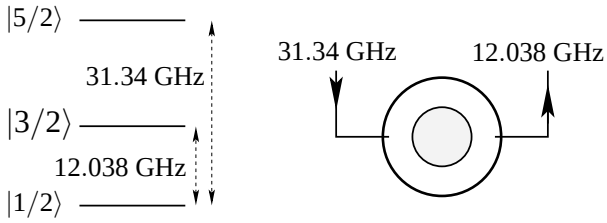


Fig. 4. *The Fe^{3+} ion energy level diagram at zero magnetic field and the maser principle.*

modes within the anticipated 50 MHz or so linewidth of the $|1/2\rangle \leftrightarrow |5/2\rangle$ transition at 31.29 GHz. We tested a few of them (they no doubt had different couplings to the pump probe and also different overlaps with the $WGH_{17,0,0}$ mode ...) and found experimentally that a -56 dBm maser signal at 12.038 GHz could be obtained with the pump generator (viz. an Agilent E8254A synthesizer) set to 31.339 GHz and an output power level of 2 dBm. The successful pump mode has yet to be identified though, based on its sharp frequency dependence, it is suspected to also be a whispering-gallery mode.

III. Preliminary performance

The whispering-gallery-mode maser oscillator's output signal (at the cryostat's top plate) was amplified by 70 dB then mixed (in a DBM) with the signal from a second microwave synthesizer (a Wiltron 69137A) referenced to a hydrogen maser. The resulting $\approx$ 90 kHz beat–note was sent to a high-resolution frequency counter (an HP 53132A).

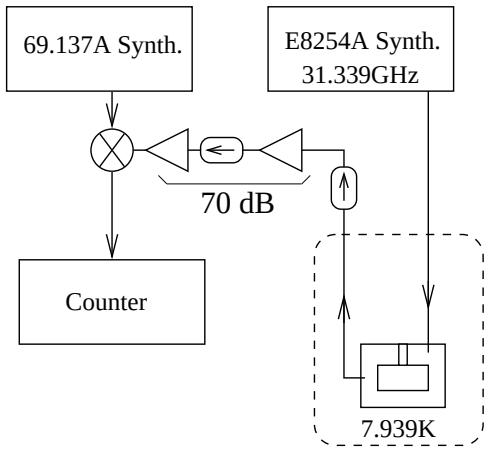


Fig. 5. *Measurement set-up*

By slowly increasing the resonator's temperature whilst recording the beat–note's frequency, we observed that the WGMO's frequency had a turn-over temperature at 7.939 K. The resonator was stabilized at this temperature, and then the beat–note measured against time; the

corresponding fractional-frequency Allan deviation was subsequently computed, with the result given in figure (6).

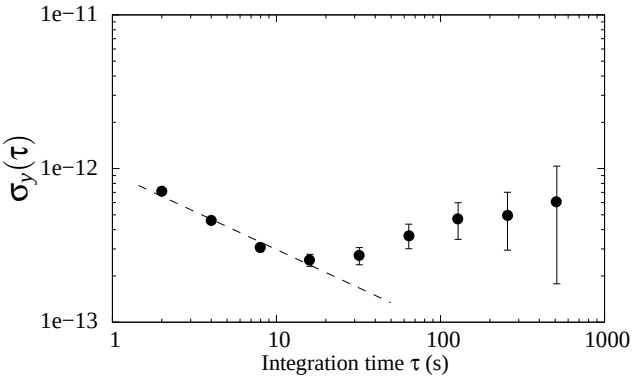


Fig. 6. *Frequency stability of the WG Maser Oscillator ; the dashed line is the intrinsic instability of the reference.*

In the same figure, we show the synthesizer's own instability as a frequency reference, which limited the measurement resolution for sampling intervals $\tau < 20$ s. For greater sampling intervals $\tau < 100$ s, the WG maser is characterized by a frequency stability better than 4×10^{-13}.

It should be noted that the microwave lines within the cryostat were not at all designed to operate above 18 GHz. Between the feedthrough connector on the insert's top flange and the corresponding connector that feeds the resonator's pump probe at the bottom of the insert, we measured before cooling down an insertion loss of 13 dB with a 31 GHz test signal. Allowing for a 1 dB reduction in the cable loss upon cooling, we crudely estimate that the forward power received by the resonator's pump probe to be no more than -10 dBm. For lack of appropriate K_A-band equipment, the power reflected back from the pump probe, hence the coupling to the 31.339 GHz pump mode, could not be determined. Moreover, no isolators (for either X-band or K_A-band) were placed in any of the microwave lines between the resonator and the top plate within the insert. Fluctuations in the standing-wave pattern in these lines could thus have significantly 'pulled' the WGMO's output frequency.

Assuming a spin-lattice relaxation time ('T_1') of around 7 ms [15], we estimate that the measured -56 dBm level of maser signal power is consistent with a concentration of Fe^{3+} ions at the part-per-billion level (it is doubted that the $|1/2\rangle \leftrightarrow |5/2\rangle$ transition was fully saturated at the 2 dBm level of applied pump power used in the reported experiment.) The origin of the long-term degradation in the frequency stability has yet to be positively determined. We plan to evaluate the WGMO's sensitivities to pertinent experimental variables including the pump power and frequency, the degree of signal- and pump-mode coupling, the VSWRs in relevant sections of transmission line (due to reflections from microwave components and feedthroughs), tilt, and either a deliberately applied and/or the incidental magnetic field.

REFERENCES

[1] A. G. Mann, "Ultrastable cryogenic microwave oscillators," in *Frequency measurement and control : advanced techniques and future trends* (A. N. Luiten, ed.), Topics in Applied Physics, v. 79, pp. 37–66, Springer, 2000.

[2] A. G. Mann, G. Santarelli, S. Chang, A. N. Luiten, P. Laurent, C. Salomon, D. G. Blair, and A. Clairon, "A high stability atomic fountain clock using a cryogenic sapphire interrogation oscillator," in *Proc. of 1998 IEEE International Frequency Control Symposium*, (Pasadena, CA, USA), pp. 13–17, 1998.

[3] G. J. Dick and D. G. Santiago, "Microwave frequency discriminator with a cryogenic sapphire resonator for ultra-low phase noise," in *Proc. of 6th European Frequency and Time Forum*, (Noodwijk, The Nederlands, 17-19 March 1992), pp. 35–39, 1992.

[4] P. Wolf, S. Bize, A. Clairon, A. N. Luiten, G. Santarelli, and M. Tobar, "Tests of lorentz invariance using a microwave resonator," *Phys. Rev. Lett.*, vol. 90, p. 060402, 2003.

[5] S. Chang, A. G. Mann, and A. N. Luiten, "Improved cryogenic sapphire oscillator with exceptionally high frequency stability," *Electron. Lett.*, vol. 36, no. 5, pp. 480–481, 2000.

[6] G. J. Dick, R. T. Wang, and R. L. Tjoelker, "Cryo-cooled sapphire oscillator with ultra-high stability," in *Proc. of 1998 IEEE International Frequency Control Symposium*, (Pasadena CA, USA), pp. 528–533, 1998.

[7] P.-Y. Bourgeois, Y. Kersal´e, N. Bazin, M. Chaubet, and V. Giordano, "A cryogenic open-cavity sapphire reference oscillator with low spurious mode density," *IEEE Trans. Ultrason. Ferroelectr. Freq. Contr.*, vol. 51, no. 10, pp. 1232–1239, 2004.

[8] A. J. Giles, S. K. Jones, D. G. Blair, and M. J. Buckingham, "A high stability microwave oscillator based on sapphire loaded superconducting cavity," in *Proc. of 43rd Annual Symposium on Frequency Control*, (Denver Marriot Hotel, Denver), pp. 89–93, IEEE, 1989.

[9] A. N. Luiten, A. G. Mann, N. J. McDonald, and D. G. Blair, "Latest results of the u.w.a. cryogenic sapphire oscillator," in *Proc. of IEEE International 49th Frequency Control Symposium*, (San Francisco, CA , USA), pp. 433–437, 1995.

[10] G. Marra, M. Oxborrow, and D. Henderson, "Reduction of residual amplitude modulation in pound-locking schemes for ultra-stable oscillators," in *Proc. of Proc. of 18th European Frequency and Time Forum*, (University of Surrey, Guildford, UK), IEE, 2004.

[11] A. E. Siegman, *Microwave Solid-state Masers*. McGraw-Hill, 1964.

[12] J. W. Orton, *Electron Paramagnetic Resonance*. London : Iliffe Books Ltd, 1968.

[13] N. Bloembergen, "Proposal for a new type solid state maser," *Phys. Rev.*, vol. 104, no. 2, pp. 324–327, 1956.

[14] L. S. Kornienko and A. M. Prokhorov, "A paramagnetic amplifi er and generator, using fe3+ ions in corundum," *Sov. Phys. JETP*, vol. 36, no. 9, p. 649, 1959.

[15] G. S. Bogle and H. F. Symmons, "Paramagnetic resonance of fe3+ in sapphire at low temperatures," *Proc. Phys. Soc.*, vol. 73, no. 3, pp. 531–532, 1959.

[16] J. E. King and R. W. Terhune, "Operation of a zero-fi eld x-band maser," *J. Appl. Phys.*, vol. 30, pp. 1844–1845, 1959.

[17] G. E. Friedman and A. W. Nagy, "Iron sapphire maser with no magnetic fi eld," *Proc. IEEE*, vol. 51, pp. 361–362, 1963.

[18] W. E. Hughes, "4c5 – a zero-fi eld maser oscillator," *IEEE J. Quantum Electron.*, vol. QE-2, no. 8, pp. 233–235, 1966.

[19] A. E. Siegman, *Lasers*. Mill Valley, CA : University Science Books, 1986.

[20] L. A. Lugiato, "Theory of optical bistability," in *Progress in Optics* (E. Wolf, ed.), vol. XXI, pp. 71–211, Elsevier, 1984.

ROTATING MICHELSON-MORLEY EXPERIMENT BASED ON A DUAL CAVITY CRYOGENIC SAPPHIRE OSCILLATOR

P.L. Stanwix*, M.E. Tobar, M. Susli, C.R. Locke, E.N. Ivanov, J. Winterflood, J.G. Hartnett, F. van Kann
School of Physics, The University of Western Australia, Crawley, Australia

P. Wolf
BNM-SYRTE, Observatoire de Paris, Paris, France
Bureau International des Poids et Mesures, Sevres, France

ABSTRACT

Recent experiments based on cryogenic microwave [1,2,3] and optical [4,5] oscillators have tested the isotropy of the speed of light (Michelson-Morley experiment) at sensitivities of the order of a part in 10^{15}. Further improvements of the accuracy in this type of experiment are not expected due to the already long data set and systematic error limits [2]. We have constructed a new rotating Michelson-Morley experiment consisting of two cylindrical cryogenic sapphire resonators. By rotating the experiment our sensitivity to light speed anisotropy is increased, as the relevant time variations are at the rotation frequency where the frequency stability of the cryogenic oscillators is the best. Preliminary data has been analyzed using the frameworks of Robertson, Mansouri and Sexl (RMS) and the Standard Model Extension (SME), and the results are presented.

1 INTRODUCTION

The Einstein equivalence principle (EEP) is at the heart of special and general relativity [6] and a cornerstone of modern physics. One of the constituent elements of EEP is Local Lorentz invariance (LLI) which, loosely stated, postulates that the outcome of any local test experiment is independent of the velocity of the freely falling apparatus (a fundamental hypothesis of special relativity). The central importance of this postulate in modern physics has motivated tremendous work to experimentally test LLI [6].

Numerous test theories that allow the modeling and interpretation of experiments that test LLI have been developed. Kinematical frameworks [7,8] postulate a simple parameterization of the Lorentz transformations with experiments setting limits on the deviation of those parameters from their special relativistic values. Also, a general Lorentz violating extension of the standard model of particle physics (Standard Model Extension, SME) has been developed [9] whose Lagrangian includes all parameterized Lorentz violating terms that can be formed from known fields.

Experiments that test the isotropy of the speed of light (Michelson-Morley experiment) have placed upper limits on any violation of LLI to a few parts in 10^{15} [1,3,4] and generally rely on the Earth's rotation. To this point reductions in the upper limit of LLI have been enabled by advances in the stability of atomic clocks and macroscopic resonators. Further improvements may be achieved by actively rotating the experiment in the lab frame [10], which is being undertaken in this work. The advantage of rotating the experiment is two-fold: the data integration rate is increased by three orders of magnitude and the relevant signals are translated to the cryogenic sapphire oscillator's optimal operating regime.

2 METHODOLOGY

A schematic of the experiment is shown in Fig. 1. The experiment can be separated into 3 main elements: the cryogenic sapphire oscillators, the rotation system and the data acquisition system.

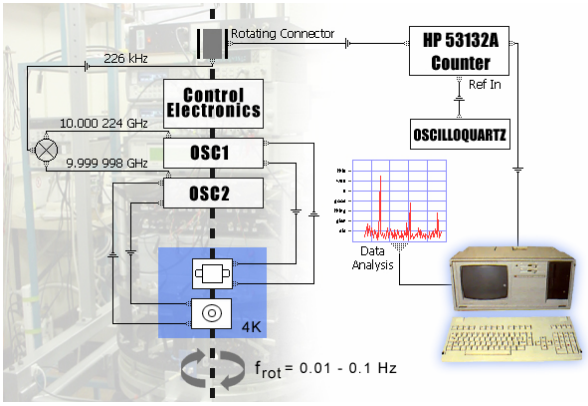


Figure 1. Experimental overview

2.1 The cryogenic sapphire oscillators

This Michelson-Morley experiment is based on two cylindrical cryogenic sapphire resonators. The sapphire cylinders each have a diameter of 31.60 mm and height of 30.05 mm. They have a spindle at either end and are supported by both spindles within superconducting niobium cavities of internal diameter 49.80 mm and height 48.90 mm. The resonators are oriented on their

sides orthogonal to each other and perpendicular to the normal to the Earth's surface. Both resonators are excited in the whispering gallery $WGH_{8,0,0}$ mode by separate oscillators with frequencies of approximately 10.000224 GHz and 9.999998 GHz at 4.2 K. Straight antenna probes mounted in each lid couple microwave power to the resonators.

The mounting structure for the resonators is shown in Fig. 2. The structure is placed inside a small stainless steel vacuum cylinder, sealed with Mylar gaskets. The small vacuum cylinder is in turn sealed inside a large stainless steel cylinder that is supported by the insert inside the dewar. The double vacuum provides thermal isolation to the sapphire loaded cavities, and avoids sapphire contamination from microwave components located between the two cylinders. The cavities are thermally connected to the liquid helium bath via a copper post that joins a stainless steel post connected to a triangular plate. This plate is bolted to three copper heat sinks which protrude into the liquid helium bath. The stainless steel post provides some thermal filtering of bath temperature fluctuations. A foil heater and a carbon-glass temperature sensor are attached to the copper post just above the cavity between the two vacuum cylinders to control the set point of operation, with mK stability.

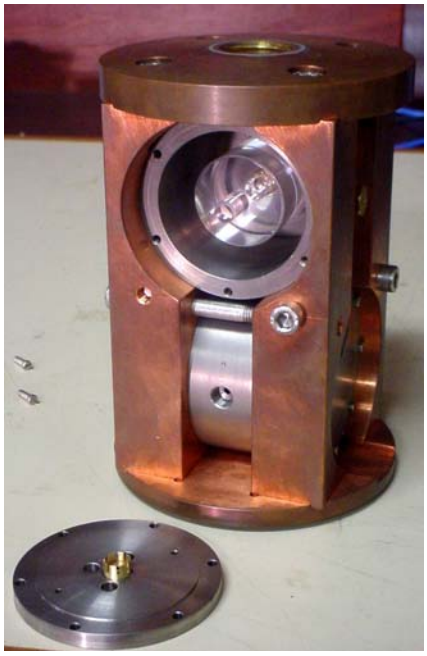


Figure 2. The two resonators are positioned orthogonal to each other in the mounting structure. One of the sapphires can be seen mounted inside the superconducting niobium cavity. The spindles are firmly held in each lid by sprung brass bushes.

2.2 Rotation System

The rotation system is shown in Fig. 3. The cryogenic dewar containing the insert is suspended within a ring bearing and hangs below the floor level of the laboratory. To avoid flexing of cables in the oscillator loop, the oscillator circuits are mounted on the suspended dewar, along with the control electronics. A detailed description of oscillator circuit operation and control systems can be found in [11].

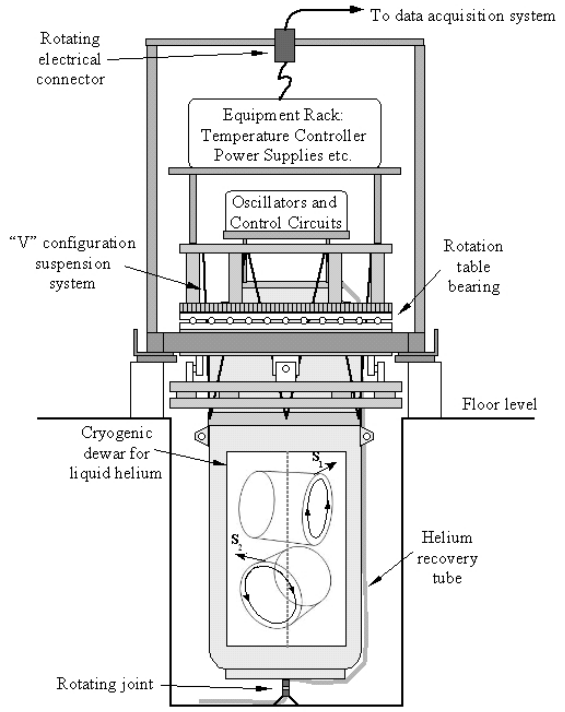


Figure 3. Rotation system for the experiment. Inside the dewar a schematic of the two orthogonally oriented resontors is shown, along with the Poynting vectors of propogation S_1 and S_2.

The suspension arrangement used is multiple "V" shaped loops of elastic shock cord, which avoids high Q-factor pendulum modes by ensuring that the cord has to stretch and shrink (providing damping losses) for horizontal motion as well as vertical. A table mounted on the bearing carries peripheral equipment such as the temperature controller and power supply. The rotation system is driven by a microprocessor controlled stepper motor, with the ability for the rotation to be automatically controlled by computer.

A rotating connector located at the bottom of the dewar allows the boil off from the liquid helium bath to be collected and re-liquefied. This connector is mounted at the bottom of the pit on a scissor action bracket, allowing the connector to move vertically with the dewar as the helium boils off. A combination of two rotating electrical connectors is located above the experiment. A commercial 18 conductor slip ring connector with a hollow through bore transfers electrical power to the rotating experiment and monitor signals to the data acquisition system. The second mercury based coaxial rotating connector is mounted inside the larger slip ring connector in the same axis of rotation. The sole purpose

of the connector is to transmit the difference signal to the stationary frequency counter.

2.3 Data Acquisition System

The data acquisition system logs the dual-oscillator difference frequency as well as monitoring systematic effects including the temperature of the resonators, liquid helium bath level, ambient room temperature, oscillator control signals and tilt. The difference frequency is measured using a HP53131A frequency counter, configured for minimum dead time, and communicated to a computer using GPIB. The data acquisition is triggered by the orientation of the experiment, determined from a series of infra-red LEDs located around the circumference of the rotation table.

3 RESULTS

3.1 Turning Point Power Dependence

The excellent performance of cryogenic sapphire oscillators is made possible by the existence of a turning point in the temperature dependence of the resonator frequency. The turning point results from the effect of paramagnetic impurities on the sapphire's susceptibility, and the changing penetration depth of the microwave cavity [12,13]. In this experiment we are able to manufacture a turning point in the difference frequency by careful selection of the microwave power incident on each resonator. Measurements have shown that the turning point temperature of each individual resonator is dependent on its incident microwave power, with a relatively small effect being observed in the other resonator. An example is given in Fig. 4. This power dependence has not been observed in previous work with single cavity cryogenic sapphire oscillators [12,13,14,15], suggesting that it may be due to temperature gradient effects between the sapphire and the cavity, enhanced by the poor thermal contact between the two. A plot of the frequency turning points is shown in Fig. 5.

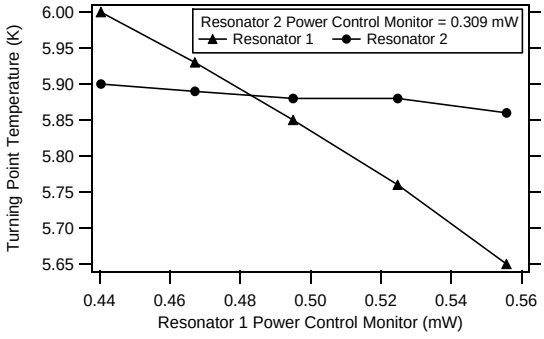


Figure 4. The effect of incident microwave power on the frequency turning point temperature. The power control monitor is proportional but not equal to the power incident on the sapphire.

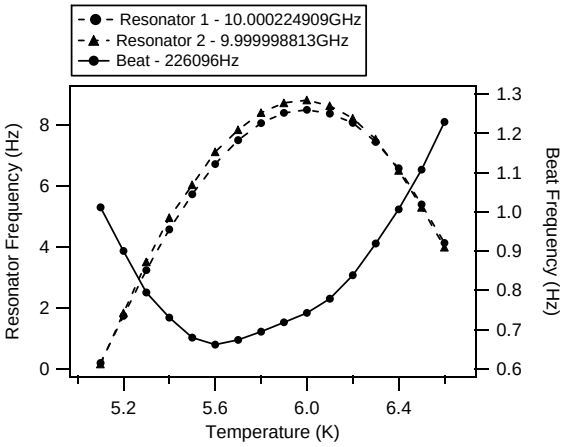


Figure 5. The frequency temperature turning points of the individual resonators and the difference frequency, measured simultaneously.

3.2 Frequency Stability

Typical frequency stability results are shown in Fig. 6. Systematic effects associated with the rotation of the experiment can be seen up to 20 seconds. For the Lorentz Invariance tests these systematics play an important role and need to be investigated and well understood. However, to evaluate the intrinsic stability the systematics were removed from the data using a least squares approach. Comparison with the raw data suggests this is the source of the short term instabilities, as the typical noise floor of a cryogenic sapphire oscillator is recovered.

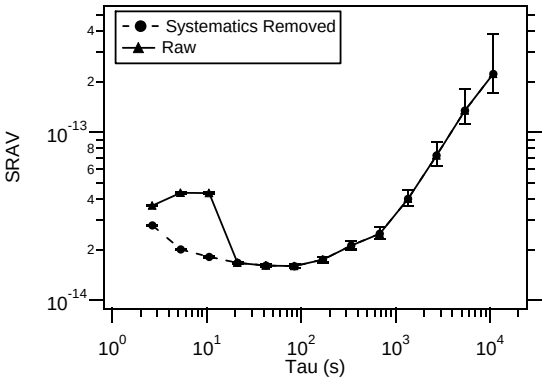


Figure 6. Square Root Allen Variance (SRAV) frequency stability measurement. The hump at short integration times is due to systematic effects associated with the rotation of the experiment. These have been removed and the SRAV replotted for comparison.

4 ANALYSIS

Inside the sapphire crystals standing waves are set up around the circumference of the two cylinders with the dominant electric and magnetic fields in the axial and radial directions respectively, corresponding to a propagation (Poynting) vector around the circumference. The observable of the experiment is the difference

frequency, and to test for Lorentz violations the perturbation of the difference frequency with respect to an alternative test theory has been derived. In the photon sector of the SME the frequency perturbation may be calculated, to first order, as the integral over the non-perturbed fields (Eq. (34) of [9]), and subsequently expressed in terms of 19 independent variables. The change in orientation of the fields due to the laboratory rotation and Earth's orbital and sidereal motion induces a time varying modulation of the difference frequency, which is searched for in the experiment. Alternatively, with respect to the RMS framework, the change in the Poynting vector direction with respect to the cosmic microwave background is analyzed as a second order effect. A complete verification of LLI in the RMS framework [7,8] requires a test of (i) the isotropy of the speed of light ($P_{MM} = \delta - \beta + 1/2$), a Michelson-Morley (MM) experiment [16], (ii) the boost dependence of the speed of light ($P_{KT} = \beta - \alpha - 1$), a Kennedy-Thorndike (KT) experiment [17], and (iii) Ives-Stillwell (IS) experiments [18,19], which measure the time dilation parameter ($P_{IS} = \alpha + 1/2$). Because our experiment compares two cavities it is only sensitive to the isotropy (MM) parameter and cannot measure the boost (KT) parameter, contrary to the previous experiment that compared a sapphire oscillator to a hydrogen maser [1,3].

The amplitude and phase of any Lorentz violating signal is determined by simultaneously fitting the parameters of equation (1) to the data, with the phase of the fit adjusted according to the test model used. A and B determine the frequency offset and drift, and C_i and S_i are the amplitude of a cosine and sine at frequency ω_i respectively.

$$\Delta\nu/\nu = A + Bt + \sum_i C_i \cos(\omega_i t) + S_i \sin(\omega_i t) \qquad (1)$$

The data used in this analysis was acquired over a 3 day period during December 2004. In the final analysis we fit 5 frequencies to the data, i={$2\omega_R$, $2\omega_R \pm \omega_\oplus$, $2\omega_R \pm 2\omega_\oplus$}, as well as the frequency offset and drift using the ordinary least squares (OLS) technique. ω_R is the rotation frequency and $\omega_\oplus$ is the sidereal frequency. Upon inspection the residuals resulting from an OLS fit exhibit a significantly non-white behavior. In the presence of non-white noise OLS is not the optimal regression method and it is more appropriate to use weighted least squares (WLS) [1]. WLS involves pre-multiplying both the experimental data and the model matrix by a whitening matrix determined by the noise type of the OLS residuals. Fig. 7 shows the total amplitude resulting from a WLS fit to the data over a range of frequencies about $2\omega_R$. It is evident that the systematic at $2\omega_R$ is affecting the fitted amplitude of the sidereal sidebands $2\omega_R \pm \omega_\oplus$.

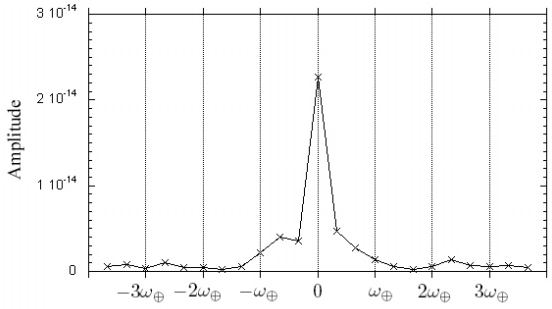


Figure 7. Spectrum of amplitudes calculated using weighted least squares, showing the level of systematicleakage about $2\omega_R$. $\omega_\oplus$ is the sidereal frequency.

In the photon sector of the SME [9], Lorentz violating terms are parameterized by 19 independent parameters, which are in general grouped into three traceless and symmetric 3 × 3 matrices ($\tilde{\kappa}_{e+}$, $\tilde{\kappa}_{o-}$, and $\tilde{\kappa}_{e-}$), one anti-symmetric matrix ($\tilde{\kappa}_{o+}$) and one additional scalar, which all vanish when LLI is satisfied. The 10 independent components of $\tilde{\kappa}_{e+}$ and $\tilde{\kappa}_{o-}$ are constrained by astronomical measurements to $< 2\times10^{-32}$ [9,20]. Seven components of $\tilde{\kappa}_{e-}$ and $\tilde{\kappa}_{o+}$ have been constrained in an optical and microwave cavity experiment [4,3] at the 10^{-15} and 10^{-11} level respectively, while the scalar $\tilde{\kappa}_{tr}$ component recently had an upper limit set of $< 10^{-4}$ [21]. The remaining $\tilde{\kappa}_{e-}^{ZZ}$ parameter could not be previously constrained in non-rotating experiments [3,4]. In this work we have measured a value of 20.1 (1.3) × 10^{-15} for $\tilde{\kappa}_{e-}^{ZZ}$. Since the result is dominated by systematic effects at $2\omega_R$, the constraint is taken as being the amplitude, 2.0 × 10^{-14}. Our experiment is also sensitive to all other seven components of $\tilde{\kappa}_{e-}$ and $\tilde{\kappa}_{o+}$ and improves present limits by up to a factor of 2, as shown in Tab. I. To calculate the expected signal we transform to the SCECF using the standard transformations as in [9], and add rotation in the laboratory frame in a similar way to [21]. A complete analysis of the $\tilde{\kappa}_{e-}$ and $\tilde{\kappa}_{o+}$ terms requires at least one year of data to decorrelate their contributions to a Lorentz violating signal. In the same way as [22] we have assumed that there is no correlation, and hence no cancellation, between the $\tilde{\kappa}_{e-}$ and $\tilde{\kappa}_{o+}$ terms, allowing us to place limits on them.

In the RMS frame-work, a frequency shift due to a putative Lorentz violation is given by equation (2) of [1], where v is the velocity of the preferred frame wrt the CMB, $\hat{\theta}_j$ is the unit vector in the direction of the azimuthal angle (direction of propagation) of each resonator (labeled by subscripts 1 and 2), and φ is the azimuthal variable of integration in the cylindrical coordinates of each resonator.

$$\frac{\Delta\nu_0}{\nu} = \frac{P_{MM}}{2\pi c^2}\oint\left(\left(\mathbf{v}.\hat{\theta}_1\right)^2 - \left(\mathbf{v}.\hat{\theta}_2\right)^2\right)d\varphi \qquad (2)$$

Perturbations due to Lorentz violations occur at the same five frequencies as the SME, but in this work we do not consider the $2\omega_R$ frequency due to the large systematic, as we only need to put a limit on one parameter. We determine a value of 2.3 (5.7) × 10^{-10} for P_{MM}, a factor of 3 better than [4].

	$\tilde{\kappa}_{e-}^{XY}$	$\tilde{\kappa}_{e-}^{XZ}$	$\tilde{\kappa}_{e-}^{YZ}$
from [3]	-5.7 (2.3)	-3.2 (1.3)	-0.5 (1.3)
this work	-0.3 (0.9)	-1.2 (1.0)	-0.9 (1.0)
	$\tilde{\kappa}_{e-}^{XX} - \tilde{\kappa}_{e-}^{YY}$	$\tilde{\kappa}_{e-}^{ZZ}$	
from [3]	-3.2 (4.6)	-	
this work	-1.43 (2.27)	20.1 (1.25)	
	$\tilde{\kappa}_{o+}^{XY}$	$\tilde{\kappa}_{o+}^{XZ}$	$\tilde{\kappa}_{o+}^{YZ}$
from [3]	-1.8 (1.5)	-1.4 (2.3)	2.7 (2.2)
this work	-2.3 (1.2)	0.7 (0.89)	-0.6 (1.4)

TABLE I. Results for the SME Lorentz violation parameters, using Lipa's method of assuming no cancellation between the $\tilde{\kappa}_{e-}$ (in 10^{-15}) and $\tilde{\kappa}_{o+}$ (in 10^{-11}) [22].

5 Conclusion

A new Michelson-Morley experiment based around two cryogenic sapphire oscillators has been described in this work, and the results of the first operation have been discussed. An analysis in both the RMS and SME frameworks has shown an improvement of a factor of 3 in the former over previous results [4], and a factor of 2 in the latter as compared to [3]. The performance of the experiment will be enhanced further by implementing tilt and environmental control, and to remove the assumption that the $\tilde{\kappa}_{e-}$ and $\tilde{\kappa}_{o+}$ do not cancel each other data integration will continue for more than a year.

6 Acknowledgment

This work was supported by the Australian Research Council. Thanks to the technical staff of the UWA School of Physics for their expertise.

* *Corresponding Author: Paul Stanwix*

Email: pstanwix@physics.uwa.edu.au

References

[1] P. Wolf, S. Bize, A. Clairon, et al., "Tests of Lorentz invariance using a microwave mesonator," Phys. Rev. Lett., vol. 90, 6, 060402, (2003).

[2] P. Wolf, M.E. Tobar, S. Bize, et al., "Whispering gallery resonators and tests of Lorentz invariance," Gen. Rel. Grav., vol. 36, 10, pp. 2351-2372, (2004).

[3] P. Wolf, M.E. Tobar, S. Bize, et al., "Improved test of Lorentz invariance in electrodynamic," Phys. Rev. D Rapid Comm., vol. 70, 051902, (2004).

[4] H. Muller, S. Herrmann, C. Braxmaier, et al., "Modern Michelson-Morley experiment using cryogenic optical resonators," Phys. Rev. Lett., vol. 91, 2, 020401, (2003).

[5] A. Brillet and J.L. Hall, "Improved laser test of the isotropy of space," Phys. Rev. Lett., vol. 42, 9, pp. 549-552, (1979).

[6] C.M. Will, *Theory and Experiment in Gravitational Physics*, *Revised Edition*, Cambridge U. Press, (1993).

[7] H.P. Robertson, "Postulate versus observation in the special theory of relativity," Rev. Mod. Phys., vol. 21, pp. 378-382, (1949).

[8] R. Mansouri and R.U. Sexl, Gen. Rel. Grav., vol. 8, 497, 515, 809, (1977).

[9] V.A. Kostelecky and M. Mewes, "Signals for Lorentz violation in electrodynamics," Phys. Rev. D, vol. 66, 056005, (2002).

[10] M.E. Tobar, J.G. Hartnett, J. Anstie, "Proposal for a new Michelson-Morley experiment using whispering spherical mode resonators", Phys. Lett. A, vol. 300, 1, pp. 33-39, (2002).

[11] C.R. Locke, S. Munro, M.E. Tobar, E.N. Ivanov, G. Santarelli, "Performance of the next generation cryogenic sapphire oscillator," Proc. Joint Meeting of IEEEIFCS/EFTF, pp. 350-354, (2003).

[12] S.K. Jones, "Frequency stabalisation with a sapphire loaded superconducting cavity," PhD Thesis, Physics Dept., The University of Western Australia, (1988).

[13] A.J. Giles, S.K. Jones, D.G. Blair, M.J. Buckingham, "A high stability microwave oscillator based on a sapphire loaded superconducting cavity," Proc. 43rd Symp. Freq. Control, pp. 89-93, (1989).

[14] A.J. Mann, A.J. Giles, D.G. Blair, M.J. Buckingham, "Ultra-stable cryogenic dielectric microwave resonators: mode-frequency temperature compensation by residual paramagnetic impurities," J. Phys. D, vol. 25, pp. 1105-1109, (1992).

[15] S. Chang, A.G. Mann, A.N. Luiten, D.G. Blair, "Measurements of radiation pressure effect in cryogenic sapphire dielectric resonators," Phys. Rev. Lett., vol. 79, 11 pp. 2141-2144, (1997).

[16] A.A. Michelson and E.W. Morley, "On the relative motion of the earth and the luminiferous ether," Am. J. Sci., vol. 34, 203, pp. 333-345, (1887).

[17] R.J. Kennedy and E.M. Thorndike, "Experimental establishment of the relativity of time," Phys. Rev., vol. B42, pp. 400-418, (1932).

[18] H.E. Ives and G.R. Stilwell, "An experimental study of the rate of a moving atomic clock," J. Opt. Soc. Am., vol. 28, pp. 215-226, (1938).

[19] G. Saathoff, *et al.*, "Improved test of time dilation in special relativity," Phys. Rev. Lett., vol. 91, 190403, (2003).

[20] V.A. Kostelecky and M. Mewes, "Cosmological constraints on Lorentz viloations in electrodynamics," Phys. Rev. Lett., vol. 87, 251304, (2001).

[21] M.E. Tobar, *et al.,* "New methods of testing Lorentz violations in electrodynamics," Phys. Rev. D, vol. 71, 025004, (2005).

[22] J.A. Lipa, *et al.*, "New limit on signals of Lorentz violation in electrodynamics, " Phys. Rev. Lett., vol. 90, 060403, (2003).

Session 5B / GNSS time scales

GALILEO SYSTEM TEST BED V1: RESULTS ON THE EXPERIMENTAL GALILEO SYSTEM TIME

F. Cordara*, R. Costa*, L. Lorini*, D. Orgiazzi*, V. Pettiti*, I. Sesia*°, P. Tavella*,

P. Elia, M. Mascarello**,**

M. Falcone†, J. Hahn†

* Istituto Elettrotecnico Nazionale "Galileo Ferraris", Torino, Italy

° Politecnico di Torino, Torino, Italy

** Alenia Spazio S.p.A., Roma, Torino, Italy

† European Space Agency (ESA/ESTEC), Noordwijk, The Netherlands

ABSTRACT

The Galileo System Test Bed (GSTB) V1 is the first experimental phase in the Galileo project supported by the European Space Agency (ESA). The aim is testing the Galileo algorithms in a mixed configuration where the space segment is given by the existing GPS constellation, while the ground segment is an experimental set-up as close as possible to the Galileo architecture. In the framework of the GSTB V1, Alenia Spazio and the Istituto Elettrotecnico Nazionale (IEN), had the responsibility to design and realize an infrastructure able to generate the Experimental Galileo System Time (E-GST) which is the reference time for all the GSTB V1. The infrastructure that allows the realization of E-GST is the Experimental Precise Timing Station (E-PTS), which is implemented at the IEN facilities, Turin, Italy.

The paper reports the main results of the E-PTS experimentation carried out during the whole 2004 addressing the E-GST overall performance with respect to the demanding GSTB V1 goals. Besides, the lessons learnt from the experiment are also pointed out in the paper, together with some recommendations to be taken into account in the design of the Precise Timing Facilities (PTFs) of the Galileo ground segment[1].

1. TIMEKEEPING INSIDE GSTB V1

During the GalileoSat Definition Phase an early experimental stage, named Galileo System Test Bed (GSTB), was defined as an integral part of the Galileo Design Development and Validation Phase in order to mitigate program risks [1,2,3].

The main goal for the GSTB first stage (GSTB V1) is to conduct a set of experiments to verify, in a real context, the Galileo concepts defined in the Galileo B2 phase [4]. These major objectives have been pursued relying on the existing GPS constellation and realising a ground infrastructure close to the final Galileo set-up where the critical algorithms (system time generation, orbit determination, time synchronisation and integrity) could be implemented and tested continuously for a significant period (12 months).

In this scenario, an Experimental Precise Timing Station (E-PTS) has been designed and implemented by Alenia Spazio, Italy and Istituto Elettrotecnico Nazionale (IEN), Turin, Italy, to test the Galileo system time algorithms in a real environmental condition.

2. THE EXPERIMENTAL PRECISE TIMING STATION (E-PTS)

The E-PTS infrastructure (Figure 1) that provides the E-GST is installed and continuously operated since 2003 at the IEN premises in Turin, Italy.

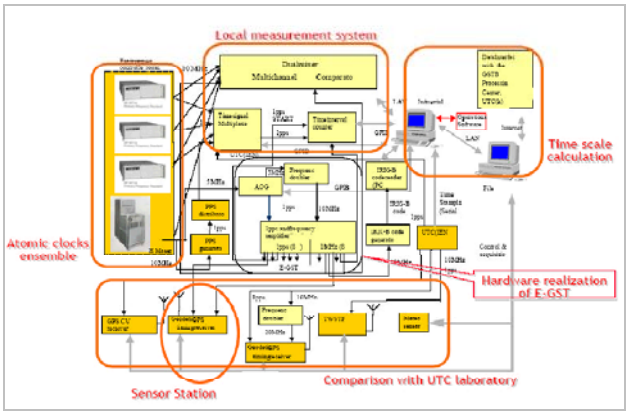


Figure 1: The E-PTS architecture and its major functions.

[1] This paper is mainly based but not exclusively on the paper [20] presented as a poster to the 36th Precise Time and Time Interval (PTTI) meeting, Washington DC, December 2004.

The E-PTS architecture partially re-uses the IEN atomic clocks ensemble, that is a hydrogen maser (H-maser) and 3 commercial Cesium (Cs) clocks [5,6]. Moreover, the remote time transfer equipment (TWSTFT station and GPS Common View receivers) for comparison with external UTC laboratories, namely PTB (Germany) and NPL (UK), are also shared with IEN institutional duties. Because both systems are driven by the local UTC(IEN) time scale, their measures are then related to E-GST by means of supplementary internal measurements.

All the other new hardware equipment especially procured for E-PTS are located in dedicated rack hosted in the IEN laboratory (Figure 2). Here two independent local measurement systems perform the measurements of the time difference between couples of clocks readings, as required by the time scale generation algorithm.

The primary measurement system is based on a 1PPS signal multiplexer followed by a high-resolution time interval counter (TIC), providing hourly measures to the E-GST generation algorithm. In addition, a multi-channel phase comparator (MCPC) acts as an auxiliary system that performs simultaneous high rate (down to 1 Hz) measurements of all the E-PTS clocks, aiming an effective real-time monitoring and analysis of the behaviour of each oscillator.



Figure 2: The E-PTS installation at the Italian timing laboratory (IEN, Turin).

The timely processing of the data provided by both measurement systems is allocated to the E-PTS operational software, which is also in charge of the following main functions:

- automated schedule and overall management of the E-PTS operation;
- *real-time* generation, steering and physical realisation of the E-GST time scale;
- algorithm testing in different scenarios (*off-line* experimentation);
- experimental results (core products) generation, formatting, archiving and transfer to the GSTB Processing Center (GPC) at ESA/ESTEC premises, Nordwijk, The Netherlands.

The frequency output of the H-maser is fed to a high accuracy phase micro-stepper (namely an Auxiliary Output Generator, AOG) which do not degrade the intrinsic short-term stability of the H-maser. A frequency correction is then imposed to the AOG according to the E-GST algorithm: the output of the AOG (5 MHz and 1PPS signals) is the physical realisation of E-GST, a real-time hardware time scale.

A dual-frequency GPS receiver (a Javad Legacy receiver) is also operated by the E-PTS as an element of the GSTB V1 sensor station network. This receiver is directly connected to the physical realisation of E-GST (10 MHz and 1PPS signals), allowing the orbitography algorithms to be referenced to E-GST within the Orbit Determination and Time Synchronisation (OD&TS) experimentation activities.

2.1 Early E-PTS Versions

During the deployment phase of the overall GSTB V1 core infrastructure, a reduced functionality E-PTS has been established to support the early operations of the GSTB V1 (integration and verification activities).

In particular, since April 2003, a first version called E-PTS V1.1 has been operated, providing raw data from an IEN-property dual frequency geodetic GPS receiver (an Ashtech Z-12T) directly referenced by the local UTC(IEN) time scale.

Once the installation and verification of the E-PTS sensor station has been completed (July 2003), an additional provisional version has been delivered to provide 1 Hz raw data referred to a free-running active H-maser of the IEN Time and Frequency laboratory. This version, called E-PTS V1.2, was used as a very stable reference during the OD&TS preliminary experimentation.

3. THE EXPERIMENTAL GALILEO SYSTEM TIME (E-GST)

As a prototype of the final Galileo System Time (GST), the E-GST needs to have optimal characteristics for both navigation and time dissemination purposes. E-GST is realised in real time and with a hardware representation.

To fulfil the time dissemination purposes of Galileo, its reference time has also to be strictly synchronised (modulo 1 second) to the international reference time, called Universal Time Coordinate (UTC). To this aim, E-GST is then steered versus the International Atomic Time, TAI (which differs from UTC by an entire number of seconds), by the use of continuous comparison measurements versus several local national time scales UTC(k). It's worth mentioning that in the final Galileo architecture the steering function to TAI will be

accomplished by the use of an external Time Service Provider (TSP) [7].

The real-time E-GST (referred as the on-line version) is realized following a basic scenario involving only the E-PTS clocks and a baseline algorithm based on the Galileo Phase B2 algorithm [8], with additional features suitably devised for the GSTB V1 purposes. The real-time E-GST generation algorithm is given by the three steps process following also the recommendations in [9]:

- at first, an ensemble time scale, called GSTR', is computed every day as a weighted average of the E-PTS Cesium clocks;
- the ensemble time scale GSTR' is in turn compared to TAI every month, as soon as the BIPM Circular T data are available, and then corrected in frequency with the aim of being in agreement with TAI in the long term. The resulting steered time scale, called GSTR, is then kept in agreement with TAI;
- the frequency drift of the E-PTS master clock (the IEN active H-maser) is daily estimated versus the steered time scale GSTR. The frequency correction is then applied to the H-maser through the AOG, providing the real-time physical realisation of E-GST.

It follows that the resulting E-GST should provide:

- the good short-term stability of the E-PTS H-maser to support the navigation tasks of Galileo (mainly orbit determination and time synchronisation;
- the medium-term stability of the Cesium clocks ensemble, allowing the estimation of the E-PTS clocks parameters (in particular, the frequency drift of the H-maser);
- the long-term stability (and the accuracy) of TAI and UTC to fulfil the time dissemination purposes.

4. E-PTS PERFORMANCE ASSESSMENT

Table 1 reports the E-PTS performance targets as required [4] in the frame of GSTB V1, together with the goals expected for the final Galileo.

	GSTB V1	Galileo
Offset of E-GST versus TAI	< 1 μs	< 50 ns (95%)
Stability @ 1 day of E-GST versus TAI	$< 5.5 \cdot 10^{-14}$ (2σ)	$< 5.5 \cdot 10^{-14}$ (2σ)
Uncertainty of the prediction of E-GST versus TAI	33 ns (2σ)	28 ns (2σ)

Table 1: E-PTS performance targets (from [4]).

Starting from March 5th, 2004, the E-PTS is in the final configuration. The hardware E-GST time scale is continuously generated in real-time, according with the computation performed by the algorithm, and it is then fed to the sensor station obtaining the GPS pseudo-ranges referenced to E-GST.

The E-PTS performances in term of E-GST figures has been assessed comparing the resulting E-GST physical time scale versus the TAI reference (available a posteriori with about 1.5 month delay, 5 days spaced) and using UTC(IEN) as an intermediate time scale. The needed time offset between E-GST and UTC(IEN) has been provided by the E-PTS measurement system with a sub-nanosecond uncertainty.

4.1 E-GST Accuracy versus TAI

The time offset between E-GST and TAI is given in Figure 3 starting from March 5th, 2004, and up to the end of November 2004 covering a period of 9 months. The behaviour of E-GST versus PTB and NPL time scales, through both TWSTFT and GPS CV, is also reported in Figure 4.

It's worth to mention that the initial value (about –30 ns) is the time offset of the E-PTS free-running H-maser versus TAI at time when the AOG was enabled. Such an offset has been maintained to avoid phase jumps on the E-GST time scale over the software-to-hardware transition, accepting then a non-zero but known initial bias between the two time scales.

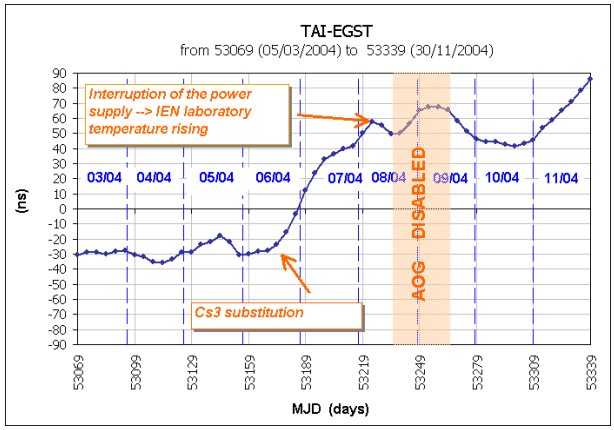


Figure 3: Time offset between E-GST and TAI. Maximum: 85.8 ns; minimum: -35.7 ns; mean: 17.5 ns. Major events occurred in the E-PTS are also marked.

Figure 3 shows that the initial offset of E-GST with respect to TAI was quite well maintained in the first three/four months of the E-PTS operation. However, because several events occurred in the E-PTS during the experiment (such as operational software troubles, Cs clock substitution and uncontrolled rising of the laboratory temperature), the E-GST versus TAI accuracy slightly degraded in the following months.

Despite these problems and the limited set of clocks in the E-PTS ensemble, a good accuracy performance has been achieved, quite promising for designing the final Galileo algorithm and its Precise Timing Facility (PTF). As reported in Figure 5, the absolute value of the E-GST to TAI offset was below 71 ns for 95% of the time (< 50 ns for 70% of time). Even if the GSTB V1 goal has been met, this performance would not fully satisfy the demanding specifications for final Galileo (50 ns, as reported in Table 1).

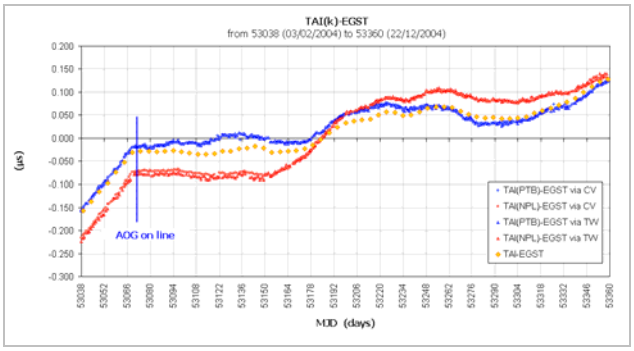


Figure 4: E-GST comparison versus TAI (yellow diamonds), PTB (blue) and NPL (red) time scales, both by GPS CV (circles) and TWSTFT (triangles), up to December 21st, 2004. Steering of E-GST started on March 5th, 2004 (MJD 53069).

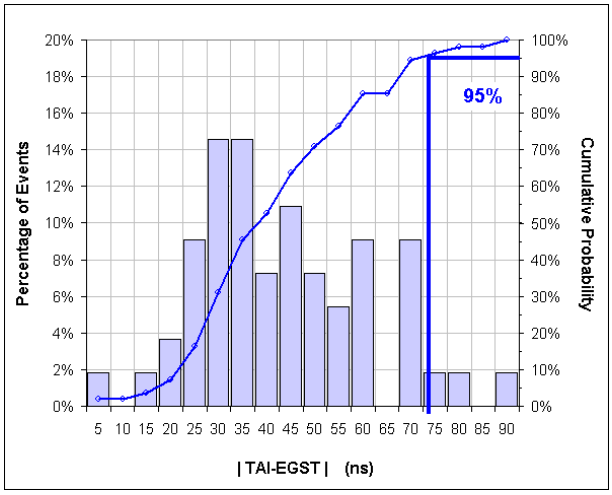


Figure 5: Time distribution of the offset between E-GST and TAI (histogram) and associated cumulative probability.

4.2 E-GST Stability (Short and Long Term)

The long-term stability of E-GST versus TAI has been estimated through the computation of the Allan deviation (ADEV) [10] of the time difference between E-GST and TAI. Following Table 2 reports the ADEV values at different observation times t, up to 50 days, over the 9 months period of interest.

τ, days	ADEV (1σ)
1	$1.45 \cdot 10^{-14}$ (✱)
5	$6.5 \cdot 10^{-15}$
10	$8.0 \cdot 10^{-15}$
20	$8.6 \cdot 10^{-15}$
50	$7.8 \cdot 10^{-15}$

Table 2: E-GST stability versus TAI in term of Allan deviation (ADEV). ✱ = extrapolated value.

Since the BIPM data are spaced by 5 days, the stability at τ = 1 day (as requested by GSTB V1 specification [4]) has been extrapolated starting from the value obtained for 5 days. Assuming that E-GST is mostly affected by a white frequency noise, the extrapolation expression is:

$$\sigma_y(1day) = \sigma_y(5days) \cdot \sqrt{\frac{(5days)}{(1day)}} = \sigma_y(5days) \cdot \sqrt{5}$$

Even if this expression seems to be pejorative (especially further to the findings below concerning the E-GST short-term stability analysis), the resulting stability of 2.9×10^{-14} (2σ) is fully compliant with the dissemination requirements reported in Table 1.

Moreover, for observation intervals up to one day, the E-GST stability is expected to be the H-maser stability apart from possible corruption due to the AOG and the signal distribution chain as well as the daily steering procedure.

To evaluate such an impact, the time offset between E-GST and the free-running H-maser has been measured by the MCPC (1 Hz acquisition rate). Since E-GST is just a frequency steered replica of the same H-maser, the difference provided by MCPC drops out the contribution of the H-maser itself. Only the augmentation of the AOG plus distribution chain plus steering procedure is then taken into account, giving a figure of the noise added by these equipment.

Stability results in term of ADEV are given in Figure 6 below for observation interval from 1 second up to about 1 day. The nominal stability of the E-PTS H-maser and the intrinsic limit of the measurement system are also reported for comparison purposes.

The measurement system acts as a lower bound for the E-GST behavior detection up to some tens of seconds. Whereas, the E-GST generation and distribution chain doesn't degrade the H-maser frequency stability, except for a "bump" around an observation time of 2000 seconds may be due to a thermal effect. Further to these notices, the E-GST short-term stability could be estimated as the thick red line plotted in Figure 6. This stability has been clipped with a $3.0 \cdot 10^{-15}$ floor according to the manufacturer specification of the H-maser [11].

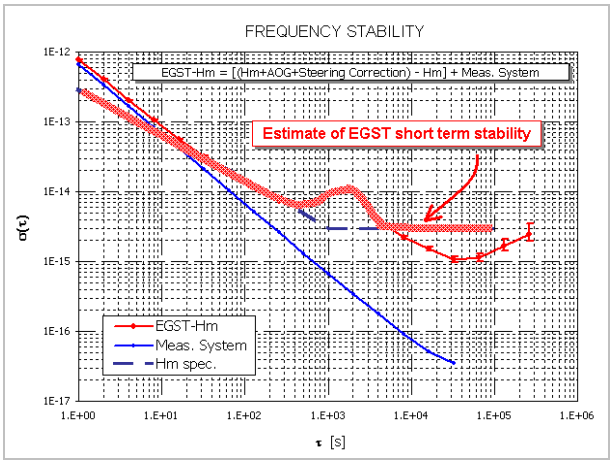


Figure 6: E-GST short-term frequency stability versus the free-running H-maser (dotted red line), together with the stability of the E-PTS H-maser as provided by manufacturer specifications (dashed blue line) and the intrinsic limit of the measurement system (solid blue line).

To summarize, the short-term stability estimated above along with the long-term stability measured versus TAI (Table 2) give then a complete characterization of the actual E-GST stability (from 1 second to 50 days). The E-GST stability over one day required by the Galileo specifications ($5.5 \cdot 10^{-14}$) and estimated by an extrapolation ($2.9 \cdot 10^{-14}$) was possibly an over-estimation, because value from real experimentation seems even better ($< 3.0 \cdot 10^{-15}$).

Besides, the E-GST short-term stability as measured inside the E-PTS is still not the last evaluation. Actually, E-GST is visible from outside the E-PTS through the GPS measurement (and the subsequent processing) provided by the dual frequency GPS receiver acting as the E-PTS sensor station in the frame of the OD&TS experimentation activities.

Aiming to get an estimate of E-GST as seen through external measurement systems, the E-PTS sensor station data have been processed using a Precise Point Positioning (PPP) technique [12,13], referring E-GST to the International GPS Service Time (IGST) scale [14]. In addition, the time offset between E-GST and the UTC(USNO) as estimated by the OD&TS process have been taken into account [15]. Stability results in term of ADEV are given in following Figure 7, starting from a 300 seconds (5 minutes) observation interval.

As expected, the plot shows that the E-GST stability is degraded by the external measurement systems (receiver equipment and data processing algorithms). In addition, also the intrinsic noise of the external time scales used as reference (namely IGST and UTC(USNO)) affects the overall E-GST stability.

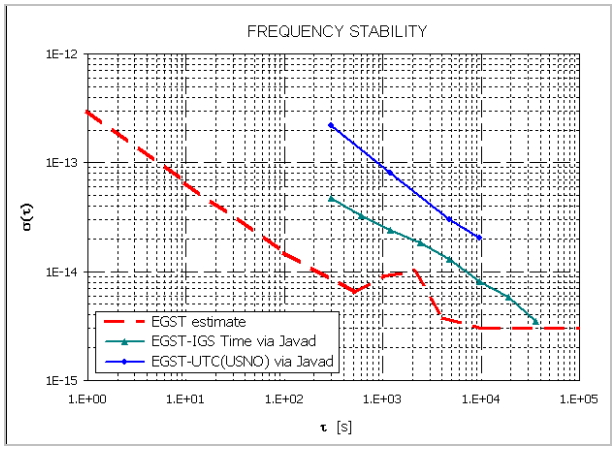


Figure 7: E-GST frequency stability as estimated versus the IGST time scale by a Precise Point Positioning, PPP, technique (green line) and versus UTC(USNO) by OD&TS processing (blue line). The E-GST stability as measured inside the E-PTS is also reported (dashed red line) as a reference.

4.3 Uncertainty of the Prediction of E.GST versus TAI

The E-GST to TAI time offset is known from BIPM data only 1.5 months a posteriori. When BIPM data are available, a prediction of the E-GST to TAI offset is evaluated for the subsequent 1.5 months (45 days). When fresh data from BIPM are available, the difference between predicted versus real measures is evaluated. This difference is the prediction error. As reported in Table 1, the GSTB V1 requirements ask to have a prediction error within 33 ns (95%).

For the 9 months of E-PTS operation, the prediction error as a function of the prediction length has been evaluated and the results are reported in Figure 8 below.

As expected, the prediction for a short period is more accurate than the prediction on a long period. In particular, it can be seen that at 45 days the prediction error covers a region of about ±40 ns, which is outside the specification. Besides, the worst prediction errors are for the months of July and August, where several troubles in the E-PTS operation has been experienced.

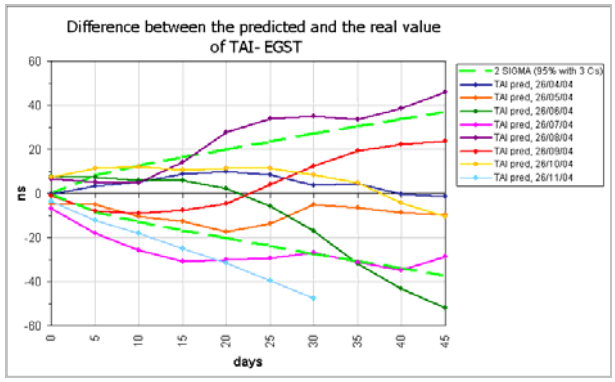


Figure 8: Prediction error on the E-GST versus TAI time offset, evaluated for the 9 months of E-PTS operation.

The 95% boundary provided by a theoretical model [16] is also reported.

In addition to these experimental results, the prediction error has been also estimated by means of a theoretical model detailed in [16]. As mentioned before, in GSTB V1 the E-GST steering versus TAI is obtained through different intermediate steps that firstly steer the Cs clocks ensemble on TAI, then steer E-GST versus the steered Cs ensemble. The evaluation of the uncertainty of the E-GST to TAI offset becomes therefore quite complicated and an approximated evaluation was performed basing on certain assumptions:

- the TAI prediction is obtained as a linear extrapolation of the past month of BIPM data;
- the medium long term stability of E-GST is completely driven by the Cesium ensemble stability (i.e., the H-maser steering algorithm is working very well following the Cesium behaviour for observation time larger than one day);
- the Cs clocks are mostly affected by a white frequency noise (WFN) of about $3 \cdot 10^{-14}$ at one day and they have a frequency offset that is estimated versus TAI on the previous month of BIPM data;
- the steering versus TAI does not corrupt the Cs ensemble stability.

Under these hypothesis, [16] shows that the uncertainty prediction for an interval t depends both on the number of Cs clocks and on the uncertainty of the measures linking E-GST to TAI (through TWSTFT or GPS CV). In particular, the uncertainty of the TWSTFT link (used to connect UTC(IEN) to UTC) was estimated to be 1.4 ns (1σ) [17,18]; therefore, the prediction uncertainty for a period of 45 days results to be about 37 ns (95%, N = 3). In case a GPS CV technique was used, the prediction uncertainty would become about 41 ns (95%, N = 3), because the higher uncertainty (5.1 ns, 1σ).

Following Figure 9 plots the prediction uncertainty for a different number of Cs clocks as estimated applying the theoretical model above and assuming to use a TWSTFT link. It can be seen that with 3 Cs clocks the uncertainty is about 37 ns (95%), while a value of about 28 ns (95%), as required in final Galileo, would be ensured by operating 6 Cs clocks, at least.

Besides, comparing the 95% boundary obtained by this approximate theoretical model with the experimental data (as reported in Figure 8), it follows that a good agreement is then achieved.

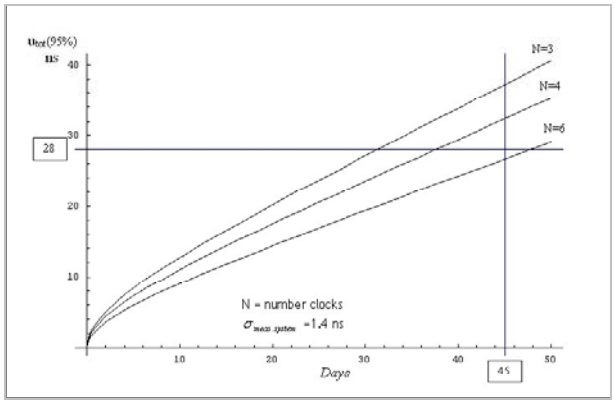


Figure 9: Uncertainty of the E-GST to TAI offset prediction for a different number of Cs clocks, as estimated using the theoretical model described in [16].

5. CONCLUSIONS

According with the experimentation conducted up to now, the results can be summarised in the following Table 3.

It follows that the GSTB V1 targets have been reached apart the uncertainty on the prediction of E-GST with respect to TAI. Moreover, the comparison respect the Galileo targets gives the confidence that with a dedicated laboratory with, i.e., an higher number of Cs clocks and redundant systems, the targets could be reached even if they are demanding.

	GSTB V1 Target	**GSTB V1 Results**	**Galileo Target**
Offset of E-GST versus TAI	< 1 μs	< 71 ns (95%)	< 50 ns (95%)
Stability @ 1 day of E-GST versus TAI	$< 5.5 \cdot 10^{-14}$ (2σ)	$< 2.9 \cdot 10^{-14}$ (2σ)	$< 5.5 \cdot 10^{-14}$ (2σ)
Uncertainty of the prediction of E-GST versus TAI	33 ns (2σ)	40 ns (2σ)	28 ns (2σ)

Table 3: E-PTS performance results with respect to targets

Many lessons have been learned by the E-PTS team from this long time, day-by-day operation and experimentation. To summarise, following valuable recommendations on hardware, software and algorithm can be then issued in view of the future work for the Galileo ground segment about the PTFs design and implementations:

- apart from redundancy needs, additional Cs clocks in each PTF ensemble is highly advised, mainly aiming to maintain the demanding agreement of Galileo system time versus TAI (50 ns, 95%). As a consequence, the future PTF architecture (in term of both hardware and software) shall be designed in order to be able to process an higher number of clocks

(12, for instance) and allow a safe management of the introduction and removal of a clock (e.g. for purpose of maintenance);

- one or even two H-masers are not able to allow the monitoring of its behaviour (i.e., performance analysis and anomalies detection). The procurement of three H-masers for each PTFs should be then taken into account during the architecture design;
- since most of H-masers allow a limited range of operating temperature only (e.g., 4°C around 22°C for the equipment used in GSTB V1), the clocks room inside PTF should be designed to avoid large temperature oscillations even in case of malfunction of the thermal control system. This will prevent that severe failure could affect the H-masers and the Cs clocks as well. Moreover, all the hardware equipment involved in the clocks measurement system and in the GST physical realization (namely the AOG and the distribution system) should be placed in the same temperature-controlled room;
- even if the E-GST baseline algorithm (as defined in Phase B2) meets most of the requirements and it is also accepted as the Galileo system time algorithm, some improvements should be addressed by next re-engineering stage. In particular, it is suggested to insert an additional check on clock anomaly (with respect to the expected behaviour) at the very beginning of the data pre-processing, to avoid inserting “bad” measures in the GST computation;
- the check on the frequency steering correction to be imposed on the H-maser through the AOG was too weak and did not protect versus large anomalies (as experienced in the last months of the experimentation). It follows that a more robust check should be addressed, maybe taking into account a configurable threshold with respect to the last imposed correction on the H-maser;
- it is possible that for some intentional or accidental reasons, a clock may suffer a time or frequency jumps. However, such jumps have to be compensated before using those clocks in the GST algorithm, not to inject false instabilities. It follows that the time and frequency jumps compensation features should be included and fully managed inside the GST processing algorithm;
- because very strong correlation between a single clock and the average time scale based on a small ensemble of clocks, a single clock may influence very much the average time scale and degrade its stability in case of clock malfunctioning. To reduce the risk of the so called “clock ensemble correlation” [19], a corrective factor should be then applied;
- the current baseline algorithm is designed to maintain E-GST in agreement with TAI but no recovery action are envisaged to cope with possible degraded situations. In case of a GST is too far away from TAI, a manual recovery action should be available, providing a configurable time or frequency step to be triggered by the PTF operator.

The main challenge of the experiment was the completely automatic generation of E-GST asking for robustness on clock measurement, clock monitoring, ensemble algorithm evaluation, loading BIPM data, steering versus TAI, evaluation of the H-maser frequency correction, real-time hardware E-GST generation, connection to a GPS geodetic receiver and automatic data exchange to Control Center, time transfer with PTB and NPL, real time monitoring of operation.

The main experienced difficulties may be summarised as follows:

- hardware failures (PPS generator, signal distributor, temperature control…);
- algorithm robustness (anomaly protection, tuning…);
- software failures (bugs, network traffic, PC virus…);

Moreover, a good early interface between T&F scientists and industrial software team (analysis, design, and development) would be advisable.

In addition, it must be noticed that some PTF key activities were not planned in the GSTB V1 experimentation, mainly devoted to test critical algorithms in a real but experimental set-up. Nevertheless, hardware matters (such as, H-masers automatic switching system, architecture redundancy and its impact on the performance, 2nd PTF switching) should be fully investigated in the early steps of Galileo program, in order to have a reliable design.

At last, it’s worth mentioning that in the final Galileo architecture the steering function to TAI will be accomplished by the use of an external Time Service Provider (TSP) [7].

ACKNOWLEDGEMENTS

GSTB V1 is an ESA program and the prime contractor is Galileo Industries (GaIn); in turn, GaIn is a joint company between Alenia Spazio, Alcatel Industries, Astrium GmbH, Astrium Ltd. and GSS (Galileo Sistemas y Servicios).

The E-PTS, one of the GSTB V1 elements, is under the Alenia Spazio responsibility. The companies and institutions involved in the E-PTS development are Alenia Spazio and the Istituto Elettrotecnico Nazionale "Galileo Ferraris" (IEN - Italy), the Space Software Italia (SSI) developed the software and the National Physical Laboratory (NPL, U.K.), the Physikalisch Technische Bundesanstalt (PTB - Germany) provided external contributions.

REFERENCES

[1] M. Falcone, R. Lucas Rodriguez, F. Emma, and J. Hahn, "The GALILEO System Test Bed (GSTB)," in Proceedings of the GNSS 2001 Conference, Sevilla, Spain, 2001.

[2] E. Breeuwer, R. Franco, M. Falcone, J.C. De Mateo, P. Erhard, and N. M. Hollreiser, "Simulation and Experimentation in the GALILEO System Design and Verification," in Proceedings of the GNSS 2002 Conference, Copenhagen, Denmark, 27-30 May, 2002.

[3] M. Falcone, F. Amarillo-Fernandez, J. Hahn, and H. Favin-Leveque, "The GALILEO System Test Bed for Early Prototyping and Experimentation of GALILEO Algorithms," in Proceedings of the ION GPS 2002 Meeting, Portland, USA, September, 2002.

[4] GSTB-V1 Statement Of Work (ESA-APPNG-SOW/00114), 5 October 2001.

[5] E. Detoma, S. Facioni, M. Gotta, G. Graglia, F. Cordara, L. Lorini, V. Pettiti, P. Tavella, and J. Hahn, "The Experimental Precise Timing Station (E-PTS) for the Galileo System Test Bed (Phase V1): Architectural Design and Experimental Goals," in Proceedings of the 34th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Reston, Virginia, USA, 2002.

[6] G. Graglia, E. Detoma, F. Cordara, R. Costa, L. Lorini, D. Orgiazzi, V. Pettiti, P. Tavella, and J. Hahn, "The Experimental Precise Timing Station and the Galileo System Time generation in the Galileo System Test Bed Phase V1," in Proceedings of the GNSS 2003 Conference, Graz, Austria, 2003.

[7] Galileo Joint Undertaking, "GALILEO Research and Development Activities, Activity E Galileo Mission Implementation," Statement Of Work (GJU/03/103/issue3/OM/ms), pp. 23-25.

[8] F. Cordara, R. Costa, L. Lorini, D. Orgiazzi, V. Pettiti, P. Tavella, G. Graglia, E. Detoma, and J. Hahn, "The generation of the Experimental Galileo System Time in the Galileo System Test bed V1," in Proceedings of the 2003 IEEE International Frequency Control Symposium (FCS) Jointly with the 17th European Frequency and Time Forum (EFTF), Tampa, USA, 2003, pp. 306-311.

[9] J. Laverty, P. Tavella, et al., "Report and recommendations of the GalileoSat Working Group on the Galileo Time Interface," ESA STR-243, ISSN 0379-4067, 2001.

[10] D. W. Allan, "Time and Frequency (Time-Domain) Characterization, Estimation, and Prediction of Precision Clocks and Oscillators," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 1987, vol. UFFC-34, n. 6, pp. 647-654.

[11] Symmetricom Inc., Specifications of MHM 2010 Active Hydrogen Maser.
(http://www.symmttm.com/products_pfr_ahm_overview.asp)

[12] J. Kouba, and P. Héroux, "Precise Point Positioning Using IGS Orbits and Clocks Products," in GPS Solutions, vol. 5 n. 2, 2001.

[13] P. Tavella, and D. Orgiazzi, "'Timing-oriented' processing of geodetic GPS data using a Precise Point Positioning (PPP) approach," in 6th Meeting of representatives of laboratories contributing to TAI, BIPM, France, 2004.
(http://www1.bipm.org/en/committees/cc/cctf/tai_symposium.html)

[14] J. Ray, and K. Senior, "IGS/BIPM pilot project: GPS carrier phase for time/frequency transfer and timescale formation," Metrologia, vol. 40, pp. S270-S288, 2003.

[15] "GSTB-V1 – Clock Reference Station Analysis" (GT-GAIN-TN-1122, issue 1.0), July 2004.

[16] G. Panfilo, and P. Tavella, "Preliminary test on the steering algorithm for keeping a time scale synchronized to UTC," in Proceedings of the 2003 IEEE International Frequency Control Symposium (FCS) Jointly with the 17th European Frequency and Time Forum (EFTF), Tampa, USA, 2003, pp. 301-305.

[17] F. Cordara, L. Lorini, V. Pettiti, P. Tavella, D. Piester, J. Becker, T. Polewka, G. De Jong, O. Koudelka, H. Ressler, B. Blanzano, and C. Karel, "Calibration of the IEN-PTB TWSTFT link with a portable reference station," in Proceedings of the 18th European Frequency and Time Forum (EFTF), Guildford, UK, 2004.

[18] W. Lewandowski, F. Cordara, L. Lorini, V. Pettiti, A. Bauch, D. Piester, and O. Koudelka, "A simultaneous calibration of the IEN/PTB time link by GPS CV and TWSTFT portable equipment," in Proceedings of the 18th European Frequency and Time Forum (EFTF), Guildford, UK, 2004.

[19] P. Tavella, J. Azoubib, and C. Thomas, "Study of the clock-ensemble correlation in ALGOS using real data," in Proceedings of the 5th European Frequency and Time Forum (EFTF), Besançon, France, 1991, pp. 435-441.

[20] F. Cordara, R. Costa, L. Lorini, D. Orgiazzi, V. Pettiti, I. Sesia, P. Tavella, P. Elia, M. Mascarello, M.Falcone, and J. Hahn, "Experimental Galileo System Time (E-GST): One Year Of Real-Time Experiment", in Proceedings of the 36th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Washington DC, 2004, pp. 105-122.

GALILEO TIME SERVICE PROVIDER: THE FOCAL POINT FOR EUROPEAN TIMING EXPERTISE

J. Hammesfahr[(1)], A. Moudrak[(1)], J. Furthner[(1)], A. Schroth[(1)], J. Dallat[(2)], R. Piriz[(3)], A. Bauch[(4)], C. Bruyninx[(5)], P. Defraigne[(5)]

(1) German Aerospace Center (DLR), Germany
(2) Computer Anwendungen für Management (CAM), Germany
(3) GMV, Spain
(4) Physikalisch-Technische Bundesanstalt (PTB), Braunschweig, Germany
(5) Royal Observatory of Belgium (ORB), Brussels, Belgium

ABSTRACT

The Galileo Time Service Provider (GTSP) is the entity that will act as the interface between Galileo and the global time standard, Coordinated Universal Time (UTC). This paper reviews the concepts behind the GTSP and identifies top-level design options. It reflects ideas of the authors regarding the work to be undertaken during the following three years of Galileo development until the In-Orbit Validation experiments.

1. INTRODUCTION

As GPS, Galileo offers not only positioning, but also time synchronization capabilities to its users. Thus, Galileo timekeeping has to deal with a specific metrological timekeeping function to enable dissemination of the Coordinated Universal Time (UTC) through Galileo. To support this function, the reference time scale of Galileo, the Galileo System Time (GST), has to be closely synchronized to UTC (but without introduction of leap-seconds that is done on UTC to maintain its conformance with the rotation of the Earth). The residual offset GST-UTC has to be determined and broadcast in the Galileo navigation message.

The continuously improved quality of GPS time dissemination services which broadcasts a real-time prediction of UTC derived with the help of the US Naval Observatory (USNO), sets the benchmark for Galileo. The ability of Galileo to offer a comparable or better accuracy of time dissemination will contribute to Galileo acceptance in the user community and widen the potential Galileo market.

To implement the metrological timekeeping function, Galileo will need to rely on European timing institutes which support the International Bureau of Weights and Measures (BIPM) in calculation and distribution of UTC. These institutes possess both scientific competence in the metrological time keeping and practical experience in maintaining real-time realizations of UTC (UTC is computed by BIPM as a "paper-clock" available only *post factum*).

Unlike the United States where hardware resources and timing expertise supporting GPS are concentrated at the USNO, these resources and expertise are spread amongst several institutes in Europe. Therefore, to meet the GPS benchmark, it has to be achieved that these institutes behave as a well-coordinated network producing high-quality timing products for Galileo. The networking effect would also stimulate concentration of the European timing expertise and, thus, ensure the competitiveness of Galileo. The cooperation between Galileo and the timing institutes will be organized and coordinated by the Galileo Time Service Provider (GTSP).

The role of the GTSP has been addressed in two definition studies:

- ESA-funded working group on Galileo timing interface (2000-2001)
- Galileo Mission Implementation Activity (2003-2005) which was launched by the Galileo Joint Undertaking (GJU) (see [2] for more details on the outcome of this study).

GTSP implementation – also funded by GJU – is to start this year.

The main objective of GTSP is the provision of a specific service to Galileo to link GST to UTC. This service comprises also coordination of the actors involved into the GTSP operation, and scientific support for Galileo on specific aspects related to the metrological timekeeping. Thus the GTSP has to bridge the commercial world of Galileo with the state-funded metrological institutes in Europe, and to become the focal point for the European timing expertise to ensure the competitiveness of Galileo. This paper presents an independent view of the GTSP architecture and organization developed by a team of research and timing institutes and commercial companies. It does not necessarily reflect the opinion of GJU.

From the team point of view, GTSP development should be driven by the following principles:

- High integration of existing European resources (expertise and facilities)

- Excellent performance (quality and availability of products etc.) of the GTSP surpassing the formulated requirements
- Meeting stringent requirements of Galileo facility operations as concerns reliability and security of GTSP services
- Close cooperation with the timing community from the outset and encouraging participation in GTSP activities for UTC laboratories worldwide (potentially, also non-European, i.e. Canadian, Russian, Chinese, Israeli laboratories etc.)
- Promotion of the Galileo brand and the Galileo services through initiation of a Galileo awareness program for the timing community and timing users
- Cost-efficient implementation and operations of the prototype GTSP through following the principles of sufficiency-for-purpose. The basic philosophy will be to acquire data from the UTC laboratories and the PTF without impacting their internal operations.

Translation of these principles into GTSP organization and architecture is discussed further on.

2. GTSP FUNCTIONS

In 2000, the GalileoSat Time Interface Working Group completed a preliminary study on the establishment of a system time scale for Galileo and its interface to the international timescale UTC. Following this activity, ESA has identified that the navigation function is safety critical (Galileo will provide Safety-of-Life navigation services) and should be kept within Galileo. On the other hand, the metrological function is non-critical for safety-of-life applications, but it is one of the key factors to ensure Galileo competitive abilities vs. GPS and, thus, its successful introduction into the market.

Separation of the responsibilities for the navigation and the metrological function was considered to be a way to allow Galileo to achieve a competitive level of timing performance. It was proposed to outsource the responsibility for the metrological function to GTSP, an external body with respect to Galileo, which has to establish a link between Galileo and the European timing community. Through this model, Galileo will be able to benefit from the long-term scientific experience and the high-quality products of the timing community.

Properly implemented, this solution can become mutually stimulating for both Galileo and the timing community, and hence will allow:

- to support the overall expertise within the Galileo program through the link with the European timing community
- to reduce Galileo operational costs and risks through re-use of products of European time laboratories, and as an additional effect
- to improve the cooperation within the European and the international timing community and create a focal point for perspective developments in the domain of timing metrology.

Obviously, the main objective of GTSP is to satisfy the relevant performance requirements, and to provide a robust, secure, and cost-effective operational solution. Also, the positive effect of GTSP activities onto the timing community needs to be carefully considered. GTSP activities may stimulate an improvement of the quality of UTC laboratories' products which in turn will improve the quality of Galileo services.

The main functions of GTSP are:

- to provide Galileo with information on GST-UTC offset to support timing services provided through Galileo
- to provide Galileo steering corrections to enable GST to stay in agreement with TAI/UTC (the timescale International Time Atomic (TAI) differs from UTC only by an integer number of seconds, these two timescales will not be distinguished further)
- to monitor the performance of the GST (and potentially, also performance of Galileo time dissemination).

The GTSP shall become responsible for organizing and processing data from the Galileo Precise Time Facility (PTF) (see [1] for details of PTF design), UTC laboratories, the BIPM and the International Earth Rotation Service (IERS) in order to fulfill these functions. Thereto, GTSP has to obtain access to existing time transfer equipment and time scales within the Galileo PTF and facilities of UTC laboratories.

Beyond the operational tasks to support GST generation through linking it to TAI, the GTSP can assist monitoring and validation of Galileo time products. GTSP has to assist in the building up of user trust in Galileo products and services through providing a competent and independent verification of the GST performance.

Finally, the contribution of the GTSP to the development of the Galileo business plan and its potential role in Galileo marketing requires careful considerations. On one side, the GTSP may allow Galileo to benefit from using the products of renowned European timing institutes; it will support establishing the Galileo's own brand-mark. On the other hand, supporting Galileo may become an important factor for the timing institutes themselves, advertising their research and innovation potential and stimulating their further development in a pan-European framework. Our philosophy is to propose an "open-architecture" GTSP which would be able to cooperate with numerous European timing institutes.

3. GTSP ACTORS AND ROLE

The GTSP operations will involve the three basic players:

- GTSP itself
- Galileo
- Timing community (represented by European UTC laboratories, the BIPM, and the IERS)

GTSP will act as an agent between Galileo and the timing community (see Figure 1) linking the GST which is produced at the Galileo PTF to the international reference time scale UTC/TAI which is computed by BIPM *a posteriori*. Real-time representations of UTC are produced by national timing institutes (further referred to as UTC laboratories).

GTSP has to collect data generated by the time transfer equipment at the PTF and UTC laboratories. From processing of these data and additional information from BIPM (offsets between the time scale of UTC laboratories and the UTC), GTSP will estimate the GST-UTC time offset and GST-UTC steering correction, and provide this information to Galileo PTF.

Beyond the exchange of data and calculation of the steering corrections, several administrative and supervising functions should be addressed:

- To forward the time transfer observations and clock measurements from the PTF to the BIPM
- To retrieve the leap second information from IERS and to provide this information to Galileo
- To arrange the collaboration with the involved timing laboratories
- To organize the operation of the TWSTFT time transfer links between PTF and the timing laboratories, including negotiations with satellite service providers
- To organize calibration exercises for initial and later periodical calibration of the GNSS Common View and TWSTFT links between PTF and the timing laboratories.
- To monitor the GST-TAI offset using historical data from BIPM

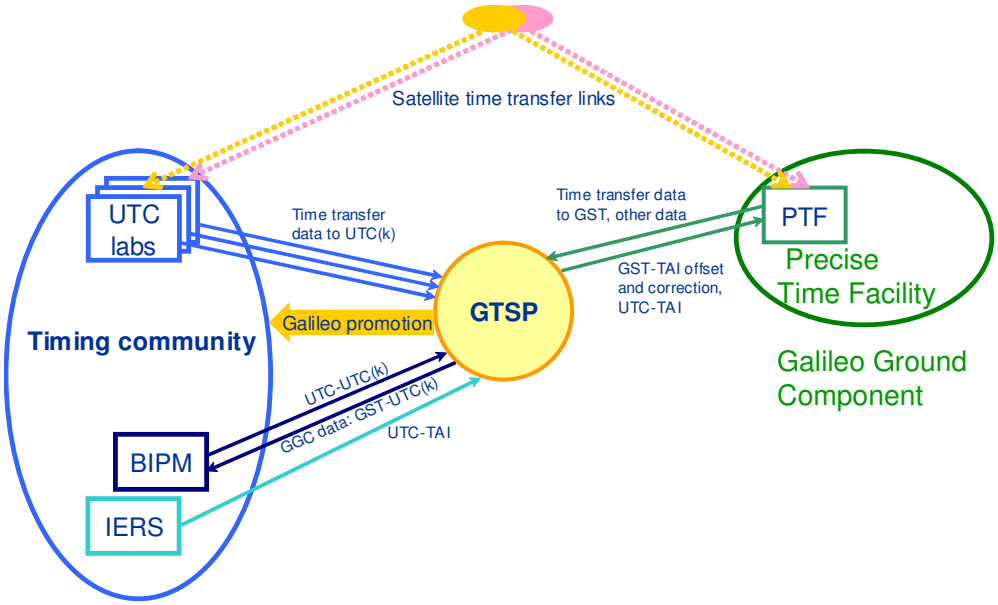


Figure 1. GTSP actors and interfaces

4. GTSP IMPLEMENTATION PHYLOSOPHY

According to our analysis, GTSP implementation activities should be concentrated on:

- ensuring the ability of the GTSP to fully support Galileo tests and verification activities during the Galileo In-Orbit Validation (IOV) phase, and to implement the presently developed version of GTSP (which is considered as a prototype) in such a way that its operation can continue, when necessary, over the IOV phase until Galileo reaches the Fully Operational Capability and potentially even longer;
- ensuring the performance (quality and availability of products etc.) of the GTSP to be considerably better than the specified performance requirements to strengthen Galileo competitive abilities;
- achieving the implementation and operations of the GTSP in a cost-efficient manner. The major goal should be to ensure that the GTSP can offer high-quality services for an attractive price to the Galileo Operating Company;

- establishing a close cooperation with the timing community from day one and making participation in GTSP activities open (also, in a flexible way) for all UTC laboratories (potentially, also non-European, i.e. Canadian, Russian, Chinese, Israeli laboratories etc) which are willing to commit to necessary performance standards;
- contribution to establishing and promotion of the Galileo brand, and promotion of Galileo services through supporting the development of the Galileo business plan and starting a Galileo awareness program for the timing community.

The following principles for implementation of the GTSP can be emphasized:

- GTSP should be integrated with the core Galileo infrastructure and shall be fully operational during the IOV phase to support the overall Galileo validation program;
- GTSP should act as an agent for the information and expertise provided by the European timing institutes which is required for Galileo to provide high-performance time services;
- GTSP should become at the same time an agent between Galileo and the community of time and frequency users;
- GTSP should be developed in such a way that its services will allow GST to meet relevant requirements immediately. In the same time, the GTSP should be able to rapidly adapt its services to growing needs of Galileo and its users and to technological developments in the timing metrology (new types of precise clocks, advanced time transfer techniques etc.).
- Preferably, GTSP should be implemented as an independent commercial entity. Accordingly, UTC laboratories should not be part of this entity, and should be linked externally to assure GTSP independency and flexibility to adapt to future developments;
- GTSP should make maximal use of already existing expertise and facilities in order to minimise development risks and costs;
- An important constraint on GTSP implementation is avoiding any negative impact on the routine operation of UTC laboratories, BIPM, Galileo and other actors involved.

5. GTSP MODELS

GTSP has to establish three important relationships:

- with Galileo,
- with UTC laboratories,
- with BIPM and IERS.

Relationship with Galileo will be governed by regulations of the Galileo Operating Company, and is a subject for negotiations later. Presently, it can be only anticipated that this relationship will have a commercial character.

BIPM and IERS are international organizations with public mandates. According to their status, it will be hardly possible for the GTSP to contract them for specific provisions related to timing products. The easiest solution for GTSP is to act on the same basis as any metrological institute presently, receiving BIPM and IERS products with no charge through public channels.

For the relation with the timing institutes one can consider the following models:

- Both, GTSP and UTC laboratories, are contracted by the Galileo operating company (GOC) directly,
- GTSP is contracted by the Galileo operating company, and the UTC laboratories are contracted in turn by the GTSP.

The second option seems to be favorable since it will simplify the task of the Galileo Operating Company, and will facilitate an effective collaboration between GTSP and the timing community.

The relationship between the Galileo Operating Company and the GTSP should be independent of any other sub-contracts placed on an operating entity – the GTSP should remain as an agent between the UTC laboratories and the Galileo Operating Company and should ideally not be subjugated to either party. This solution can ensure the quality of GTSP products and the ability of the GTSP to perform its work in a commercial manner, but with the necessary scientific independence so essential for the long-term evolution of this activity.

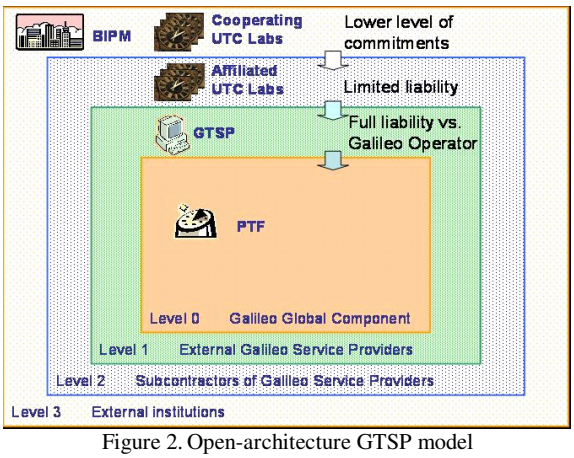


Figure 2. Open-architecture GTSP model

We have elaborated an "open-architecture" GTSP model. The corresponding liability layers are illustrated in Figure 2. The proposed model foresees two kinds of relationship with UTC laboratories: affiliation and cooperation.

Affiliated laboratories are to be contracted to ensure GTSP performance. They should guarantee a certain level of product quality, and accept certain regulations and restriction on operations of the time transfer equipment and data transfer to GTSP. Affiliated laboratories are to be selected based on the performance of their timescales. It is deemed that four laboratories are sufficient to ensure the necessary level of GTSP performance and to provide the necessary redundancy.

Cooperating laboratories are those volunteering to support GTSP operations. These laboratories could be involved on a free-of-charge basis to further increase the quality of the GTSP products and to facilitate its consolidating effect on the European timing community. To enable this model of involvement of cooperating laboratories, GTSP has to ensure promotion of its activities, so that participation in them becomes a sign-of-quality for timing institutes.

6. GTSP ARCHITECTURE

The basic idea of GTSP operations should be to acquire data from the UTC laboratories and the Galileo PTF without disturbing their internal operations.

GTSP will perform the necessary processing on datasets provided by these entities and provide its products to Galileo as required. Another responsibility of the GTSP is delivering of timing products from Galileo PTF to BIPM in order to ensure GST visibility for the metrological community.

A potential sharing of responsibilities between the actors involved into the GTSP operations is illustrated in Table 1.

The GTSP may include the following elements:

- GTSP processing facility and GST monitoring facility being responsible for
  - control and monitoring of GTSP operations
  - acquisition of relevant data from specific sinks: Galileo, UTC laboratories, BIPM, IERS
  - data processing and generation of GTSP products
  - distribution of GTSP products and other necessary data to clients (Galileo, UTC laboratories, BIPM, other potential users)
- GST monitoring facility being responsible for the *a posteriori* monitoring of the GST performance vs. TAI
- Time transfer links between UTC laboratories and PTF
- Interfaces (data links) to UTC laboratories, Galileo, BIPM, IERS, and potential users

A cost-effective solution for implementation of time transfer links between PTF and UTC laboratories would be to re-use already available time transfer equipment at UTC laboratories (TWSTFT stations and/or GPS Common View receivers which are used for time transfer in the frame of regular operations of the laboratories). Compatible time transfer equipment meeting metrological standards and practices should be installed at Galileo PTF. In addition, the laboratories and the PTF should be responsible for operation of their time transfer equipment and provision of the time transfer data to GTSP.

Another cost effective option is to organize the data provisions from the PTF and the laboratories through password-protected FTP servers which can be accessed by the GTSP (this is also a standard practice of BIPM and IERS for making their products available to the metrological community). GTSP products may be made available to Galileo in a similar manner.

To support Galileo promotion and dissemination of GTSP products to Galileo user community, GTSP may establish a dedicated Web-interface (a dedicated Web server) where these products will be made available to potential users. The server may become a part of the GTSP processing facility.

Responsibility of laboratories	Responsibility of GTSP	Responsibility of Galileo
Ensure availability of time transfer equipment	Coordinate time transfer observation (negotiate schedule)	Ensure availability of time transfer equipment
Establish data links	Coordinate/organize calibration campaigns; support calibration of time transfer equipment at PTF	Provide access to PTF time transfer equipment for calibrations
Support calibration of time transfer equipment	Collect data from UTC laboratories, Galileo, BIPM and IERS	Establish data links to GTSP
Perform time transfer observations with PTF following the schedule	Compute GST-to-TAI offset and GST-to-TAI steering correction	Perform time transfer observations with UTC laboratories following the schedule
Make time transfer data available to GTSP	Make computational results available to Galileo	Make time transfer data (and other relevant data) available to GTSP

Responsibility of laboratories	Responsibility of GTSP	Responsibility of Galileo
Monitor performance of real-time UTC representations and report failures and/or unavailability to GTSP	Provide PTF clock measurements and time transfer data to BIPM	Inform GTSP about operations planning
Support trouble-shooting	Monitor GST performance vs. TAI	Support trouble-shooting
	Monitor quality of data from PTF and UTC laboratories and provide feed-back	
	Coordinate operations with Galileo; lead trouble-shooting	

Table 1. Sharing of responsibilities between GTSP actors

The GTSP overall architecture is illustrated in Figure 3. The proposed solution relies on long-term practices in timing metrology and will have minimal impact onto the regular operations of all GTSP actors, thus leading to minimization of GTSP development costs and associated risks in the spirit of fit-for-purpose philosophy adopted by the authors.

Furthermore, it may be emphasized that the GTSP should make use of the primary products of UTC laboratories – the real-time representation of UTC they maintain. The mandate of the laboratories is to ensure a high quality and availability of these real-time representations. Through their utilization, GTSP can benefit from the expertise and technical excellence of the laboratories, since in the leading laboratories the real-time UTC representations are generated from novel ultra-stable clocks (like Caesium fountains) or by elaborated timescale algorithms which compute an average of multiple commercial clocks. In addition, making use of the real-time UTC representations will allow the GTSP to become a natural extension of the European metrological infrastructure, and achieve a clear allocation of responsibilities between itself and the timing laboratories without violating their intellectual property rights.

The technical feasibility of this approach and the ability of the GTSP to meet and even to best the required performance level have been demonstrated in [3].

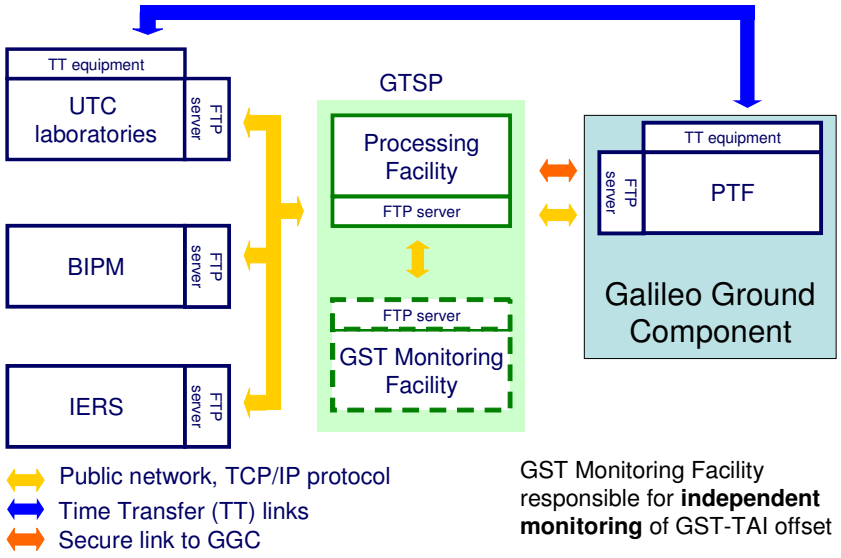


Figure 3. GTSP architecture

REFERENCES

[1] S. Bedrich, A. Bauch , J. Laverty, A. Moudrak, W. Schäfer, “Design of the Galileo Precise Time Facility”, Proceedings 18th EFTF, 2004.

[2] J. Laverty, “The Galileo Time Service Provider: A progress report”, Proceedings 19th EFTF, 2005.

[3] A. Bauch, A. Sen Gupta and E. Staliuniene, “Time scale prediction using information from external sources”, Proceedings 19th EFTF, 2005.

THE GALILEO TIME SERVICE PROVIDER: A PROGRESS REPORT

John Laverty
National Physical Laboratory, Hampton Road, Teddington, Middlesex, TW11 0LW, United Kingdom

Andreas Bauch
Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany

Patrizia Tavella, Franco Cordara, Diego Orgiazzi
IEN Galileo Ferraris, Strada delle Cacce, 91, 10135 Torino, Italy

ABSTRACT:
In this paper we describe recent work undertaken as part of the EC Framework Programme 6 to better define the Galileo Time Service Provider (GTSP), the entity that will act as the interface between Galileo, Europe's planned contribution to global satellite navigation systems, and Coordinated Universal Time (UTC), the global standard for time.

1. INTRODUCTION

The Galileo Mission Implementation (GEM) project in EC Framework Programme 6 is led by France Development Conseil (FDC) and provides Galileo with wide-ranging support on standardisation and certification issues, radio spectrum allocations and mission development issues. Here we report on one of the GEM work packages concerning the Galileo Time Service Provider (GEM WP5210), which was undertaken jointly by IEN, PTB and NPL. The GEM study was a precursor for the development and implementation of a GTSP for the in-orbit validation (IOV) phase of Galileo planned for 2007/2008. As such, this paper can be viewed as an interim report on the progress towards an operational GTSP.

2. THE GALILEO TIME SERVICE

2.1 An Overview of Galileo's Timing Functions

As with other Global Navigation Satellite Systems, Galileo will establish its own reference system time, known as Galileo System Time (GST). GST will be formed from a composite of the clocks that are operated in the Galileo Precise Timing Facilities (PTF), Galileo's own timing stations based on hydrogen masers and caesium atomic clocks. Other clocks in the Galileo system, such as those on-board the satellites, will be synchronised to GST via pseudo-range measurements linked to the PTF and processed by Galileo's Orbit Determination and Time Synchronisation (ODTS) software. We understand that the possible inclusion of these other clocks in the generation of GST will be considered after the ODTS process is operational and its estimation errors have been fully characterised.

The GST specifications state that it will be steered to TAI with a maximum offset of 50 ns for 95% of the time (over any yearly time interval) and that the difference will be known with an uncertainty of 28 ns (95% confidence). The uncertainty stems mainly from the need to predict GST-TAI for real time applications, because International Atomic Time (TAI) is known definitively only in deferred time with a maximum delay of 45 days. The offset between GST and UTC will be an integer number of seconds, information that Galileo will broadcast as part of its time service. In terms of the links between GST and TAI, it has been decided that the Galileo Operator will be responsible for systems necessary for the navigation mission, while a Time Service Provider will be responsible for the additional facilities necessary to provide a UTC time service via Galileo. The metrological time-keeping function for the Galileo time service will be performed through co-operation with the Galileo Time Service Provider (GSTP) and with selected European UTC(k) laboratories.

2.2 Galileo development phases

The main development phases going forward are:

In-Orbit validation (IOV) Phase: Its scope is to early validate/verify Galileo (segments, operations and service provision) before the full deployment of the overall system. The IOV results will finalise the definition and development of the segments and operations. The Galileo Development and In-Orbit Validation (IOV) Phase is subdivided into three main sub-phases, namely:

- Phase C (detailed design)
- Phase D (manufacturing)
- Phase E1 (initial deployment, system validation and operations)

During the IOV phase, there will be only one operational PTF, which will be physically implemented at the Galileo Control Centre (GCC). In addition, there will only be a limited number of Galileo satellites during the IOV phase. The consequences on the GTSP design from these constraints are that:

- The GTSP is likely to have an important role in validating the PTF at IOV, via comparisons with UTC(k) laboratories;

- It will be difficult to test the GTSP's role in PTF-PTF handover procedures at IOV;
- The GTSP will not be able to rely on Galileo common-view receivers for time transfer in IOV, and hence will have to make use of either TWSTFT or GPS common-view techniques.

Long Term Operations Phase: Its scope is to provide guaranteed services to end users, dealing with the provision of the full Galileo service and the maintenance/replenishment of the constellation. The current baseline design considers a minimum phase of 20 years operation. During the eventual Full-Operational Capability (FOC) phase, a hot back-up PTF, identical to the master PTF will be implemented in a hot back-up GCC and steered to the master PTF. At FOC, it will be possible for the Galileo operator to switch the role of master or slave between the two PTFs.

The impact on the GTSP design when moving from IOV to FOC is than that:

- The GTSP will have an important role in assessing GST-TAI offsets, but Galileo will be able to cross-compare PTF performances and therefore to detect and act on faults itself;
- It will be possible to test the GTSP's role in PTF-PTF handover procedures at FOC;
- The GTSP will be able to rely on Galileo common-view receivers for time transfer in FOC, allowing a choice between TWSTFT, Galileo and GPS common-view techniques.

2.3 Galileo Time Service Provider Developments

The GTSP issue was considered by ESA's GalileoSat Working Group on the Galileo Time Interface (GWGGTI) between 1999 and 2000 [1] and provided a number of recommendations on the way forward for the GTSP.

The Galileo System Test Bed V1 (GSTB V1) ran from 2002 to 2005 and was used to test time scale algorithms for the Galileo PTF, both for GST generation and for steering GST to TAI. The core of the GSTB V1 timing experiment is based at IEN, building on existing UTC(IEN) facilities. The UTC(PTB) and UTC(NPL) facilities are also employed in the experiments, mainly in support of calibration experiments (PTB) and for TAI steering experiments (both PTB and NPL). The results of these GSTB V1 experiments are reported elsewhere at EFTF05 [2].

The GEM Project reported here is a precursor study before the main development of the GTSP for the Galileo IOV Phase, which is the subject of detailed contract negotiations at the time of writing.

3. THE GTSP REQUIREMENTS

The Galileo Time Service Provider (TSP) is the entity that will act as the interface between Galileo and the global time standard, Coordinated Universal Time (UTC). In its simplest form the key GTSP requirements can be stated as follows:

(i) To provide Galileo with UTC information for the Galileo time service;
(ii) To provide steering corrections to enable GST to stay in agreement with TAI;
(iii) To monitor the performance of the Galileo Time Service.

In practice, the UTC information will be passed to Galileo as a TAI-GST offset and a UTC-TAI offset (i.e. leap second information). The accuracy requirements for the Galileo Time Service dictate that the GTSP will have to predict TAI in advance of the official publication of the BIPM Circular T. In the remainder of this section, we identify the key components needed to achieve these top-level requirements, according to four functional categories, and comment on which components were analysed further in order to guide the GTSP design process.

Data Measurements: The raw data for the GTSP function are the offsets between various clocks in the UTC and Galileo timing systems, based on measurements of time transfer links and phase offsets between the clocks at each participating laboratory. Together, they provide a data set that maps the behaviour of all the measured clocks relative to each other, which can be used to meet objectives (i) and (ii) above.

- UTC(k) and PTF clock data
- UTC(k) and PTF time transfer data

The first set of data is needed only for composite clock schemes while the second set is needed for all GTSP data processing schemes – see section 5 below for more explanation. In addition, to validate the Galileo Time Service from a user perspective, it is necessary to measure the following parameter:

- UTC(k) – GST data from Galileo signal-in-space measurements

System Reports: In addition to the raw measurements, the GTSP will need to track system changes that introduce phase discontinuities (e.g. introducing a different phase delay in a measurement when replacing a cable; recalibrating a time transfer receiver; etc.)

- UTC(k) and PTF clock steering reports
- UTC(k) and PTF configuration change reports
- Calibration reports

Also, data will have to be gathered from other organisations such as the BIPM or IERS:

- UTC – TAI reports from the IERS Bulletin C
- UTC – UTC(k) reports from the BIPM

Data Processing: The computation of the TAI predictions is an essential precursor to computing the products identified in objectives (i) and (ii) above:

- TAI predictions (e.g. TAIp – UTC(k))

The outputs associated with objectives (i) and (ii) above need to be computed:

- UTC – GST offset computations *(e.g. based on UTC(k) – GST measurements, TAI predictions and reports from the IERS and BIPM)*
- GST steering computations

Finally, the output associated with objective (iii) needs to be computed:

- UTC(k) – GST computations from Galileo signal-in-space measurements

Data Management: Finally, arrangements need to be in place to share, validate and archive the various data sets and products associated with the GTSP:

- data communications (to/from Galileo, BIPM, IERS, UTC(k), users, etc.)
- data validation
- data archive

Many of the components identified above are well known and – from a top-level view – offer a clear design path to be followed. For example, the IERS and BIPM services in support of UTC are well established and therefore the Galileo GTSP will simply adapt to those existing services. Equally, there are some components that offer detailed design choices, but where top-level input here is not required. For example, the GTSP can leave the choice of counter timers and phase comparators for clock data measurements at the UTC(k) and PTF to the designers of those systems without undue risk. Also, the details of the data management functions are assumed to be uncontroversial and hence were not analysed further in this study. However, there are some issues at the interfaces of the Galileo and UTC laboratories where the GTSP needs to ensure that proper coordination takes place. This is particularly true of the time transfer links. A superficial view suggests simplicity, but a more careful analysis reveals potential for unnecessary duplication of equipment and effort and, at worst, incompatible time transfer implementations at the UTC(k) and PTF. This issue is explored further in section 4. In addition to the GTSP interfaces, there are some core GTSP responsibilities that would benefit from a top-level design justification. Specifically, we have in mind the algorithms for TAI predictions and GST steering computations. These are discussed further in section 5.

Finally, we note that the GEM study reviewed these ideas in relation to the following Galileo Mission Segment (GMS) documents: The GMS Requirements Document [3], the GMS External Interface Requirements Document [4], and the GMS Time Service Provider Interface Control Document [5]. The length of this conference paper prevents a detailed review of the different GTSP requirements. However, we note that the requirements as currently formulated remain high-level and leave considerable room for interpretation. It seems inevitable that the GTSP requirements will evolve with the PTF development during CDE1 phases of the GalileoSat programme, and be defined in more detail to address issues such as service availability, data provision delay, and detailed data contents (detailed format, resolution of measurement values, report of measurement uncertainties etc.).

4. SELECTION OF TIME TRANSFER LINKS

The GEM study considered two issues in relation to time transfer links: (i) whether the selected time transfer techniques were fit for purpose and (ii) how to avoid the unnecessary duplication of requirements for time transfer facilities for Galileo, GPS and UTC?

The first issue was addressed by comparing the stability of the caesium clocks and hydrogen masers against the candidate time transfer techniques: Two-Way Satellite Time and Frequency Transfer (TWSTFT), Geodetic GNSS Time Transfer and GNSS Common-View. The different time transfer techniques were adequate for the steering targets for GST (|TAI-GST|<50 ns). However, there are implications on how often steering decisions can realistically be made. Noise processes in active hydrogen masers start to become visible across TWSTFT and Geodetic time transfer links after averaging times of one day. Across a common-view link, the averaging time is extended to between 3 and 10 days. Hence, while common-view links should be able to support the GST-TAI target of 50 ns, it is better to use TWSTFT and/or Geodetic GPS with time scales based on active hydrogen masers – particularly if the Galileo time service is to emulate the performance of GPS.

The second issue was addressed in a review the various Galileo and UTC interfaces that require time transfer links – including some outside the strict scope of the study but which may influence design choices for the GTSP. The interfaces considered were:

- The Galileo-to-GPS interface, to measure the Galileo/GPS Time Offset (GGTO) parameter [6]

- The PTF-to-PTF interface, to synchronise the GST outputs from the two Galileo PTF at FOC

- The PTF-to-ODTS interface, to synchronise the Galileo satellite clocks to GST via pseudo-range measurements taken by Galileo Sensor Stations (GSS)

- The UTC(k)-to-UTC(k) interfaces, to measure the time offset between the various UTC(k) time scales.

The baseline for each of these links is shown in Table 1, both for Galileo IOV and FOC phases.

A concern here is that there are potentially three different TWSTFT networks (i.e. for the PTF, GGTO and UTC) and that the GTSP might generate a need for a fourth TWSTFT network. The cost of TWSTFT satellite time can be up to 2 MEuro p.a. for a dedicated transponder (although actual cost can be much less depending on the details of the transponder usage). Similarly, although less costly than TWSTFT, there may be a need to duplicate GNSS timing receivers at the PTF and UTC(k) laboratories unless arrangements are made share data from receivers and to harmonise on common data formats.

Discussions with the Galileo Project Team provided guidance that has helped to narrow down these options. Firstly, it became clear that Galileo Sensor Stations (GSS) would not be made available for the GTSP application, but data from the GPS timing receivers installed at the PTF for the GGTO application could be provided. Secondly, Galileo would not lease a TWSTFT transponder for the GTSP in the IOV phase, although the TWSTFT equipment installed at the PTF for the GGTO measurements could, in principle, be made available for GTSP applications. Similarly, it was hoped that equipment sharing at UTC(k) laboratories might be negotiated so that the GTSP would be able to draw on data from the PTF and UTC(k) for its computations.

The solution identified for the PTF-UTC(k) links following this analysis draws on the facilities already in place for UTC and for Galileo's other time interfaces. For the Galileo IOV phase, the hope is that the single Galileo PTF will join the UTC TWSTFT network if agreed with Intelsat and UTC laboratories, with the GTSP responsible for processing the sub-set of data necessary for its application. This preferred solution recognises that the European and North American UTC laboratories have access to their transponder on a "pre-emptable" basis (i.e. the UTC(k) laboratories can lose that access if demand rises from other fee-paying customers). As a back up for the Galileo IOV phase, the GTSP would aim to process GPS common-view data from receivers at the PTF and UTC(k) laboratories. For the Galileo FOC phase, the hope is that the Galileo TWSTFT network will be used, with selected UTC(k) laboratories joining that network. The change of TWSTFT network from the IOV phase would offer the GTSP greater security than that available on a "pre-emptable" basis. As a back up for the Galileo FOC phase, the GTSP would aim to process GPS and Galileo common-view data from receivers at the PTF and UTC(k) laboratories.

5. GTSP DATA PROCESSING APPROACH

Four approaches to the TAI predictions were considered in the study.

"EGNOS model": EGNOS Network Time (ENT) is computed on the basis of all clocks in the system and steered to GPS system time. Using this model as an analogy, the core GST algorithms would then take responsibility for managing the steering process, with the GTSP tasked to provide reliable TAI predictions as an input to the GST algorithm. However, in practice, Galileo is working to a different model and therefore this approach was excluded.

"UTC(k) reference model": A second approach would be to use a UTC(k) reference laboratory to determine the offset between GST and UTC, via time transfer link measurements and the BIPM Circular T. The data could then be used to predict future values of the (TAI – GST) time offset, using the stability of GST or the UTC(k) as the underpinning 'flywheel' until the next BIPM Circular T is published. This particular approach has merit because of the relative simplicity of the scheme. Some care is needed to avoid single point failure (e.g. more than one UTC(k) needs to be used as a reference) and to allow sufficient redundancy for quality control of the data (i.e. that one has sufficient independent reference points – clocks and/or time scales - to detect anomalous behaviour in advance of the next BIPM Circular T). The main limitation of the technique is that it uses only a sub-set of the clocks in the participating laboratories for its (TAI - GST) predictions.

"Distributed GTSP Processing model": A third approach would be to distribute the TAI prediction responsibilities between a number of laboratories, with the GTSP required to carry out the computations for the PTF and to then combine the predictions to a common conventional value. In this scenario, each UTC(k) laboratory supporting the GTSP would generate a free running time scale based on *all* of its atomic clocks, along with a prediction of the offset of this composite timescale from TAI. GST would then be steered to a weighted mean of each of the UTC(k) laboratories' prediction of TAI, along with the TAI offset predictions obtained from the two PTF laboratories. In principle, this approach uses the stability of all the available clocks to minimise the offset of GST from TAI. The division of responsibilities also allows for consistency checks in the various TAI predictions to improve the integrity of

the overall prediction (although instabilities in the base time scales dictate that there will always be small scale divergences). This particular approach is feasible in principle but, was excluded because it is not currently a standard approach in UTC(k) laboratories and therefore would require significant coordination effort to implement effectively.

"Centralised GTSP Processing Model": A fourth approach would be for the GTSP to operate a single TAI prediction algorithm based on all the clocks in the two PTF and the participating UTC(k) laboratories. Each UTC(k) laboratory and the two PTF laboratories would report individual clock data to the GTSP. A free running GTSP composite clock would then be generated. The (TAI–GST) timescale offset is then predicted each month shortly after the publication of the BIPM Circular T, and GST is then steered to this prediction of TAI. This method should produce a more accurate prediction of TAI since we are combining all clocks in a single algorithm, although its value at short averaging time may be reduced because of correlated noise in common time transfer systems. The approach has merit in terms of more reference points for quality control and the use of all available clocks in the prediction to achieve the most accurate (TAI – GST) prediction possible. However, these benefits have to be weighed against the greater data processing complexity than that of the other viable scheme, the UTC(k) reference model.

In conclusion, our analysis presented two viable schemes with which the GTSP concept could be developed: The **UTC(k) reference model** and the **Central GTSP Processing model.** The choice between them is a balance between a relatively straightforward implementation of the GTSP concept but without using the full potential for the underlying hardware, versus a more complex data processing scheme that can take advantage of that underlying hardware.

6. FUNCTIONAL ANALYSIS AND ARCHITECTURE DEFINITION

The study structured information on the GTSP in terms of the different functions to be undertaken and the architecture necessary to carry out those functions. This information format is designed to guide the engineering design work to follow. It helps to clarify how the GTSP requirements (in term of operations, performance, liability, etc.) map to specific GTSP functions and then onto its architecture sub-systems. A diagram showing the core GTSP architecture is presented in Figure 1. Ensuring a simple but modular framework, here each facility implements one of the core functions allocated to GTSP, such as the data management (T-DMF), the GST to TAI steering correction computation (T-GPF) and the independent monitoring of the resulting steering performance (T-IPMF). In particular, the latter one could actively contribute to the GTSP tuning process. Unfortunately, the length of this conference paper prevents a detailed explanation of each element but it is hoped that the titles are self-explanatory.

7. OTHER GTSP EXTERNAL INTERFACES

While this study concentrated on the core interfaces for the GTSP (i.e. BIPM, UTC(k), IERS and Galileo links), it also noted the following as yet undefined interfaces that may be considered as the GTSP is developed into an operational entity:

Users: The GTSP is currently specified as a service to Galileo rather than to users of the Galileo time service. Galileo documents [4] imply that information on its services will be provided to users via a Galileo service centre rather than via the GTSP. It is not obvious from a first principles approach that this is the intended solution, a point flagged in the study reports;

EGNOS: The interface with the EGNOS time provision is relevant to the users of Galileo and EGNOS, but it has not been defined in this study;

EUROMET is a forum for discussing time metrology activities in Europe, and hence it may be appropriate to establish liaison mechanisms in future, even if such links are not essential for the technical operation of the GTSP;

Standards and Accreditation Bodies: We noted that various standards organisations refer to UTC and GPS for time traceability in their recommendations. As with EUROMET, it may be appropriate to establish liaison mechanisms in future even if such links are not essential for the technical operation of the GTSP.

8. CONCLUSIONS

In this paper we have described the work undertaken in the GEM Project of EC Framework Programme 6, the precursor study before the main development of the GTSP for the Galileo IOV Phase. That next phase is the subject of detailed contract negotiations at the time of writing. We note that the Team selected to take the GTSP development forward to the Galileo IOV phase in 2007/2008 will need to address the following:

- Detailed interpretation of Galileo requirements;
- Review and refinement of the Architecture and Development plan;
- Development and implementation of the GTSP, including communication links and data processing algorithms;
- The active involvement of the BIPM, IERS and UTC(k) laboratories, in the GTSP operations.

REFERENCES

[1] J. Laverty, P. Tavella and J.Hahn, UTC and Galileo Time Services: a report from the Galileosat Working Group on the Galileo Time Interface, in Proceedings of the EFTF Conference, 2001, pp. 261-266.

[2] F.Cordara, R.Costa, L.Lorini, D.Orgiazzi, V.Pettiti, I.Sesia, P.Tavella, P.Elia, M.Mascarello, M.Falcone and J.Hahn, Galileo System Test Bed V1: results on the Experimental Galileo System Time, in Proceedings of the EFTF Conference, 2005.

[3] Ground Mission Segment Requirements, Galileo reference RQS/GAL/0165/GLI, version 9, 26/03/2004.

[4] GMS External Interface Requirements, Galileo reference GAL-IDB-GLI-SYST-A-0369, version 4 Draft A, 28/07/2004.

[5] GMS-Time Service Provider Interface Control Document, Galileo reference GAL-ICD-GLI-SYST-I/0677, version 1A, 06/08/2004.

[6] GPS/Galileo Time Offset Preliminary Interface Definition, Issue 01 Rev01, 20/03/2003.

Link\Status	Time Transfer methods at Galileo IOV		Time Transfer methods at Galileo FOC	
	Primary	Backup	Primary	Backup
PTF-ODTS	Galileo GSS		Galileo GSS	
PTF-PTF	not applicable		TWSTFT (PTF network)	Galileo GSS
PTF-GPS	TWSTFT (GGTO network)	GPS/Galileo Geodetic Time Transfer	GPS/Galileo SIS comparisons	TWSTFT (GGTO network)
UTC(k)-UTC(k)	TWSTFT (UTC network)	GPS CV	TWSTFT (UTC network)	GPS/Galileo CV
PTF-UTC(k)	TWSTFT (UTC network)	GPS CV	TWSTFT (PTF network)	GPS/Galileo CV

Table 1: Time transfer links as foreseen at Galileo's In-Orbit Validation (IOV) phase and for Full Operational Capability (FOC).

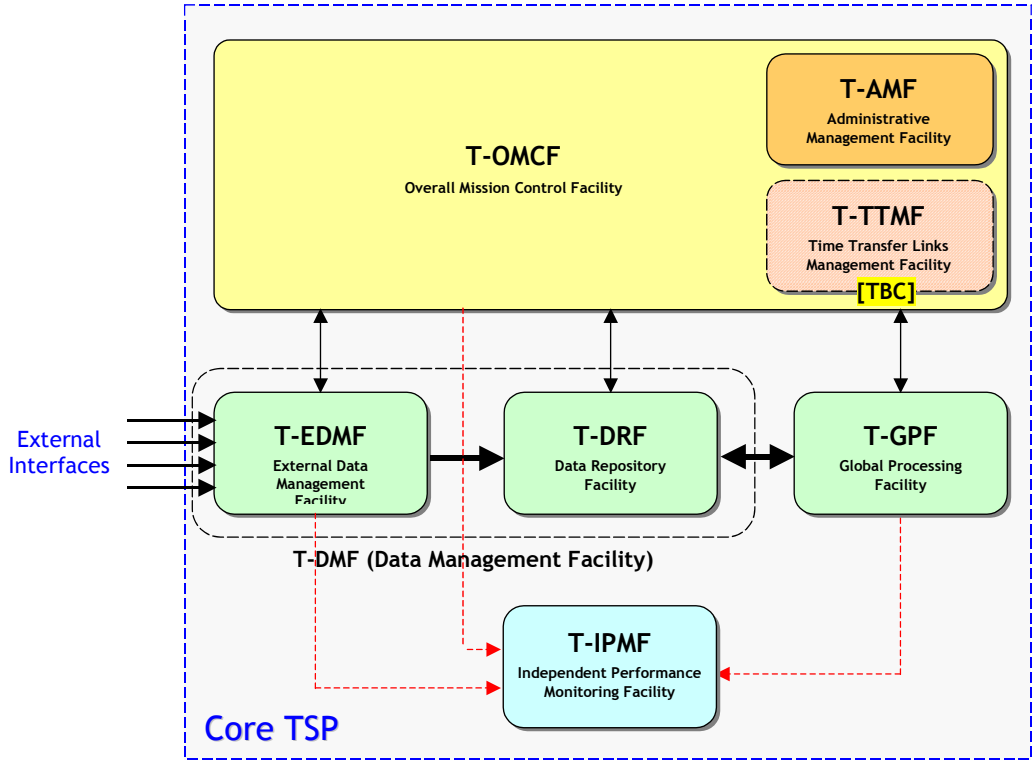


Figure 1: Core GTSP architecture overview

CALIBRATION OF GPS TIME RECEIVERS IN POLAND

J. Nawrocki[1], P. Nogaś[1], A. Foks[1], P. Mielnik[2], W. Lewandowski[3]

[1] Astrogeodynamical Observatory Borowiec, Space Research Centre, 62-035 Kórnik, Poland; e-mail: nawrocki@cbk.poznan.pl
[2] Military University of Technology, 00-908 Warsaw, Poland, e-mail: pmielnik@wel.wat.edu.pl
[3] Bureau International des Poids et Mesures, Pavillon de Breteuil, 92310 Sèvres, France; e-mail: wlewandowski@bipm.org

Abstract

Time and frequency activities in Poland are conducted within a consortium of seven time laboratories belonging to public service and research, industry, and military. National time laboratory of Lithuania contributes also to this work. The main research work concentrates on development of stable and accurate national time scales UTC(PL), UTC(AOS) and TA(PL). Contributing laboratories have a strong interest in access to accurate time with an uncertainty of a few nanoseconds. This requires differential calibration of GPS time receivers used in these laboratories against a reference receiver. This paper reports on the calibration exercise conducted by the Astrogeodynamical Observatory (AOS) using a TTS-2 portable receiver. The reference receiver was another TTS-2 operating at the AOS. This receiver was calibrated several times in the past by the BIPM. Estimated uncertainty of the reported here differential calibration ranges from 3 ns to 5 ns.

1. INTRODUCTION

Polish Atomic Time Scale TA(PL) was officially started on June the 1st, 2001. At present it is based on the indications of about 14 atomic clocks installed at several public and commercial institutions [1]. In the realization of the TA(PL) also participates the Semiconductors Institute (LT) from Vilnius, Lithuania, with two cesium standards. The participating laboratories are referenced to Central Office of Measures (GUM) using multi-channel GPS receivers – TTS-2 developed by the Astrogeodynamical Observatory in Borowiec (AOS) in co-operation with the Bureau International des Poids et Mesures (BIPM) [2]. Using the same TTS-2 receiver Time and Frequency Laboratory of GUM is connected to the UTC system through the Physikalish-Technische Bundesanstalt (PTB) in Braunschweig, Germany – Figure 1 [3].

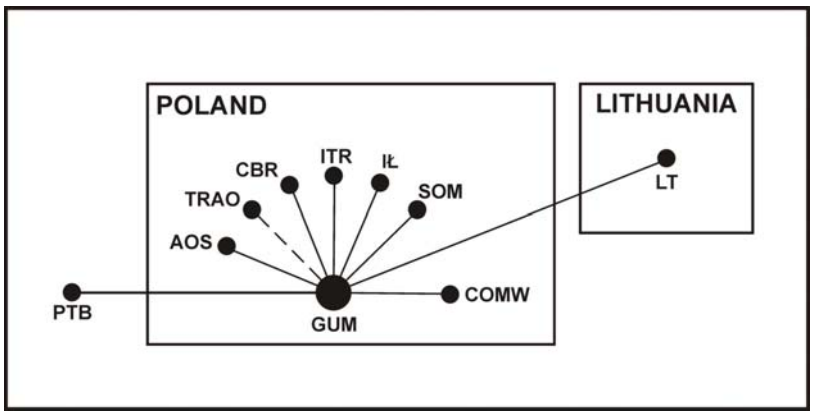


Figure 1. Organization of TA(PL) links.

TA(PL) is realized as weighted average of the participating clocks, using algorithm similar to ALGOS [1]. The main research work concentrates on development of stable and accurate national time scales UTC(PL), UTC(AOS), and TA(PL). TA(PL) is obtained as the weighted average of about twelve atomic clocks working at seven laboratories, and two clocks in Lithuania. The long-term stability of the TA(PL) is in the range of 2*E–15. Now, developments concentrate on predicting and forming UTC(PL) and UTC(AOS) using TA(PL). Soon, instead present monthly, a weekly solution for TA(PL) will be used. This should allow keeping Polish realizations of UTC within twenty nanoseconds of UTC computed by the BIPM. A contribution of Polish UTCs to GALILEO Time System Provider is also considered.

The participation to TA(PL) network allows connection of local atomic clocks to Polish Official Time UTC(PL) [4]. Contributing laboratories have a strong interest in access to official time with an uncertainty of a few nanoseconds. This requires differential calibration of GPS time receivers used in these laboratories against a reference receiver.

Three of the participating laboratories the AOS, GUM and LT provide their own realizations of UTC(k) [5] and directly participate in TAI and UTC and were several times calibrated by the special BIPM 'travelling' receiver (Table 1.).

Table 1. Clocks and time transfer receivers at laboratories participating to TA(PL), and dates of past calibrations.

Laboratory	Clock	Receiver	Rec. SN	Dates of Past Calibrations
Space Research Centre (AOS)	1 Cs	TTS-2	Prototype 3	1996*, 2000*, 2002*, 2003*
Polish Telecom (CBR)	2 Cs	TTS-2	12	-
Central Office of Measures (GUM/PL)	3Cs, 1 Hm	TTS-2	23	1996*, 2002*
Institute of Telecommunication (IŁ)	2 Cs	TTS-2	11	-
Institute of Tele- and Radio Electronics (ITR)	1Cs	TTS-2	22	-
Air Defence Centre of Metrology (SOM)	1Hm	TTS-2	30	2003**
Metrology Centre of Polish Army (COMW)	1Cs	TTS-2	34	2003**
Semicoductors Institute (LT), Vilnius, Lithuania	2 Cs	TTS-2	20	2002*

* Calibrated by the traveling BIPM receiver.
** Calibrated after assembling at the AOS.

This paper reports on the calibration exercise conducted by the AOS using a TTS-2 portable receiver. Estimated uncertainty of the reported here differential calibration ranges from 3 ns to 5 ns.

2. DIFFERENTIAL CALIBRATION EXERCISE

From 17 June 2003 to 10 July 2003 the AOS has conducted a differential calibration exercise using a special TTS-2 portable receiver with antenna and antenna cable. The reference receivers were two other TTS-2 operating at the AOS and GUM. These receivers were calibrated several times in the past by the BIPM [6]. Visited laboratories during this exercise were the IŁ, CBR, GUM, and AOS.

The SOM and COMW have the receivers, antennas and the connecting cables, assembled and calibrated at the AOS before shipment, and did not participate to this exercise. The ITR could not participate due to some technical difficulties at the time of this calibration.

At each side, the portable receiver antenna was placed in the vicinity of the local antenna with its coordinates estimated with a few centimeter uncertainty, and the portable receiver was connected to the same clock as local receiver. The block diagram of the calibration set-up at the AOS is presented in Figure 2.

All receivers involved in this exercise used 0.5 V trigger level for measuring the beginning of the rising 1pps pulse. The antenna cable delay was measured using pulse, method with the Stanford Research SR-620 counter steered by the frequency from cesium HP standard [7].

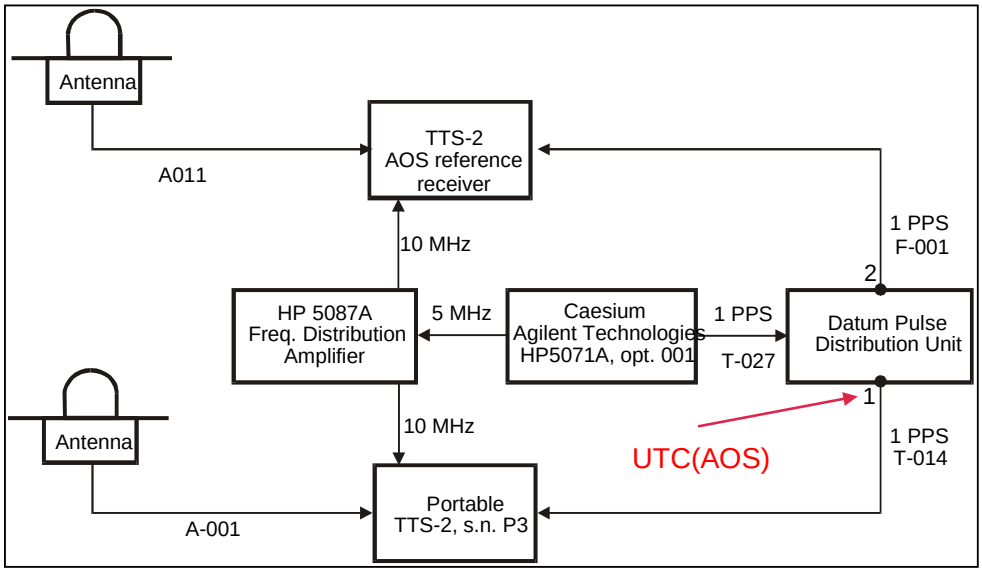


Figure 2. The block diagram of the calibration set-up at the AOS.

3. RESULTS

Classic common view method was used. First, for each track *i* observed in laboratory *l*, the time differences were computed between portable calibration receiver CAL and local receiver LOC:

$$dt_{l,i} = (Clock(l) - GPS\ time)_{CAL,i} - (Clock(l) - GPS\ time)_{LOC,i}$$

Then, for each of the visited laboratories, the noise exhibited by the time series dt_l was analysed, by the use of the modified Allan variance. In each case, white phase noise was exhibited up to an averaging interval of at least one day (Figure 3.).

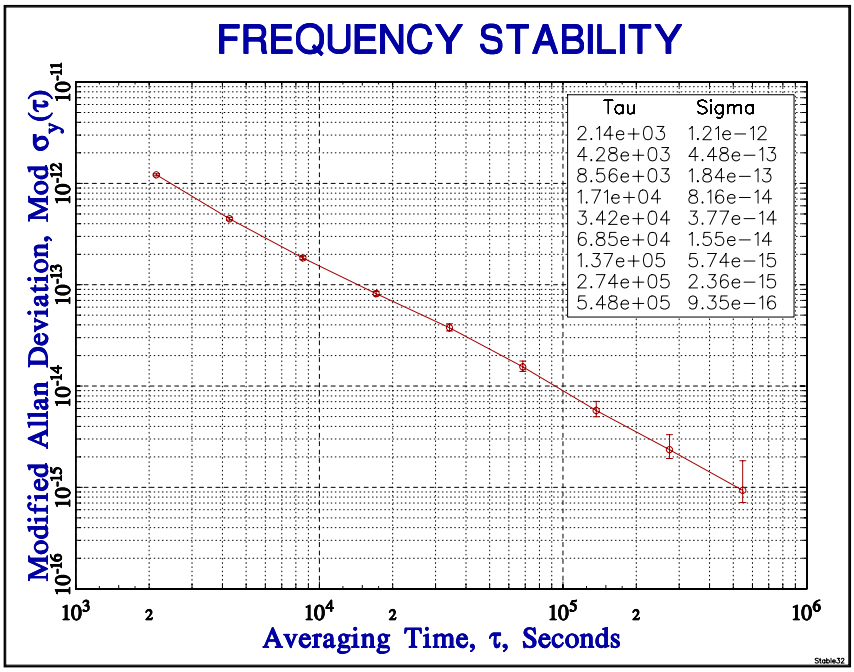


Figure 3. Square root of the modified Allan variance of time series dt_{AOS} for the period 4 July 2003 (MJD 52824) to 10 July 2003 (MJD 52830).

The averages for one day are shown in Figure 4. The uncertainties of the one-day averaging are, as well as obtained offsets are given in Table 2. At the end, the mean offsets for full interval of observation at each laboratory were computed (see Table 3).

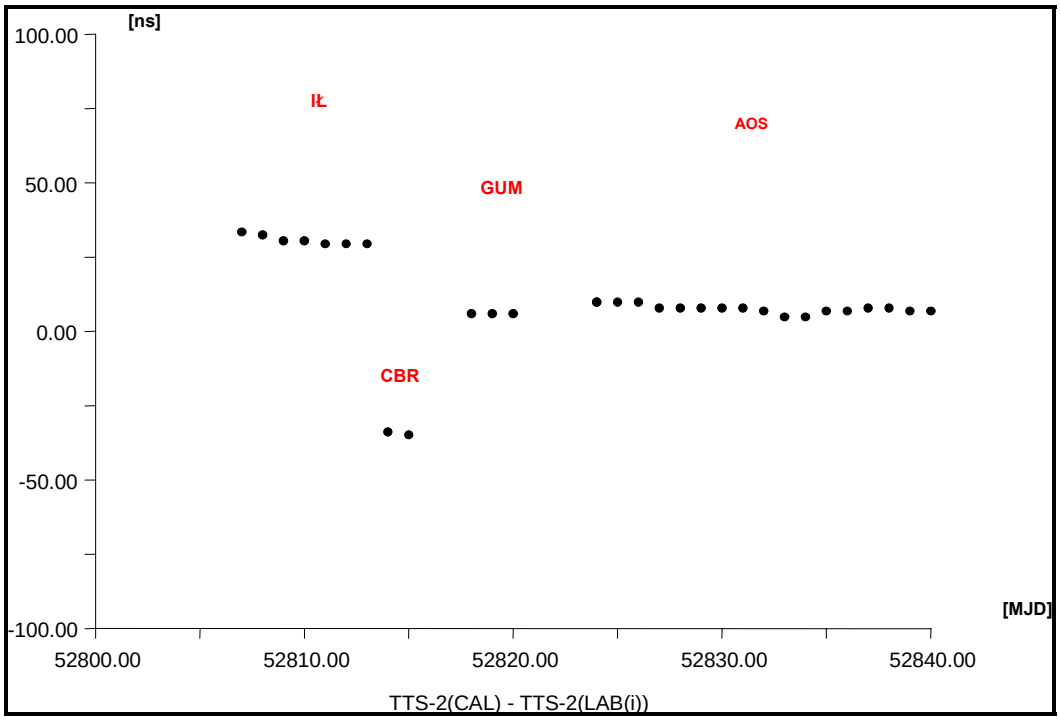


Figure 4. Daily averages of $dt_{l,i}$.for each laboratory.

Table 2. One-day offsets obtained for each laboratory during the campaign.

Institute of Telecommunications - IŁ			
MJD	**Mean offset /ns**	**RMS /ns**	**Common views**
52807	34	3	275
52808	33	2	471
52809	31	2	445
52810	31	3	459
52811	30	3	507
52812	30	2	442
52813	30	2	170
Research and Development Centre of Polish Telecom - CBR			
MJD	**Mean offset /ns**	**RMS /ns**	**Common views**
52814	–34	5	174
52815	–35	4	208
Central Office of Measures – GUM/PL			
MJD	**Mean offset /ns**	**RMS /ns**	**Common views**
52817	6	3	200
52818	6	3	483
52819	6	3	532
Astrogeodynamical Observatory - AOS			
MJD	**Mean offset /ns**	**RMS /ns**	**Common views**
52824	10	2	433
52825	10	2	434
52826	10	2	127
52827	8	2	241
52828	8	2	419
52829	8	2	466
52830	8	2	455

Table 3. Mean differences between portable and local time receivers for the full duration of comparison.

CAL – LAB(i)	Mean offset /ns	RMS /ns
CAL – IŁ	+31.0	2
CAL –CBR	–34.0	4
CAL – GUM/PL	+6.0	3
CAL – AOS	+9.0	3

Then we derived from Table 3 differential time corrections which should be made (added) to time differences derived during the GPS comparisons of the clocks kept by the laboratories. The results are summarized in Table 4. We have chosen as reference laboratory the GUM/PL which realizes Polish Official Time UTC(PL) [4]. The GPS equipment of this laboratory is

also well calibrated by the BIPM. Corrections related to another laboratory, for example the AOS, might be derived from Table 3 in the similar way. The uncertainties given in this table are conservative. They are mainly driven by the uncertainty due to the instability of GPS receiving equipment hardware. The uncertainty for the CBR is higher. This might be to some changes in the hardware during the calibration.

It is interesting to notice that the delays of GPS receiving equipment of two laboratories the GUM/PL and the AOS are in agreement within involved uncertainties. The equipment of both laboratories were calibrated several times in the past by the BIPM (see Table 1).

Table. 4. Differential time correction *d* to be added to[*Laboratory clock (i)* – *UTC*(PL)], and its estimated uncertainty *u*(*d*) for the period of comparison (1σ).

[*Laboratory clock* (*i*) – *UTC*(PL)]	d/ns	*u*(*d*)/ns
[*Clock*(IŁ) – *UTC*(PL)]	+25	3
[*Clock*(CBR) – *UTC*(PL)]	–40	5
[*UTC*(AOS) – *UTC*(PL)]	+3	3

4. Preparation For Absolute Calibration

Space Research Centre together with the Military University of Technology in Warsaw prepare the absolute calibration of GNSS timing receivers. The aims are:

- To prepare absolutely calibrated TTS-2 C/A receiver for future calibration trips.
- To estimate how the different hardware configurations change the internal delay of the TTS-2 receivers.
- To absolutely calibrate the new TTS-3 receiver. This is a multi-system (GPS/GLONASS/EGONS:WAAS), double frequency unit able to carry the C/A and P-code observations. To obtain the assumed uncertainty of one ns for time transfer, it is necessary to measure the internal delays for C/A code, P1 and P2 observations. The delays for L1 and L2 frequencies are especially important for the P3 ionosphere-free solution (to remove the errors caused by the ionosphere for one have to now the L1 and L2 delays. GLONASS satellites use not only L1 and L2 frequency channels, each satellite uses its own different frequency.

For this absolute calibration the Spirent Communications STR-4760 is used. This is a 16-channel, single frequency simulator with C/A and P1 code. The L2 (P code) is expected to be added in 2005. The calibration will be done separately for the receivers, for the antenna cables and for the antennas.

5. Conclusions

The differential GPS calibration exercise conducted in Poland in June and July 2003 provided valuable results. Involved laboratories have now accurate access to Polish Official Time UTC(PL) [4] with an uncertainty of several nanoseconds. It is expected that in the near future other calibration exercises will be organized to check reported here results and calibrate laboratories which were not yet visited.

Concerning absolute calibration of TTS-2 and TTS-3 receivers, there is a good progress, and in the late 2005 first results are expected.

6. REFERENCES

1. J. Azoubib, J. Nawrocki, W. Lewandowski, “Independent atomic time scale in Poland – organization and results, *Metrologia*, 2003, 40, issue 3, pp. S245-S248.

2. J. Nawrocki, W. Lewandowski, J. Azoubib,"Time Transfer with GPS multi-channel Motorola Oncore Receiver using CCDS standards", *Proceedings of the 29th Annual Precise Time and Time Interval (PTTI) Meeting*, 2-4 December 1997, pp. 319-328.

3. Annual Report of the BIPM Time Section (2003), 2004.

4. Government act on official time in Republic of Poland from 10 December 2003, *Journal of Official Gazette,* 2004, number 16, position 144.

5. BIPM *Circular T* (monthly).

6. W. Lewandowski, "Determination of the differential time correction between GPS time equipment located at the Observatoire de Paris, Paris, France, and the Astronomical Latitude Observatory, Borowiec, Poland ", *BIPM Report -95/12*, September 1995, 11 pages.

7. G. de Jong, "Measuring the propagation time of coaxial cables used with GPS receivers," *Proc. 17th PTTI,* pp. 223-232, December 1985.

DEVELOPMENT OF THE $^{113}Cd^{+}$ ION ATOMIC CLOCK

B. M. Jelenković, S. Chung, J. D. Prestage, L. Maleki
Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91108, USA.

ABSTRACT

We describe the set up and operation of a lamp-based, buffer-gas cooled $^{113}Cd^{+}$ ion apparatus that is developed towards the realization of an atomic frequency standard. We present measurements of the microwave spectra of the $^{113}Cd^{+}$ hyperfine resonance. The resonance width obtained during 5 sec microwave interrogation time is 0.17 Hz at 15.1999862906 GHz. The estimated short-term stability of the $^{113}Cd^{+}$ ion clock is 5 x $10^{-13}\tau^{-1/2}$.

1. INTRODUCTION

The motivation for the development of the new Cd^{+} atomic clock is to test for a possible variation of the fine structure constant, as previously proposed for the SpaceTime mission [1]. Flying two (or more) atomic clocks through the strong gravitational field of the Sun and measuring changes in the ratio(s) between the frequencies of different clocks, operating based on hyperfine transitions, will be a unique test of the variation of the fine structure constant. As shown [2] a search for a variation in α on the gravitational potential is possible by comparing the rate of drift of two clocks that are based on two elements. A time variation of the ratio of transition frequencies, i.e., the hyperfine interaction constants A, for two elements with atomic numbers Z_1 and Z_2 is

$$\frac{d}{dt}\ln\frac{A_1}{A_2} = [F(Z_1) - F(Z_2)]\frac{1}{\alpha}\frac{d\alpha}{dt}$$

F(Z) is a strong function of α for high Z nuclei that can be used to detect the temporal or spatial (d/dt replaced by d/dU, where U is the solar gravitational potential) variation in α . For two clocks, based on Cd^{+} and the existing Hg^{+} [2] clock the difference $F(Z_{Hg}) - F(Z_{Cd}) = \sim 1.7$ [1].

The lamp-excited atomic frequency standards based on buffer gas cooling of trapped ions in a rf linear trap have been demonstrated to be amongst the most stable available standards [3]. The prerequisite for the lamp-excited frequency standards is the efficient optical pumping of the clock ion. This requires an element with one isotope with nuclear spin I=0 (lamp ion) and another with a simple hyperfine level structure (clock ion). Also, for efficient optical pumping the excited levels of both isotopes have to spectrally overlap within the Doppler broadened emission from the lamp and absorption by the trapped ions. Figure 1(a) shows the schematic energy level diagram of $^{113}Cd^{+}$ and $^{106}Cd^{+}$ isotopes. Figure 1(b) illustrates an estimated overlap of the trapped ion $^{113}Cd^{+}$ ion absorption profile with the $^{106}Cd^{+}$ ion emission profile. The first is based on an assumed trapped-ion temperature of 500 K and the second on an assumed lamp-ion temperature of 1000 K. The two graphs in Fig. 1(b) are for transitions from $^{2}P_{1/2}$ and from $^{2}P_{3/2}$. The dotted curves on the graphs are for the Doppler broadened absorption lines from excited hyperfine levels of the ^{2}P state. Pumping of the $^{113}Cd^{+}$ ion to F=0 hyperfine level of the ground state can occur via both D1 and D2 line. Based on this, Cadmium is a promising candidate for the lamp-based trapped ion frequency standard, similar to the Hg^{+} atomic clock [3].

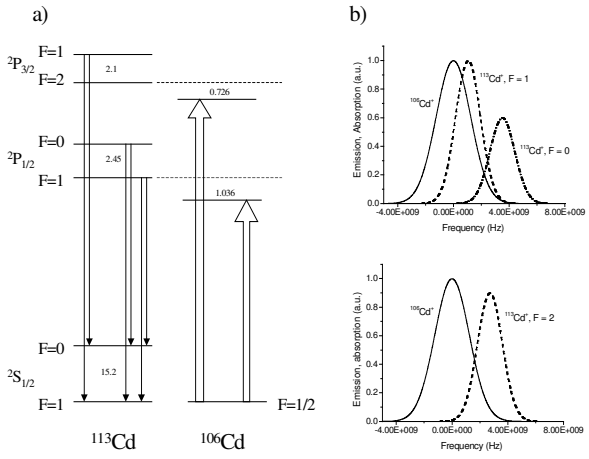


Figure 1. a) $^{113}Cd^{+}$ and $^{106}Cd^{+}$ ion energy level diagram and the pumping scheme The numbers indicate energy separation in GHz. b) Emission from $^{106}Cd^{+}$ lamp, illustrated at 1000 K (solid curve) and trapped $^{113}Cd^{+}$ ion absorption to $^{2}P_{1/2}$ (top) and $^{2}P_{3/2}$ levels (bottom graph).

2. EXPERIMENT

2.1 Ion trap

The $^{113}Cd^{+}$ ions are confined in a linear quadrupole Paul trap. We use a similar linear trap geometry as that described in [4]. The ions are produced by electron ionization of Cd atoms vaporized from a isotopically enriched solid piece of ^{113}Cd placed in an oven. The oven is typically heated to temperatures ranging from 50-80 ^{0}C. A piece of ~5 mg ^{113}Cd metal lasts for many months of nearly continuous oven operation. Typically, the trap rf amplitude was at 150 V_{pp} at 1 MHz for radial confinement of ions, with 12 V on the end caps to confine ions in the axial trap direction. The trap was inserted into the stainless steel vacuum chamber evacuated to a background pressure below 6 x 10^{-8} Pa. For the data presented here a He buffer gas was used at a pressure range of ~ 3 x 10^{-3} Pa to aid ion loading and cooling. Microwave radiation entered the trap region through a vacuum window and propagates through the trap nearly perpendicular to the trap axes and to the direction of the UV light from the rf lamp. The trap region was not shielded from the ambient magnetic field.

2.2 Cadmium rf lamp

Our Cd rf lamp is a quartz bulb (12 mm in diameter and 30 mm long) inserted into a copper rf cylindrical resonator with resonant frequency at ~160 MHz. The bulb has ~ 1 Torr of Ar and 1- 2 mg of ^{106}Cd metal piece. The UV spectra from the lamp are shown in Fig. 2. These spectra were obtained in a separate experiment by sending the UV lamp light to a monochromator and subsequently to a photomultiplier detector. Cd I line is much stronger than CdII D1 and D2 lines and is

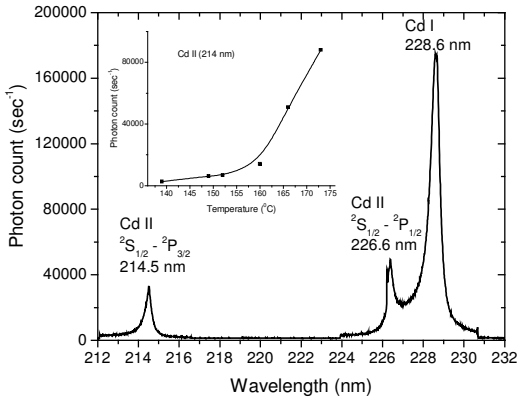


Figure 2. UV spectrum obtained from ^{106}Cd lamp. The insert shows the intensity of Cd II line at 214.5 nm vs the lamp temperature.

separated from the D2 line at 226.6 nm by only 2 nm, and can be a source of a high UV stray light background.

2.3 Optics and detection

Two six inch diameter spherical mirrors were used to focus the UV light from the ^{106}Cd lamp to the trap region and to collect the scattered light from the ions on the detector (Fig. 3). The mirrors were coated [5] so that their reflectance has a maximum of ~97% at 214 nm, 50% at 226 nm and 20 % at 228 nm. This reduced the light from Cd I line to ~ 4%, in two reflections from these mirrors. The light signal was detected using a photomultiplier. Figure 3 shows schematically the whole setup of the Cd^{+} ion atomic clock., optics, detection, microwaves, and the ion trap cross section.

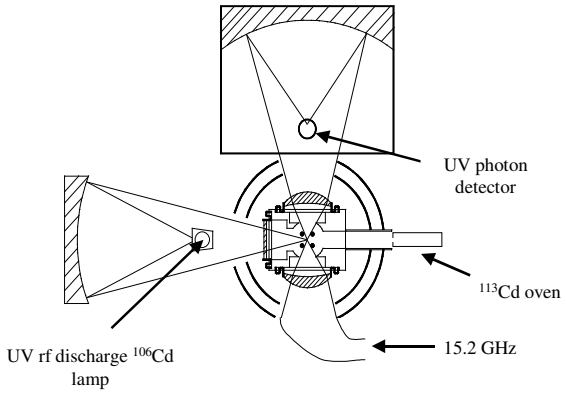


Figure 3. Schematic of the Cd^{+} ion lamp-based atomic clock

The microwave interrogation signal at ~15.2 GHz was produced by summing a 950 MHz signal power with ~ 7 MHz signal. The resulting signal was subsequently introduced into a step recovery diode and the 16-th harmonic of the signal was used. The + 7 MHz sideband of the 16-th harmonic signal was swept for frequency scan around the hyperfine resonance. Both the 950 MHz and the signal at ~7MHz were referenced to a hydrogen maser.

3. MICROWAVE SPECTROSCOPY

Figure 4a) shows the fluorescence sequence as the ions are first loaded in the trap and then the electron gun was switched off. The 1/e change of the ion number occurs during ~ 85 min. The lamp rf power, its temperature, and therefore the light intensity were

kept constant throughout the sequence. At the end of the sequence, after ions were dumped from the trap by grounding an end cap electrode, the fluorescence level returns to nearly the same level as at the beginning of the sequence, indicating stable lamp operation. In Fig. 4b) the microwave signal at ~15.2 GHz was turned on for 2 s while the lamp was dim. After the lamp goes bright again the fluorescence shows an overshoot which happens while the ions relaxed into the F=0 level. The counting gate was 0.5 s.

The high power line shape of the microwave spectra of $^{113}Cd^{+}$ ions in the rf trap at ~15.2 GHz is shown in Figure 5. The figure shows the broad spectrum of $^{113}Cd^{+}$ at different microwave power levels. The ion interrogation time was 1 sec. The top graph shows the motional side bands, the lower frequency axial sidebands near the resonance, and much weaker high frequency sidebands due to the transverse secular motion across the cloud diameter. The middle graph, at reduced power shows only near-in motional

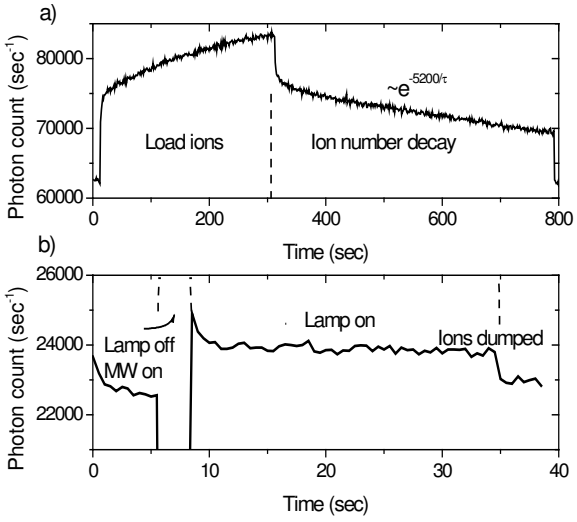


Figure 4. Fluorescence sequence of a) ion loading with electron current at 90 μA and ion number decay measurements and b) the end of ion loading, lamp dim and microwave on for 2 sec, then lamp on and photon count with 0.5 sec gate.

sidebands, only the Doppler broadened pedestal due to propagation of component(s) along the direction of the trap axes. Under certain conditions, i.e, the specific propagation and polarization direction of the microwave field with respect to the trap axis and external dc field, the sideband structure will vary [6]. The spectrum in the lower graph, obtained with a lower microwave power, shows the narrow resonance due to the Lamb-Dicke regime that applies to all ions excited by

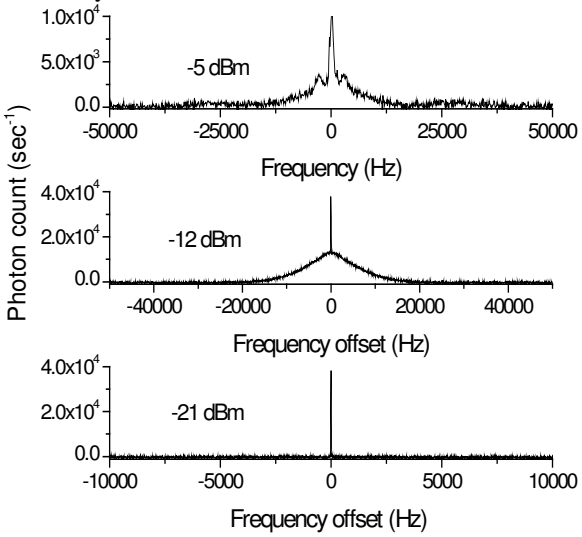


Figure 5. Spectra of the hyperfine transitions around 15.2 GHz for various microwave power.

the field components propagated perpendicular to the trap axes. There is no pedestal here since the applied power is too low to detect transitions from Doppler broadened different velocity classes, with reduced numbers of ions.

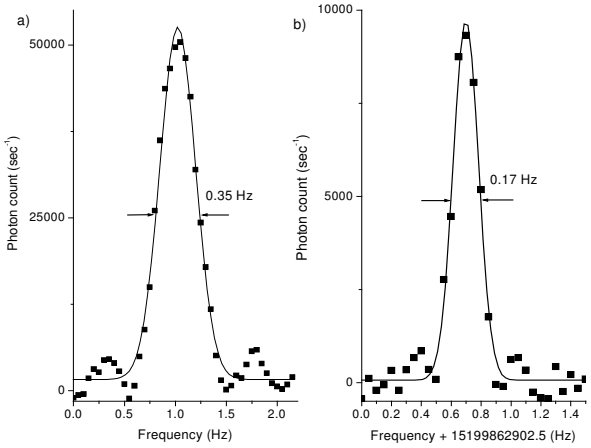


Figure 6. Rabi microwave resonance for ion interrogation time of a) 2.5 sec and b) 5 s. The microwave frequency steps are 50 ms and the counter gate was 1 s. There is no averaging for the data on graphs.

4. RABI RESONANCES

Rabi-type microwave-optical double resonances spectra are shown in Figure 6. Each point on the graphs was obtained after applying the following sequence at each frequency. For the signal plus background: lamp bright (optical pumping) – lamp dim – microwave interrogation– lamp bright – photon count. The background was measured with the same sequence, but microwave power off, and then subtracted from the photon count with microwaves on. The resonance in Figure 6a) was obtained with the microwaves applied for 2.5 s and its width is 0.35 s. For the resonance in Figure 6b), microwaves were applied for 5 sec and show width of 0.17 s. The resonance frequency is 15.1999862906 GHz. This value is 48 Hz higher then the value previously reported for zero magnetic field [7] and is due to unshielded ambient magnetic fields from the Earth's field and nearby electronics, etc. The short term clock stability estimated from σ_y

$$\sigma_y(\tau) = \frac{\Delta f}{\pi f_0 SNR}\sqrt{\frac{T_c}{\tau}}$$

is 5 x 10^{-13} $\tau^{-1/2}$. Here SNR is determined from the signal size measured at half maximum of the resonance and T_c the time of the total measuring cycle.

4. CONCLUSION

We have described the construction and the performance of a lamp-pumped trapped $^{113}Cd^+$ ions apparatus which will be used as a frequency standard. We have shown that optical pumping time with the rf lamp with the $^{106}Cd^+$ rf lamp pumping the $^{113}Cd^+$ clock ions to the dark F=0 state in about one second. The preliminary investigation of the performance of the clock give very promising characteristics like a good short-term stability of the order of 5 x 10^{-13} $\tau^{-1/2}$.

ACKNOWLEDGMENTS

We would like to thank Maggie Beach for her help with the lamps, and Robert Hamell for his help with microwaves. One of us (BMJ) thanks the National Research Council for Senior NRC Fellowship through the NASA-NRC program. The research in this work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with NASA.

REFERENCES

[1] L. Maleki, "Clock Test of Relativistic Gravity at Four Solar Radii", in 220th WE-Heraeus, Gyros, Clocks, and Interferometers: Testing General Relativity in Space, 22-27 August, 1999, Bad-Honeff, Germany. In "Testing Realtivisit Gravity in Space: Gyroscopes, Clocks, Interferometers", C. Laemmerzahlm C. W. F. Everitt, F. W. Hehl eds., (Berlin, Springer-Verlag 2001); J. D. Prestage and L. Maleki, "Precision clocks in space and α- variation", Proc. 2nd International Conference on Particle and Fundamental Physics in Space, Washington, D.C., USA, Nuclear Physics B (Proc. Suppl.) 134, 163 (2004).

[2] J. D. Prestage, R. L. Tjoelker and L. Maleki, Phys. Rev. Lett. 74, 3511 (1995).

[3] R. L. Tjoelker, J. D. Prestage, G. J. Dick and L. Maleki, "Recent Stability Comparisons with the JPL Linear Trapped Ion Frequency Standard", Proc. 1994 IEEE Frequency Control Symposium, p. 739.

[4] J. D. Prestage, R. L. Tjoelker, G. J. Dick and L Maleki, J. Mod. Optics 39, 221 (1992).

[5] Acton Research Corporation.

[6] V. Ender, Ph Courteille, W. Newhauserand, R. Blatt, J. Mod. Optics 39, 325 (1992).

[7] U. Tanaka et al., Phys. Rev A 53, 3982 (1996).

Session 1A / Piezoelectric resonators, filters and materials

FREQUENCY-TEMPERATURE CHARACTERISTIC OF FLEXURAL RESONATORS USING THE MODEL OF TIMOSHENKO

H. HANTZ, F. STHAL and R. BOURQUIN
FEMTO-ST Institute, UMR CNRS 6174, Dept. LCEP
ENSMM, 26 chemin de l'Epitaphe, 25030 BESANCON Cedex, FRANCE
hhantz@ens2m.fr

ABSTRACT

The classical Bernoulli's approximation is usually used to compute the frequency of a flexural resonator (like tuning fork resonator) and to determine its frequency-temperature behaviour. This approximation is valid for beams with a large length/width ratio. For short beams, the effects of shear stress and rotatory inertia may play a significant role for temperature compensated resonators. These effects have been taken into account by Timoshenko for the prediction of frequency of resonance of isotropic beam. The aim of this paper is to discuss the extension of Timoshenko's shear coefficient in the case of an anisotropic material and to compute the frequency-temperature characteristic of a X + θ cut resonator when the shear stress and the rotatory inertia have been taken into account. Comparisons between the classical (approximated) and the present "exact" treatments are given for quartz. Furthermore, we compare the numerical predictions obtained by using the different sets of data available for thermal sensitivities of elastic coefficients.

1. INTRODUCTION

For beam vibrating in flexure mode, when the length/width ratio becomes small, the influence of shear stress cannot be neglected, mainly if we want to consider the thermal behavior of such resonator. This kind of problem has been first treated by Timoshenko [1]-[2] for isotropic beam. Even if the correction on the value of the resonant frequency calculated from the classical Bernoulli's theory is small, it can induce a significant change of frequency temperature behaviour for quasi thermally compensated resonators. The main difficulty to extend the Timoshenko's theory for anisotropic materials, like quartz, is to compute the shear coefficient in such a case. There exist several generalizations of the shear coefficient. In this paper, we consider two of them. The first one is given by Kawashima [3], who use the Cowper's definition [4] of the shear coefficient in the case of a singly rotated beam around the X axis (X-Cut) vibrating in the XY' plan. The thermal coefficient of a resonator predicted by using this coefficient is given as a function of length/width ratio and of cut angle θ. The results are compared with a computation of the second one shear coefficient based on an energy definition [4] in the case of a beam purely oriented along the Y axis.

2. THERMAL BEHAVIOUR OF A RESONATOR ACCORDING TO TIMOSHENKO THEORY

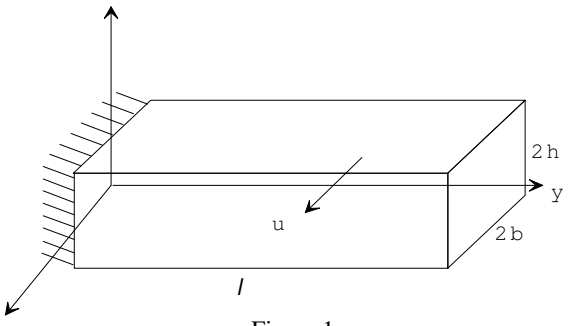


Figure 1

In static Timoshenko's flexure theory, the rotation of the cross section of the beam is taken into account and the shear stress components T_4 and T_6 are related to the bending force P. Since the resonant frequency of flexural resonator is low, one considers that these relations, established in the static case, remain valid. By definition of the Timoshenko's shear coefficient, one has

$$\int_A T_6 \, dA = k \, A \, G \, S_6 \qquad (1)$$

where A is the cross section and G the shear modulus.
The vibration of the beam at the frequency ω is governed by the equation [2]:

$$\frac{\partial^4 u(y)}{\partial y^4} + \left(\frac{m}{kAG} + \frac{m r^2}{Y I_z} \right) \omega^2 \frac{\partial^2 u(y)}{\partial y^2} + \frac{m^2 r^2 \omega^4}{kAGY I_z} u(y) = 0 \qquad (2)$$

where u(y) is the mechanical displacement, Y the Young modulus, and

$$I_z = \frac{4}{3} b^3 h$$
$$m = \rho A \qquad (3)$$
$$r^2 = \frac{b^2}{3}$$

the rotating inertia, the mass per unit length and radius of gyration, respectively. By putting

$$u(y) = u_0 \, e^{(\lambda \, y / l)}$$
$$\alpha = \frac{r^2}{l^2}$$

$$\eta = \frac{Y I_z}{kAG\, l^2} = \frac{Y\alpha}{kG}$$
$$\Omega = 3\frac{\rho\, \omega^2 l^4}{Y\, b^2} \qquad (4)$$

in (2) one obtain the relation between the reduced frequency Ω and λ:

$$\lambda^4 + \Omega(\eta+\alpha)\lambda^2 - \Omega + \Omega^2\eta\alpha = 0 \qquad (5)$$

This leads to four roots for λ and the general solution for u(y) is

$$u(y) = a\sin(\lambda_1 s) + b\cos(\lambda_1 s) + c\sinh(\lambda_2 s) + d\coh(\lambda_2 s)\,\Omega \qquad (6)$$
$$s = y/l$$

The boundary conditions, for a perfectly clamped beam, are

$$\left.\begin{aligned} & u(s) = 0 \\ & \Psi = \frac{kAG}{l(kAG - \omega^2 m r^2)} \times \\ & \left(\eta\frac{\partial^3 u(s)}{\partial s^3} + \left(1+\eta^2\Omega\right)\left(\frac{\partial u(s)}{\partial s}\right)\right) = 0 \end{aligned}\right\} \; s = 0$$
$$\left.\begin{aligned} & M = \frac{Y I_z}{l^2}\left(\frac{\partial^2 u(s)}{\partial s^2} + \eta\,\Omega\, u(s)\right) = 0 \\ & T = \frac{kAG}{l^3}\left(\frac{\partial^3 u(s)}{\partial s^3} + \Omega(\eta+\alpha)\frac{\partial u(s)}{\partial s}\right) = 0 \end{aligned}\right\} \; s = 1 \qquad (7)$$

Ψ, M and T are the rotation of a section, the bending moment and the shear force, respectively. The boundary conditions (7) and the frequency equation (5) allow to determine Ω as function of η and α. Once Ω determined, the resonant frequency is known by

$$\omega = \frac{b}{l^2}\sqrt{\frac{\Omega Y}{3\rho}} \qquad (8)$$

The limit case (Bernoulli's approximation) is obtained when α (and η) tends to zero. In order to compute the thermal sensitivity of ω, we have computed numerically the following expression for Ω, for the first mode of vibration

$$\Omega = 12.3624 - 57.454(\eta+\alpha) + 247.48(\eta^2+\alpha^2) + 806.3\eta\alpha - 906(\eta^3+\alpha^3) - 7160(\eta\alpha^2+\eta^2\alpha) \qquad (9)$$

To determine the thermal dependence of Ω it is necessary to know the thermal sensitivity of the shear coefficient k. Kawashima has given the following expression of k for a singly rotated bar which is independent of the cross section ratio [3]

$$k_k = \frac{10c'_{22}}{11c'_{22} + \sqrt{2\hat{c}_{44}(\hat{c}_{44} + c'_{66})}} \qquad (10)$$

Definitions of the coefficients are given in appendix I. The second expression of k used in this paper is based upon an energy definition. For a beam along the Y axis, the corresponding expression is:

$$k_1 = \frac{T_2^2}{A\int_{-h}^{+h}\int_{-b}^{+b}\left((T_6)^2 + (T_4)^2\right)dx_1 dx_3} \qquad (11)$$

The expressions of T_2, T_4 and T_6 are given in appendix II. In this case, k is function of the cross section ratio.

The variation of k_k with respect to the cut angle θ is plotted in Fig. 2. The two curves are obtained by using the numerical of elastic coefficients according to Bechmann [5] and Koga [6].

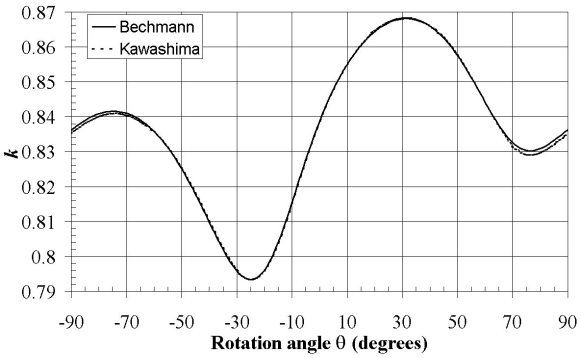
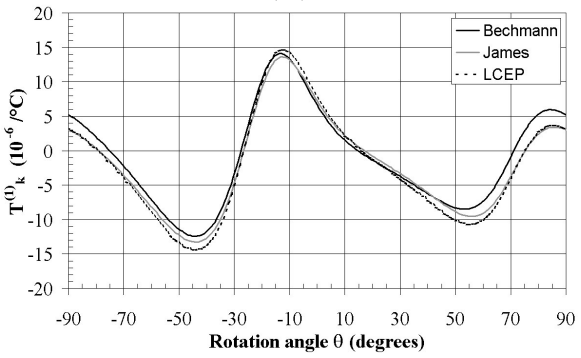


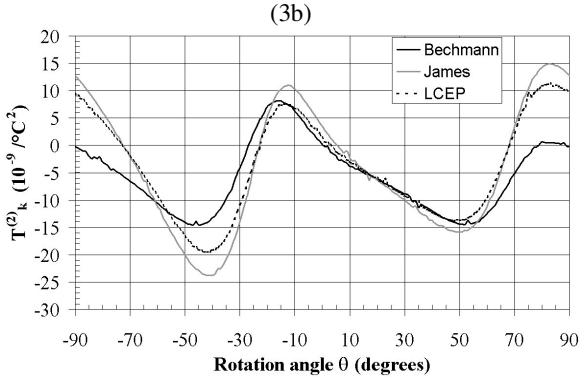
Fig. 2: Variation of Timoshenko's shear coefficient with respect to the cut angle θ.

The first, second and third order thermal coefficients (TCs) of k_k are plotted in Fig. 3. Three sets of data of thermal coefficients of elastic constants have been used: the classical data of Bechmann, Ballato and Lukasek [7], the values published by James [8], and the data published in [9]-10] referred as LCEP (See Appendix III). The thermal coefficients have been computed by a numerical method and the second and third order curves smoothed in order to reduce the noise of computation.

(3a)



(3b)



(3c)

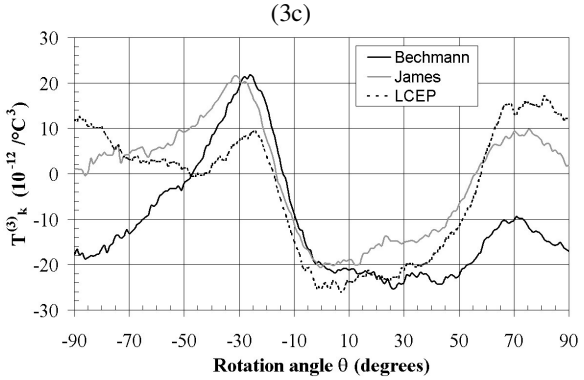


Fig. 3: First, second and third order TCs of k_k.

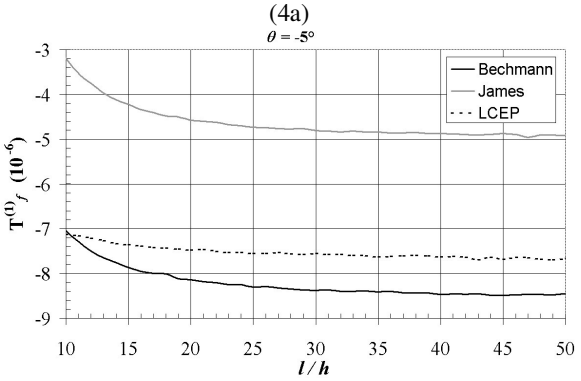
A comparison of thermal sensitivity of both definitions of k is given in Table 1 for square cross section and θ = 0°. The comparison is made for coefficients computed with the numerical values of ref. [7].

$T^{(1)}_{k\,k}$ (10^{-6}/°C)	6.6
$T^{(2)}_{k\,k}$ (10^{-9}/°C)	0.24
$T^{(3)}_{k\,k}$ (10^{-12}/°C)	-24.7
$T^{(1)}_{k\,l}$ (10^{-6}/°C)	5.3
$T^{(2)}_{k\,l}$ (10^{-9}/°C)	10
$T^{(3)}_{k\,l}$ (10^{-12}/°C)	-55.3

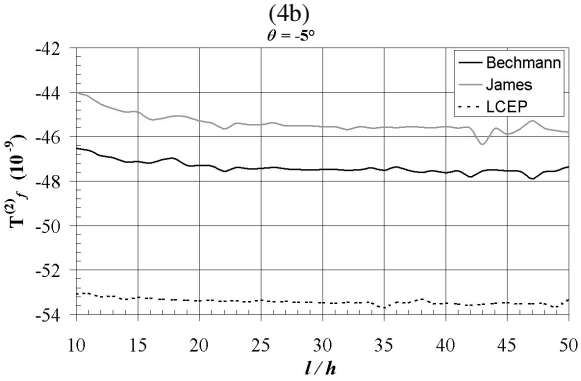
Table 1: Comparison of thermal sensitivity of both definitions of shear factor (square cross section, θ = 0°).

The first, second and third order temperature coefficients of the frequency of a beam as functions of the length/width ratios are plotted in Fig. 4 to 7 for various cut angle (θ = -5°, 0°, +5° and +10°). We may note significant differences between the values computed by using the three sets of thermal coefficients. For a length/width ratio greater than 50 the values are close to that obtained with the simplified model of Bernoulli.

(4a)



(4b)



(4c)

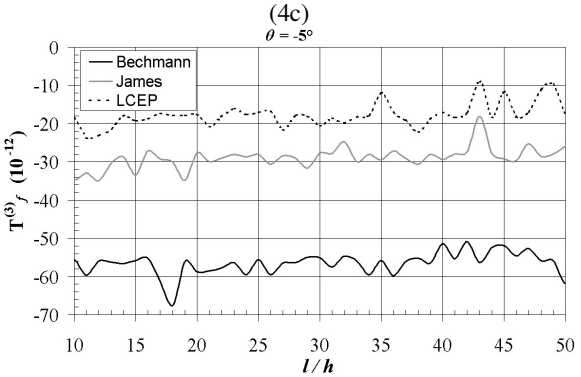
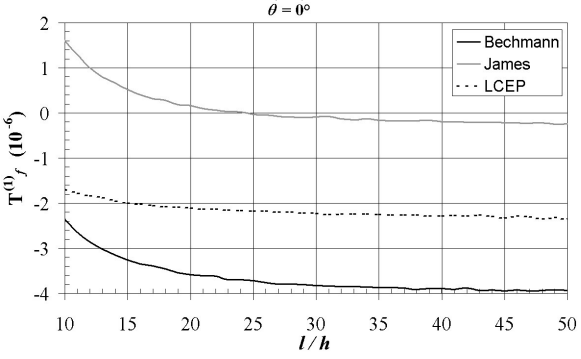
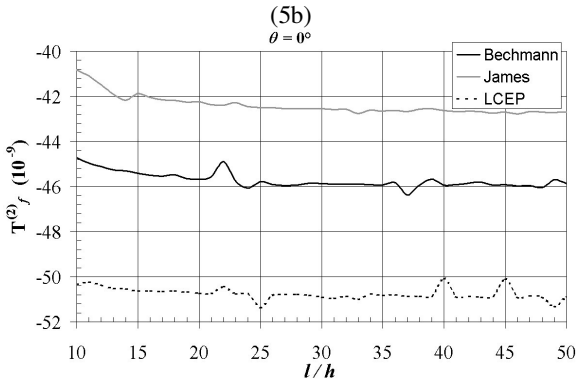


Fig. 4: First, second and third order TCs of frequency (θ = -5°).

(5a)





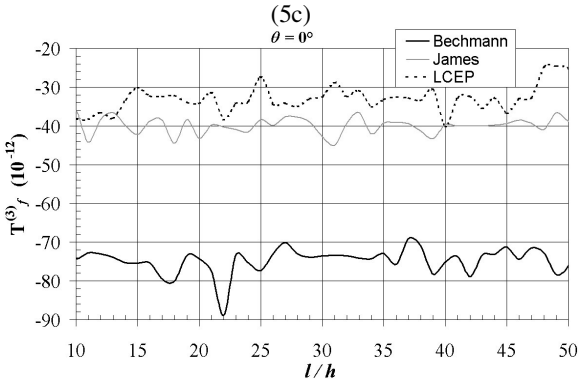
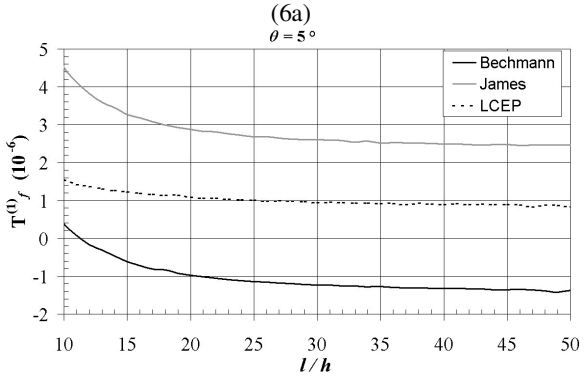
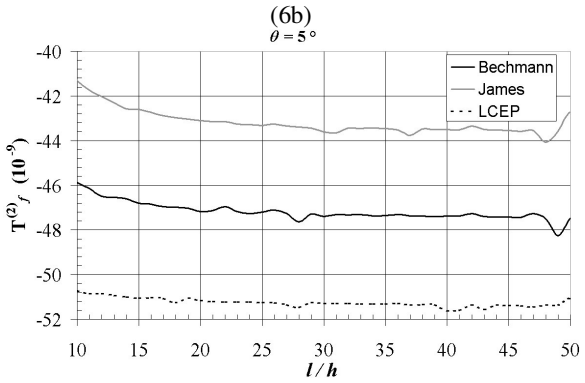


Fig. 5: First, second and third order TCs of frequency (θ = 0°).





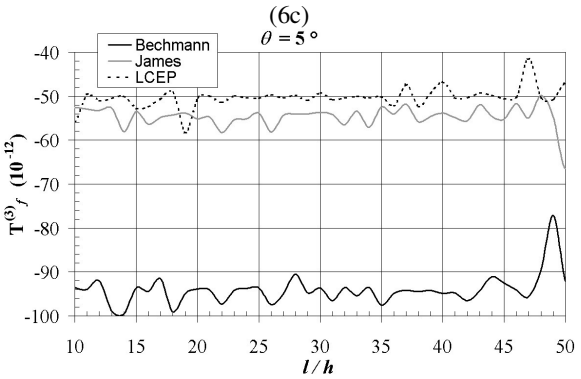
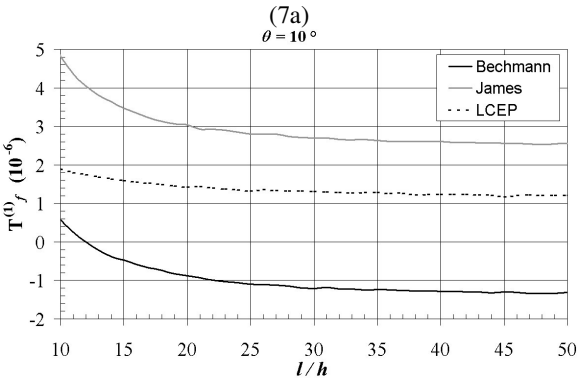
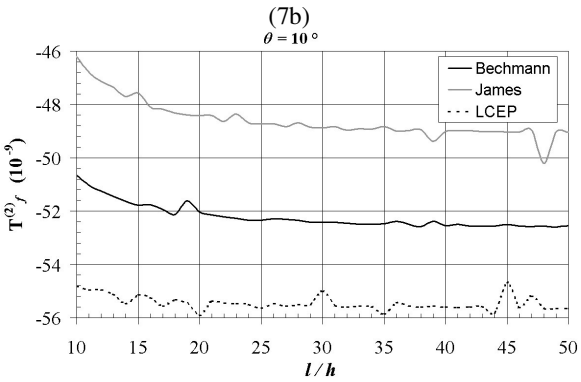


Fig. 6: First, second and third order TCs of frequency (θ = +5°).





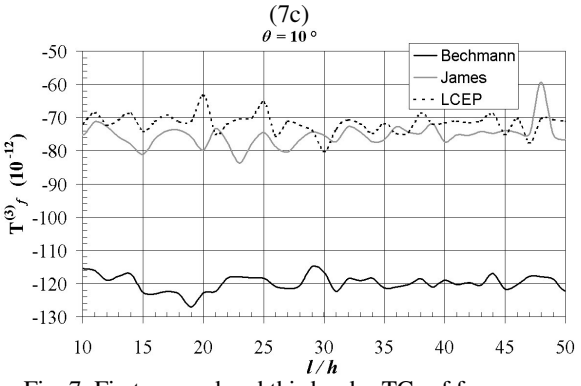


Fig. 7: First, second and third order TCs of frequency (θ = +10°).

The frequency versus temperature curve of a resonator for which the length/width ratio is equal to 12.5 is plotted in Fig. 8a, while the corresponding relative curve (normalized with respect to the frequency at 25 °C) is reported in Fig. 8b. For comparison, curves predicted by the Bernoulli model with the same set of data are also plotted. All curves are computed by using the shear factor k_k of Kawashima. We may notice that in this last case, the turnover point occurs at a lower temperature than for the Timoshenko's model. We may also notice that there is a significant difference between curves computed according to the several sets of thermal coefficients of elastic constants.

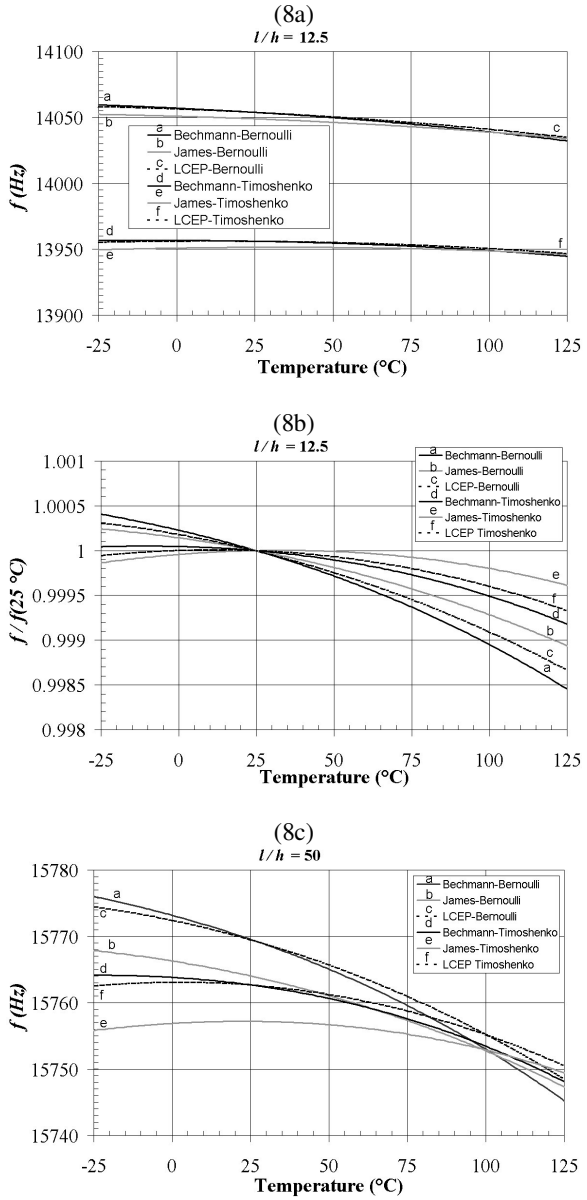


Fig. 8: Frequency vs. Temperature curves computed with Kawashima shear factor k_k ($\theta = 0°$).

The use of the second definition of the shear factor k_l (eq. 10) leads to a slightly different thermal behavior. With this second definition, the curves are not also depending of the length/width ratio, but are also functions of the cross section ratio (b/h). The temperature of the turnover point increase when the length/width ratio is lowered and the cross section ratio increased.

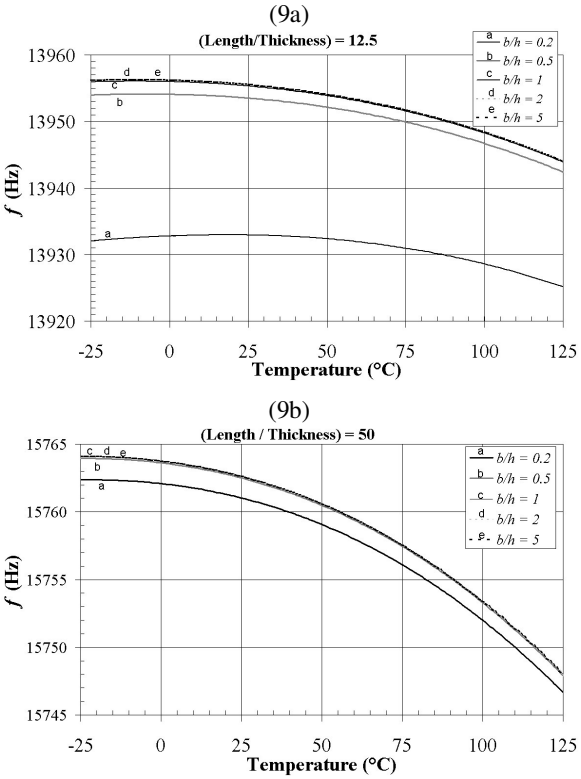


Fig. 9: Frequency vs. Temperature curves computed with shear factor k_l of eq. (10) ($\theta = 0°$).

3. CONCLUSION

For temperature compensated resonator the Timoshenko's model gives a better agreement with the experimental results. The turn over point is predicted by Bernoulli model at a lower temperature. Significant differences on the curvature of frequency vs. temperature curves are obtained from the different sets of material coefficients.

4. ACKNOWLEDGMENTS

This work is supported by the DGA.

APPENDIX I

DEFINITION of COEFFICIENTS APPEARING in (10)

According to [3]

$$
\begin{aligned}
&c'_{22} = \bar{c}_{22} - \bar{c}_{12}^2/\bar{c}_{11}; \quad \bar{c}_{22} = c_{22} - c_{23}^2/c_{33}\\
&c'_{44} = \bar{c}_{44} - \bar{c}_{14}^2/\bar{c}_{11}; \quad \bar{c}_{44} = c_{44} - c_{34}^2/c_{33}\\
&c'_{24} = \bar{c}_{24} - \bar{c}_{12}\bar{c}_{14}/\bar{c}_{11}; \quad \bar{c}_{11} = c_{11} - c_{13}^2/c_{33}\\
&\hat{c}_{44} = c'_{44} - c'^2_{24}/c'_{22}; \quad \bar{c}_{12} = c_{12} - c_{13}c_{23}/c_{33}\\
&c'_{66} = c_{66} - c_{56}^2/c_{55}; \quad \bar{c}_{14} = c_{14} - c_{13}c_{34}/c_{33}
\end{aligned}
\tag{1a}
$$

APPENDIX II

DERIVATION OF THE SHEAR COEFFICIENT

The stress appearing in (11) are computed when a static load P is applied at the free end of the beam (fig 10)

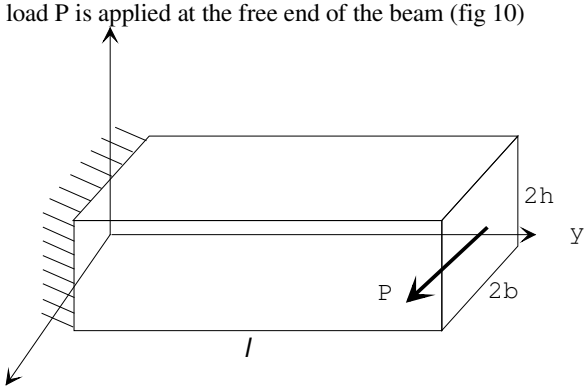


Figure 10

From a general point of view the stress must verify equation of equilibrium are

$$
\begin{aligned}
&T_{1,1} + T_{6,2} + T_{5,3} = 0\\
&T_{6,1} + T_{2,2} + T_{4,3} = 0\\
&T_{5,1} + T_{4,2} + T_{3,3} = 0
\end{aligned}
\tag{2a}
$$

in addition the relationships between strain components are

$$
\begin{aligned}
&S_{1,22} + S_{2,11} = S_{6,12} \quad 2S_{1,23} = S_{5,12} + S_{6,13} - S_{4,11}\\
&S_{2,33} + S_{3,22} = S_{4,23} \quad 2S_{2,13} = S_{4,12} + S_{6,23} - S_{5,22}\\
&S_{3,11} + S_{1,33} = S_{5,13} \quad 2S_{3,12} = S_{4,13} + S_{5,23} - S_{6,33}
\end{aligned}
\tag{2b}
$$

The strain-stress relations for a singly rotation about X axis are

$$
\begin{aligned}
&S_1 = s_{11}T_1 + s_{12}T_2 + s_{13}T_3 + s_{14}T_4\\
&S_2 = s_{12}T_1 + s_{22}T_2 + s_{23}T_3 - s_{14}T_4\\
&S_3 = s_{13}T_1 + s_{23}T_2 + s_{33}T_3\\
&S_4 = s_{44}T_4\\
&S_5 = s_{44}T_5 - 2s_{14}T_2\\
&S_6 = 2s_{14}T_5 + s_{66}T_6
\end{aligned}
\tag{2c}
$$

For quartz material we neglect the terms containing s_{14} in the above relations. Accordingly to the classical treatment of bending of bars, one can put

$$
\begin{aligned}
&T_2 = -\frac{P(l - x_2)\,x_1}{I_{x3}}\\
&T_1 = T_3 = 0
\end{aligned}
\tag{2d}
$$

Putting (2c) into (2b) with the help of derivatives of (2a), one can check that the equations (2c) are verified, excepted the fourth and the sixth which give the following relations

$$
\begin{aligned}
&s_{44}T_{4,11} + s_{55}T_{4,22} + s_{66}T_{4,33} = -(s_{66} + 2s_{12})T_{2,23}\\
&s_{44}T_{6,11} + s_{55}T_{6,22} + s_{66}T_{6,33} = 0
\end{aligned}
\tag{2e}
$$

These relations can be rewritten

$$
\begin{aligned}
&\frac{s_{44}}{2s_{12} + s_{66}}\left(T_{4,11} + T_{4,22} + \frac{s_{66}}{s_{44}}T_{4,33}\right) + T_{2,23} = 0\\
&s_{44}\left(T_{6,11} + T_{6,22} + \frac{s_{66}}{s_{44}}T_{6,33}\right) = 0
\end{aligned}
\tag{2f}
$$

The equilibrium equations (2a) reduce to

$$
\begin{aligned}
&T_{6,2} = 0 \qquad T_{4,2} = 0\\
&T_{6,1} + T_{4,3} = -\frac{P}{I_{x3}}x_1
\end{aligned}
\tag{2g}
$$

The integration of these equations gives

$$
\begin{aligned}
&T_6 = \Phi_{,3} - \frac{P x_1^2}{2I_{x3}} + f(x_3)\\
&T_4 = -\Phi_{,1}
\end{aligned}
\tag{2h}
$$

Introducing

$$
\alpha^2 = \frac{s_{66}}{s_{44}} \qquad k = \frac{2s_{23} + s_{44}}{s_{44}}
\tag{2i}
$$

and putting (2h) into (2g), one obtain

$$
\begin{aligned}
&\left(\Phi_{,11} + \alpha^2\Phi_{,33}\right)_{,3} + (k-1)\frac{P}{I_{x3}} + f_{,33} = 0\\
&\left(\Phi_{,11} + \alpha^2\Phi_{,33}\right)_{,1} = 0
\end{aligned}
\tag{2j}
$$

Integrating (2j-1)

$$
\Phi_{,11} + \alpha^2\Phi_{,33} = (1-k)\frac{P}{I_{x3}}x_3 - f_{,3}
\tag{2k}
$$

The boundary conditions are

$$
\begin{aligned}
&T_6 = 0 \qquad x_1 = \pm b\\
&T_4 = 0 \qquad x_3 = \pm h
\end{aligned}
\tag{2l}
$$

In view of (2h)

$$
f(x_3) = \frac{P b^2}{2I_{x3}}
\tag{2m}
$$

Making the substitution

$$
x_3 = \frac{z}{\alpha} \qquad x_1 = x
$$

one find that Φ is solution of

$$\frac{\partial^2\Phi}{\partial x^2}+\frac{\partial^2\Phi}{\partial z^2}=\frac{1-k}{\alpha}\frac{P}{I_{x3}}z$$
$$\Phi(x,z)=0 \quad \begin{vmatrix} x=\pm b \\ z=\pm\alpha h \end{vmatrix} \qquad (2n)$$

The expression for Φ is given in classical text book [1]. Thus

$$\Phi=\frac{\nu}{1+\nu}\frac{P}{I_{x3}}\frac{8\alpha^3h^3}{\pi^4}\times$$
$$\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}\frac{(-1)^{m+n}\cos\frac{(2m+1)\pi x_1}{2b}\sin\frac{n\pi x_3}{h}}{(2m+1)\,n\left[(2m+1)^2\frac{\alpha^2h^2}{4b^2}+n^2\right]} \qquad (2o)$$

where

$$\nu=\frac{1-k}{\alpha+k-1}=-2\frac{s_{23}}{\sqrt{s_{44}s_{66}}+2s_{23}} \qquad (2p)$$

Finally, the stress T_4 and T_6 are

$$T_4=\frac{\nu}{1+\nu}\frac{4}{\pi^3}\frac{\alpha^3h^3}{b}\frac{P}{I_{x3}}\times$$
$$\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}\frac{(-1)^{m+n}\sin\frac{(2m+1)\pi x_1}{2b}\sin\frac{n\pi x_3}{h}}{n\left[(2m+1)^2\frac{\alpha^2h^2}{4b^2}+n^2\right]}$$
$$T_6=\frac{P\left(b^2-x_1^2\right)}{2I_{x3}}+$$
$$\frac{\nu}{1+\nu}\frac{8\alpha^3h^2}{\pi^3}\frac{P}{I_{x3}}\times \qquad (2q)$$
$$\sum_{m=0}^{\infty}\sum_{n=1}^{\infty}\frac{(-1)^{m+n}\cos\frac{(2m+1)\pi x_1}{2b}\cos\frac{n\pi x_3}{h}}{(2m+1)\left[(2m+1)^2\frac{\alpha^2h^2}{4b^2}+n^2\right]}$$

APPENDIX III

THERMAL COEFFICIENTS FROM REF. [9] and [10]

Thermal coefficients listed below are computed from those in Lagrangian description in [9] and [10] by using the connecting formulas. First order coefficients $T^{(1)}$ are given in 10^{-6}/°C, second order $T^{(2)}$ in 10^{-9}/°C and third order $T^{(3)}$ in 10^{-12}/°C.

	c_{11}^E	c_{12}^E	c_{13}^E	c_{14}^E	c_{33}^E	c_{44}^E
$T^{(1)}$	-41.2	-2488	-693.6	88.2	-196.5	-179.1
$T^{(2)}$	-120.7	-2884	-1044	-52.6	-160.6	-230.6
$T^{(3)}$	-91.9	-1734	-40.8	-525	167	-162

	s_{11}^E	s_{12}^E	s_{13}^E	s_{14}^E	s_{33}^E	s_{44}^E
$T^{(1)}$	12.2	-1069	-322	154	167.8	212.8
$T^{(2)}$	93.4	-1245	-773	134	148	276
$T^{(3)}$	56	-1725	-475	-453	-101	126

4. REFERENCES

[1] S.P. Timoshenko and J.N. Goodier, Theory of Elasticity, 3rd edition, McGraw-Hill, 1970.

[2] S. P. Timoshenko, "On the Correction for Shear of the Differential Equation for Transverse Vibrations of Prismatic Bars," Phil. Mag. Ser. 6 [41], pp. 774-746 (1921).

[3] H. Kawashima, The shear coefficient for quartz crystal of rectangular cross section in Timoshenko's beam theory, IEEE Trans. Ultrason. Ferroelec. and Freq. Cont., Vol. 43, n°3, pp. 434-440, May 1996.

[4] G.R. Cowper, The shear coefficient in Timoshenko's beam theory, J. Appl. Mechanics, Vol. 33, pp. 335-340, 1966.

[5] R. Bechmann, Elastic and piezoelectric constants of alpha quartz, Phys. Rev. Vol. 110, pp.106-1061, June 1958.

[6] I. Koga, Maruga, Y. Yoshinaka, Theory of plane elastic waves in a piezoelectric crystalline medium and determination of the elastic and piezoelectric constants of quartz, Phys. Rev. Vol. 109, pp. 1467-1473, 1958.

[7] R. Bechmann, A.D. Ballato, T.J. Lukasek, Higher-order temperature coefficients of the elastic stiffnesses and compliances of alpha-quartz, Proc. of the I.R.E., pp. 1812-1822, August 1962.

[8] B. James, Determination of complete fully consistent data set of the fundamental properties of high purity synthetic quartz, Proc. 5th EFTF, pp. 221-223, Besançon, 1991.

[9] B. Dulmet, R. Bourquin, E. Bigler, S. Ballandras, Lagrangian effective material constants for the modelling of thermal behavior of acoustic waves in pizoelectric crystals. Part II. Applications and numerical values for quartz. J. Acoust. Soc. Am., Vol. 110, n°4, pp. 1800-1807, Oct. 2001.

[10] R. Bourquin, B. Dulmet, Thermal derivatives of the compliances, piezoelectric coefficients and permittivities in Lagrangian description, Proc. of 2003 IFCS and 17th EFTF, pp. 723-729, Tampa, May 2003.

A PERTURBATION METHOD FOR PREDICTING THE MASS LOADING AND TEMPERATURE SENSITIVITIES OF QUARTZ VIBRATING STRUCTURES SIMULATED BY FINITE ELEMENT ANALYSIS

S. Ballandras
Institut FEMTO-ST, UMR CNRS UFC ENSMM UTBM 6174, Département LPMO, 32 Avenue de l'Observatoire, F-25044 Besançon Cedex

Abstract :

The optimisation of parametric sensitivity of vibrating acousto-electric devices still represent a critical point for the fabrication of robust devices. A variational perturbation method is presented in this paper to compute the thermal and gravimetric sensitivity of any crystal structure, taking advantage of finite element analysis for simulating its initial dynamic behaviour. The theoretical approach is detailed and illustrated for both 2D and 3D Quartz devices.

I. Introduction

Theoretical methods developed for the prediction of the temperature sensitivity of vibrating structures cannbe easily counted. The most popular, simple and often implemented approach is conform to the celebrated Campbell & Jones approach for surface acoustic waves, consisting in using effective temperature constants to represent the evolution of a material properties versus thermal variation [1]. This enables one to compute the actual material coefficients (elastic, piezoelectric, dielectric, mass density) at a given temperature and then solve the equilibrium equations, giving access to the resonance frequency or the phase velocity. One then can extract temperature coefficients of frequency (TCF) by solving the problem for different temperature and by a second or third order polynomials fitting procedure. This approach is quite efficient for devices exhibiting strong temperature sensitivities but is revealed not accurate enough to predict temperature effect compensation phenomena comparable to those arising in quartz [2]. Dulmet and Bourquin [3] have proposed a revised version of this model for which the intrinsic limitation of this method (mainly related to the tensorial behaviour of these effective constants) are corrected. However, it requires the capability to solve the dynamics equation with reduced symmetry properties, yielding for instance the dissociation of the dynamical displacement gradients ($\partial u_i/\partial x_j \neq \partial u_j/\partial x_i$). This kind of computation is not supported by standard computation tools nor finite element codes.

Another way to address the problem relies on the perturbation method proposed by Sinha and Tiersten [4]. This approach is based on the use of a variationnal formulation of the dynamics problem, combined with a rigorous development of thermodynamics properties of the considered material, introducing the first temperature derivatives of the fundamental elastic constants [5]. A pertubation tensor then can be computed for the whole structure or even locally and mixed with dynamics equations within the variationnal formulation. By properly combining variationnal equations of the perturbed and non perturbed configurations of the vibrating device, one can extract a frequency shift proportional to the change of physical properties of the device by only computing its non perturbed vibration distribution. The main limitation of this approach is related to the accessibility of the material constants required for the calculation of the perturbation tensor. Consequently, this method is limited to the prediction of first order temperature effects on quartz devices for which Sinha and Tiersten have provided an accurate set of data [5]. One can also consider some isotropic materials for which the first temperature derivatives of the elastic constants have been computed [6]. This is particularly attractive when computing thermal sensitivities of devices taking into account their metallization pattern [7].

In this paper, we propose to use the variationnal formulation of finite element analysis (FEA) which appears very close to the one used in the Sinha-Tiersten perturbation method for the computation of the thermal sensitivity of any quartz vibrating structure, and the computation of stress sensitivity coefficients as well as gravimetric effects due to electrode loading for instance. The idea is that the integration of the dynamical terms of the perturbation method is already performed in the finite element analysis. The perturbation computation then simply reduces in replacing the elastic constants by the perturbation tensor in the variationnal FEA formulation. The frequency shift then is proportional to the ratio between the perturbed stiffness matrix and the non perturbed mass matrix. Consequently, once the dynamic displacement distribution of any vibrating structure has been computed, it is possible to use the intermediate FEA computation results to calculate the sensitivity coefficients.

The first section of the paper describes the fundamentals of the proposed approach. The second part of the paper shows some validation results for different structures such as bulk acoustic wave (BAW) resonators or tuning fork devices. Finally, further developments of the proposed approach are discussed

2. Description of the proposed approach

2.1. Fundamentals

In the proposed approach, only the perturbation of the elastic constants are considered. For instance, we do not take into account the contribution of piezoelectricity nor dielectric constant variations due to thermo-elastic effects [8, 9, 10, 11]. The 3-D variationnal formulation of a purely elastic problem is written in Lagrangian coordinates as follows :

$$\iiint_{\Omega}(-\rho_0\omega_0^2 u_i^0\delta u_i^* + \frac{\partial \delta u_i^*}{\partial a_j}C_{ijkl}\frac{\partial u_l^0}{\partial a_k})dV = \iiint_{\Omega}\delta u_i^* F_i^0 dV + \iint_{\Gamma}\delta u_i^* T_{ij}^0 n_j dS \quad (1)$$

The upperscript 0 denotes unperturbed fields (displacement u_i, surface stress $T_{ij}.n_j$, volume forces F_i) and values (angular frequency ω, mass density ρ), i.e. without any static deformation. This formulation can be used even for complex fields (* holds for complex conjugation). C_{ijkl} are the elastic constants and δu_i is the variationnal unknown which can take any value, a priori non dependent of the non perturbed solution (it will coincide with the solution in the Galerkin approach [12]). The same formulation can be written in the same system of axes, but for a wave or vibration taking place in a medium submitted to a parametric perturbation of small magnitude, for which a local linear approximation of the non linear properties can be used without loss of accuracy. This reads :

$$\iiint_{\Omega}(-\rho_0\omega^2 u_i\delta u_i^{0*} + \frac{\partial \delta u_i^{0*}}{\partial a_j}\bar{A}_{ijkl}\frac{\partial u_l}{\partial a_k})dV = \iiint_{\Omega}\delta u_i^{0*} F_i dV + \iint_{\Gamma}\delta u_i^{0*} T_{ij} n_j dS \quad (2)$$

Since the mass conservation has to be respected, one could expect the mass density to differ in eq.(2) from eq.(1). Actually, we consider that the ratio between mass density in perturbed and non perturbed states corresponds to the Jacobian of the geometrical transformation from one system of axes to another [13], and that this ratio can be approximated to 1 in the case of small amplitude fields superposed on a bias [4,13]. The variational unknown and the solution are then confounded (Galerkin) in both equations. We retrieve eq.(1) to (2), and then we express the difference between the angular frequencies in perturbed and non perturbed states [4,7,8,9,10,11]. In that matter, we assume that $\omega+\omega_0=2\omega_0$, which is acceptable because this term only appears as a weighting coefficient and thus only introduces a second order bias (considered negligible). We finally obtain the so-called Sinha-Tiersten perturbation equation [4] :

$$\frac{\omega-\omega_0}{\omega_0}=\frac{\Delta\omega}{\omega_0}=\frac{\iiint_{\Omega}(\frac{\partial u_i^{0*}}{\partial a_j}\bar{H}_{ijkl}\frac{\partial u_l^0}{\partial a_k})dV}{2\rho_0\omega_0^2\iiint_{\Omega}u_m^{0*}u_m^0 dV} \quad (3)$$

In order to compute eq. (3), the vibration distribution in the non-perturbed regime must be known as well as the steady pre-deformation. Equation (3) was established assuming stress-free border conditions and the absence of internal forces F_i. The perturbation tensor resulting from the difference between the elasticity tensor with and without pre-deformation also can depend on the space coordinates. In the general thermo-elastic problem, this tensor reads :

$$\bar{H}_{ijkl}=\delta_{ik}\bar{T}_{jl}+C_{ijkluv}\bar{S}_{uv}+C_{pjkl}\frac{\partial \bar{u}_i}{\partial a_p}+C_{ijql}\frac{\partial \bar{u}_k}{\partial a_q}+\frac{dC_{ijkl}}{d\theta}(\theta-\theta_0) \quad (4)$$

where C_{ijkluv} are the 3rd order (non linear) elastic constants, $\bar{u}_i$, $\bar{S}_{ij}$ and $\bar{T}_{ij}$ are the space-dependent quasi-static displacements, linear deformations and stresses respectively. These terms also can depend on the temperature θ, which then must be known at each point of the medium. The dependence of the fundamental elastic constants versus temperature must be taken into account carefully, according to Tiersten et Sinha's developments [5] via its 1st order derivatives versus temperature. The expression (4) of the perturbation tensor must be consider for instance in the case where differential thermal stresses arise (for instance a compound substrate with at least two material exhibiting different thermal expansion coefficient). Simpler expressions of this tensor can be derived when considering structures composed of only one material and submitted to slowly varying perturbation phenomena, as shown in the next section.

2.2. Structures with infinitely thin electrodes.

We assume first an homogeneous temperature distribution within an isolated structure only composed of quartz, the electrodes are assumed infinitely thin (we neglect the mass loading). In that case, we can write for the whole device :

$$T_{ij}=0 \Rightarrow S_{ij}=\alpha_{ij}(\theta-\theta_0) \quad (5)$$

We add one more non-rotation condition which allows to substantially enhance the symmetry of the operator, then yielding :

$$\frac{\partial u_i}{\partial a_j}=S_{ij} \quad (6)$$

consequently, (4) can be written as follows :

$$\bar{H}_{ijkl}=(C_{ijkluv}\alpha_{uv}+C_{pjkl}\alpha_{pi}+C_{ijql}\alpha_{kq}+\frac{dC_{ijkl}}{d\theta})(\theta-\theta_0) \quad (7)$$

The perturbation tensor then does not depend on space variables any more, and (3) can be split in "static" and "dynamic" parts as :

$$\frac{\Delta\omega}{\omega_0}=U_{ijkl}\,{}^{\theta}H_{ijkl}(\theta-\theta_0) \quad (8)$$

with :

$${}^{\theta}H_{ijkl}=C_{ijkluv}\alpha_{uv}+C_{pjkl}\alpha_{pi}+C_{ijql}\alpha_{kq}+\frac{dC_{ijkl}}{d\theta} \quad (9)$$

and

$$U_{ijkl}=\frac{\iiint_{\Omega}(\frac{\partial u_i^{0*}}{\partial a_j}\frac{\partial u_l^0}{\partial a_k})dV}{2\rho_0\omega_0^2\iiint_{\Omega}u_m^{0*}u_m^0 dV} \quad (10)$$

Practically, one easily can estimate the values of ${}^{\theta}H_{ijkl}$ by a simple tensorial rotation (for a given crystal cut, one

rotates all the constants of (9)). However, a dynamic model is required to access the displacement field to correctly compute (10). We further show how this perturbation method is combined with FEA.

2.3. Influence of the electrodes

The previously obtained perturbation method is valid for structures with infinitely thin electrodes. In the real case, massive electrodes are deposited on the structure to excite and detect high frequency vibrations or electrical charges induced by a given perturbing effect. As shown in [7], one also can treat the influence of the electrode using a perturbation method. In that case, (3) is split in two terms termes respectively representative of the volume and the device surface. We then express the continuity of the stresses and displacements at the crystal/metal. The corresponding term will be null for non metallized surfaces. At room temperature, only the surface term persists :

$$\frac{\Delta\omega}{\omega_0}=\frac{\iint_\Gamma u_i^{0*}T_{ij}^0 n_j dS}{2\rho_0\omega_0^2\iiint_\Omega u_m^{0*}u_m^0 dV} \qquad (11)$$

the dynamic stresses of (11) can be easily accessed using FEA, providing a relation between stresses and displacements at the quartz/metal interface at ω_0 :

$$\{T_{ij}.n_j\}=[Z_M(\omega_0)]\{u_i\} \qquad (12)$$

It is of course possible to directly solve the whole problem including the electrodes but it still represents a costly operation (particularly in 3D) since the electrode are generally thin compared to the author dimensions of the vibrating device, obliging to densely mesh the metal media to preserve accuracy. Predicting the evolution of the electromechanical coupling also would consists in an exciting issue. In that approach, one has to first compute the resonance f_r then the anti-resonance f_a and then determine the frequency variations due to the electrodes for a given configuration. The actual coupling strength then is evaluated using the usual definition of $K_s^2= 2\times(f_a-f_r)/f_a$. This would significantly help in reducing the computation duration and allow for characterizing the computation stability.

3. Perturbation method coupled to FEA

We now describe the coupling between the exposed perturbation method and the FEA. We take advantage of the coincidence between (1) and the variationnal formulation implemented in the FEA (also called the Lagrangian) which describes the energy equilibrium of an isolated body. The main fundament of the finite element method consists in digitalizing the integration domain in such a way that continuous integrals in (1) are decomposed in sums of elementary integrals corresponding to the different elements of the digitalized domain. It then reads :

$$\sum_{e=1}^{E}\iiint_{\Omega_e}(-\rho_e\omega_0\delta u_i^{*e}u_i^e+\frac{\partial\delta u_i^{*e}}{\partial a_j}C_{ijkl}^e\frac{\partial u_l^e}{\partial a_k})dV= \sum_{e=1}^{E}\left(\iiint_{\Omega_e}\delta u_i^{*e}F_i^e dV+\iint_{\Gamma_e}\delta u_i^{*e}T_{ij}^e n_j dS\right) \qquad (13)$$

where the index e signifies that terms are related to the eth element of the whole domain composed of E elements. We then assume that the solution for each element can be written as a polynomial interpolation built on the N nodal values $u_i^{(n)}$ of the displacements (the independent variables) :

$$\begin{Bmatrix}u_i\\ \delta u_i\end{Bmatrix}^e=\sum_{n=1}^{N}P_n(x,y,z)\begin{Bmatrix}u_i^{(n)}\\ \delta u_i^{(n)}\end{Bmatrix}^e \qquad (14)$$

Note that if the displacement distribution corresponds to an eigen-mode of the structure, the $u_i^{(n)}$ are real.

We remark that similar forms appear in the perturbation equation (10) and in the FEA formulation (12), particularly the mass matrix and some terms of the stiffness matrix. Rewriting (10) accounting for the FEA digitalisation yields :

$$\frac{\Delta\omega}{\omega_0}=\frac{\sum_{e=1}^{E}\iiint_{\Omega_e}(\frac{\partial u_i^{0*(e)}}{\partial a_j}\overline{H}_{ijkl}^e\frac{\partial u_l^{0(e)}}{\partial a_k})dV}{2\rho_0\omega_0^2\sum_{e=1}^{E}\iiint_{\Omega_e}u_m^{0*(e)}u_m^{0(e)}dV} \qquad (15)$$

The denominator of (15) actually corresponds to the mass matrix in non-perturbed regime, whereas the numerator requires the dynamic displacement gradients computed for the stiffness matrix and for which the elastic constants have been replaced by the perturbation tensor. One must be able to re-use the corresponding elementary matrices defined in (13) to implement the perturbation method together with the FEA. However, it must be underlined that the perturbation tensor does not exhibit the symmetry properties of the elastic constant. Consequently, the denominator must be computed taking accounting for that specific point. Finally, one can write (14) under the following matrix form :

$$\frac{\Delta\omega}{\omega_0}=\frac{\sum_{e=1}^{E}\langle u_i^{0*}\rangle^e[\overline{K}_{ijkl}]^e\{u_l^0\}^e}{2\omega_0^2\sum_{e=1}^{E}\langle u_m^{0*}\rangle^e[M]^e\{u_m^0\}^e}=\frac{\langle u_i^{0*}\rangle[\overline{K}_{ijkl}]\{u_l^0\}}{2\omega_0^2\langle u_m^{0*}\rangle[M]\{u_m^0\}} \qquad (15)$$

where $[\overline{K}]$ represents the perturbed stiffness and $[M]$ the unperturbed assembled mass matrix (the one of the FEA). It appears that solving (15) only requires the nodal unperturbed values of the displacement field, the assembled mass matrix and the dynamic displacement gradients to compute the regarded frequency variation. this formulation can be implemented to simulate temperature effects but also the stress sensitivity coefficients and more generally any parametric quasi-static perturbation modelled using a linear model.

4. Numerical applications

The first example presented in this section concerns the well-established problem of BAW resonators on singly rotated quartz cut. Our model is expected to predict the TCF of fundamental pure shear waves piezoelectrically excited between two flat surfaces (see fig.1).

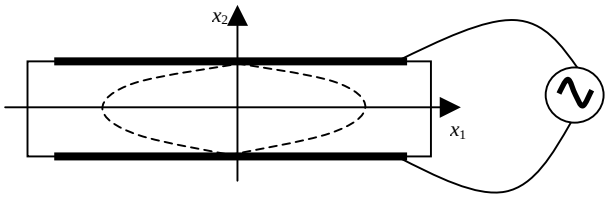


Fig.1 Geometry of the BAW resonator considered for the model validation

A very simple mesh was implemented with sufficient number of elements along the thickness to properly predict the resonance frequency (at the present time, the perturbation model is implemented using 1st degree polynomials). The resonance was fit for each angle configuration using an approximate form of the electric admittance written $Y_p/(\omega-\omega_0)$ via a 3-point method [14]. The perturbation is then applied at resonance, yielding the following curves (fig.2).

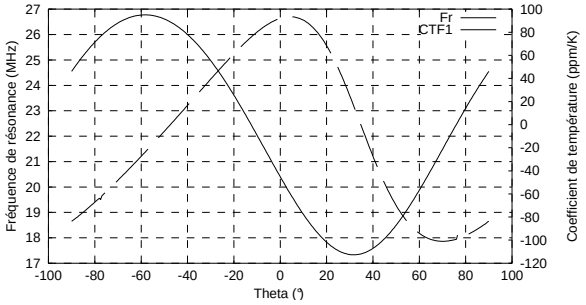


Fig.2 Resonance frequency and TCF fur shear BAW excited in a 100µm thick singly rotated quartz plate

It is seen from fig.2 that of course the resonance is well predicted by FEA but also the TCF is nicely found in agreement with literature (see refs). This simple helped in setting the computation procedure, and gave confidence in the proposed approach. However, for the sake of demonstrating the interest of the proposed approach, it is not the better choice. We then moved toward the simulation of 3D structures. In that matter, we have chosen to compute the TCF of quartz tuning fork, which actually represents a complex three dimensional structure. We of course have taken into account the electrode pattern and then we did compute the resonance frequency as well as the TCF. Figure 3 shows the implemented mesh. It still consists in 1st degree interpolation elements, then yielding sensible deviation between theory and experiments. However, it is shown in fig.4 that the loci of 1st order TCF cancelling are well predicted by the model. We actually find a zero TCF close to the Z cut, which is one of the most used crystal orientation for such structure.

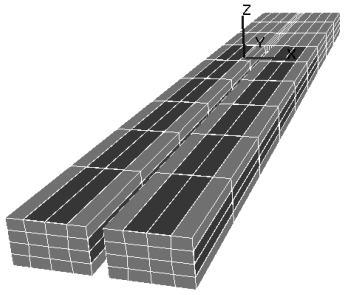
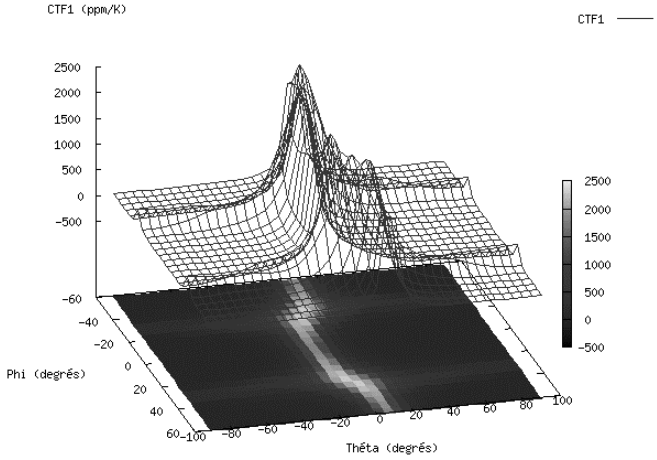
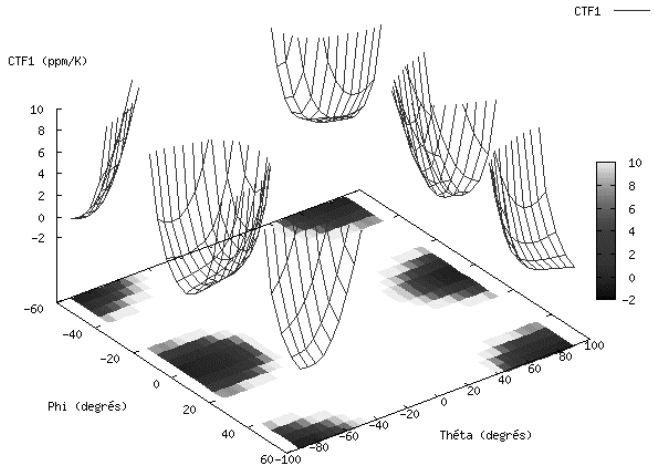


Fig.3 FEA mesh of a Z-cut quartz tuning fork resonator



(a)



(b)

Fig.4 TCF of the quartz tuning fork resonator simulated using the mesh of fig.3
(a) TCF for any quartz cut (b) focus on the zero TCF loci

Finally, we want to illustrate the possibility to simulate mass loading effects using the proposed model. We turn back to the classical BAW resonator, for which we want to demonstrate that massive electrode effects can be predicted by only computing the unperturbed resonance of the resonator and then perturbing it using the acoustic contribution of the electrode, according to the above-described procedure. We then consider meshes of fig.5 that have been magnify for the sake of readability.

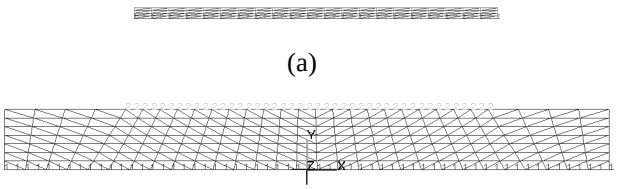


Fig.5 FEA meshes used to validate the mass-loading computation via the perturbation method
(a) gold electrode (b) quartz resonator (AT cut)

Figure6 shows the result of the procedure. One of course can remark that the resonance frequency decreases with the mass loading (the electrode thickness was varied from 0.1 to 10 µm). One would have expect a quadratic behaviour, and then more attention has to be paid to this aspect of the model before announcing any validation. The only conclusion about that particular work is that the computation procedure can be operated without any unwanted numerical defect.

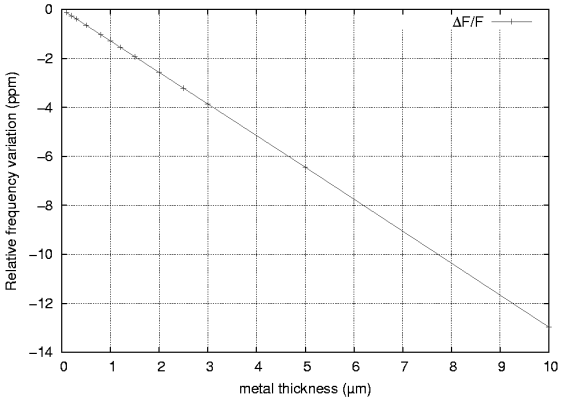


Fig.6 Mass loading effects due to gold electrodes atop AT quartz BAW resonator

5. Conclusion

In this paper, the way FEA can be coupled with a variational perturbation method has been presented in detail. This perturbation method has been implemented to predict 1st order TCF of various 2D and 3D structures. The capability of the model to predict mass loading effects also has been tested, but more data are needed to definitely consider the later approach as fully validated. However, one can note that such an approach has provided satisfying results for finding out the zero TCF loci even for complicated 3D structures. The model is actually limited to quartz for the simulation of temperature effects, but not for mass loading. One also should note that stress sensitivity coefficient can be derived in a similar way. Since 3rd order elastic constants are known for many materials (quartz, $LiNbO_3$, Silicon, Sapphire, etc.), the approach sounds attractive for the optimisation of vibrating mechanical sensors.

Références :

[1] J.J. Campbell, W.R. Jones, “A method for estimating crystals cuts and propagation direction for excitation of piezoelectric surface waves”, IEEE Trans. On Sonics and Ultrasonics, Vol. 15, pp. 209-217, 1968

[2] 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 .

[3] B. Dulmet, R. Bourquin, “Lagrangian effective material constants for the modeling of thermal behavior of acoustic waves in piezoelectric crystals”, J. of Acoust. Soc. Of Am., Vol. 110, n°4, pp. 1790-1799, 2001

[4] Sinha, B.K., Tiersten, H.F., “On the temperature dependence of the velocity of surface wave in quartz”, J. Appl. Phys., Vol.51, n°9, pp. 4659-4665, 1980

[5] B.K. Sinha, H. F. Tiersten, “First derivatives of the fundamental elastic constants of quartz”, J. Appl. Phys., Vol.50, n°4, pp. 2732-2739, 1979

[6] S. Ballandras, E. Gavignet, E. Bigler, E. Henry-Briot, "Temperature derivatives of the fundamental elastic constants of isotropic materials", Applied Physics Letters, Vol. 71, n°12, pp. 1625-1627, 1997

[7] E. Henry-Briot, S. Ballandras, G. Marianneau, G. Martin, "Influence of metallization on temperature stability of SAW devices", IEEE Trans. on UFFC, Vol. 48 ,n°3, pp.538-546, 2000

[8] B.K. Sinha, “A stress and temperature compensated orientation and propagation directionfor surface acoustic wave devices”, IEEE Trans. on UFFC, Vol. 34, n°1, 1987

[9] 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, 1992

[10] R.M. Taziev, “Extreme pressure-sensitive cuts for surface acoustic waves in α-quartz”, Proc. of the Int. Freq. Cont. Symp., pp. 345-350, 1994

[11] J.T. Stewart, D.P. Chen, “Finite element modelling of the effects of mounting stresses on the frequency temperature behaviour of surface acoustic wave devices”, Proc. of the Ultrasonics Symp., pp.105-111, 1997

[12] O.C. Zienkiewickz, “The finite element method in structural and continuum mechanics”, Mc Graw Hill Ed., 1979

[13] R.N. “Thurston, Wave propagation in fluids and solids”, in *Physical Acoustics*, Vol.1, W.P. Mason Ed., Academic Press, 1964

[14] Y. Fusero, S. Ballandras, J. Desbois, J.M. Hodé, P. Ventura, "SSBW-to-PSAW conversion in SAW devices using heavy mechanical loading", IEEE Trans. on UFFC, Vol. 49, n°6, pp. 805-814, 2002

THE INFLUENCE OF THE MATERIAL CONSTANTS ON LANGASITE DEVICES

Irina Mateescu[1], Cristina Bran[1], Serban Georgescu[2]

[1]National Institute of Materials Physics, PO Box MG 7, 77125, Bucharest - Magurele, Romania

[2]National Institute for Lasers, Plasma and Radiation Physics, PO Box MG-36, 77125, Bucharest - Magurele, Romania

e-mail: imate@infim.ro

Abstract

Based on Tiersten's analysis [1] an accurate calculation of the coupling coefficient dependence on geometrical dimensions and mass-loading effect for quartz and langasite monolithic structures has been performed [2].
In this paper we have studied the influence of the material constants on the attenuation characteristics of two-pole langasite monolithic filters by calculation of the dependence between the coupling coefficient and elastic constant c_{66}, piezoelectric constant e_{26} and Voigt's elastic constant γ_{11}. The relation between the coupling coefficient and the material constants is discussed on crystal grown by various laboratories. The influence of the material constants on electrical characteristics of langasite resonators is also presented.

1. Introduction

In some of our previous contributions [3,4,5] we have elaborated an accurate method to design narrow-band quartz monolithic filters and wide-band langasite (LGS) monolithic filters. We have shown that in order to get a much better agreement between input data and experimental results than that obtained with the relation currently used [6] it is necessary to calculate rigorously the dependence of coupling coefficient on geometrical dimensions and mass-loading.
The calculation of geometry of the two – pole structure using the relations between parameters of electric equivalent circuit, deduced from synthesis and geometrical dimensions represents an important step of the monolithic filters design. A relation between coupling coefficient and dimensions (plate thickness, electrode lengths, electrode spacing) is proposed in [3] and another one is deduced by us using an original program based on Tiersten's analysis. This one allows a better correlation of the input data of filters with experimental results.
In the method described by Tiersten the three-dimensional equations of elasticity are converted into a series of two-dimensional equations by expanding the mechanical displacement and electrical displacement in a series of powers of the thickness coordinate of the plate. It was considered an infinite Y-cut rotated plate of thickness t with two pairs of strip electrodes deposited on it. For a narrow frequency range near the cut-off fundamental thickness-shear frequency only the contributions of the fundamental thickness-shear and fundamental flexure mode are retained.
Based on this new and accurate theoretical calculation of the coupling coefficient dependence on structure geometry and mass-loading, in this paper we evaluated the influence of several material constants (elastic constant c_{66}, piezoelectric constant e_{26} and Voigt's elastic constant γ_{11}) on the coupling coefficient of two-pole monolithic filter. The results show that a small dispersion of material constants values for langasite crystals grown in various laboratories strongly modifies the bandwidth of monolithic filters. The influence of the material constant on electrical characteristics of langasite resonators is also discussed.

2. The dependence of coupling coefficient of monolithic filter on material constants

The physical properties of the crystal strongly influence the relation between the coupling coefficient k and the geometrical dimensions of the bipolar structure. The behaviour of coupling coefficient is influenced by the material constants, but the effects of these constants are very different. For applications is interesting to know the weight of each constant on control of the coupling coefficient value. That is why we tried to make a systematic study of the influence of material constants on the piezoelectric coupling coefficient behaviour.
The calculations were performed using the relation deduced from Tiersten's analysis in which was retained only thickness - shear and flexure modes of vibration. As a consequence, only the influence of elastic constant c_{66}, piezoelectric constant e_{26} and Voigt's elastic constant γ_{11} was studied.
The study was performed for two-pole langasite monolithic filter. In Figure 1 is presented the structure of the two-pole monolithic filter with the next sizes:
t - langasite plate thickness; t_e – electrode thickness;
d - distance between electrodes;

l_x - electrode length in X direction;
l_z – electrode length in Z direction.

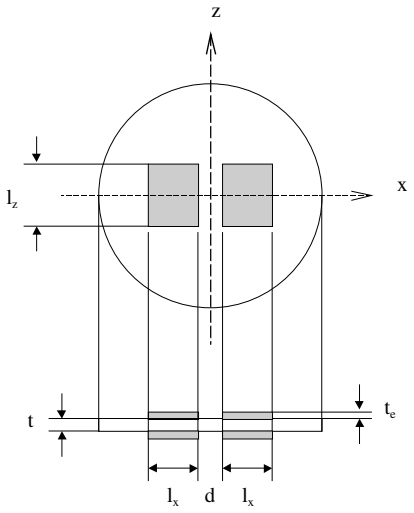


Figure 1. Two-pole monolithic structure

To evaluate the spread effect of these material constants we defined three sensitivity coefficients: $(\Delta k/k)/(\Delta c_{66}/c_{66})$, $(\Delta k/k)/(\Delta \gamma_{11}/\gamma_{11})$ and $(\Delta k/k)/(\Delta e_{26}/e_{26})$. These coefficients are calculated for various electrode spacing d of langasite two-pole structure with mass-loading R, l_x and plate thickness t (R=0.0015, l_x=9t).
The relative variation of k versus relative variation of material constant c_{66} is practically the same regardless the electrode spacing values, but changes with the electrode spacing values in case of e_{26} and γ_{11}.
The sensitivity coefficient of k with c_{66} is –1.246 in the left side of (0, 0) point and –0.6 in the right side (Fig. 2).

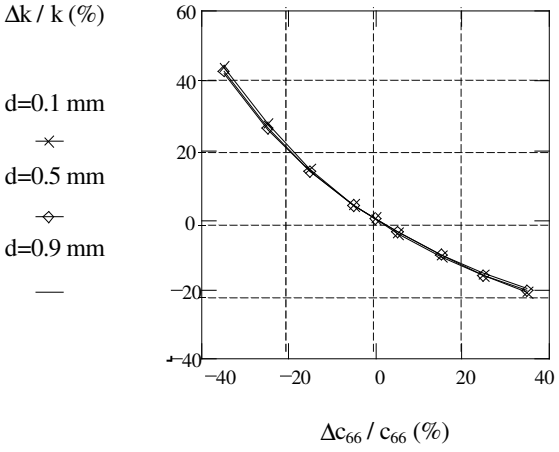


Figure 2. The relative variation of k versus relative variation of elastic constant c_{66}.

In Fig. 3 is presented the dependence of relative variation of k on the relative variation of e_{26} for three values of electrode spacing. The highest sensitivity of k function of the piezoelectric constant e_{26} is obtained for small electrode spacing values.

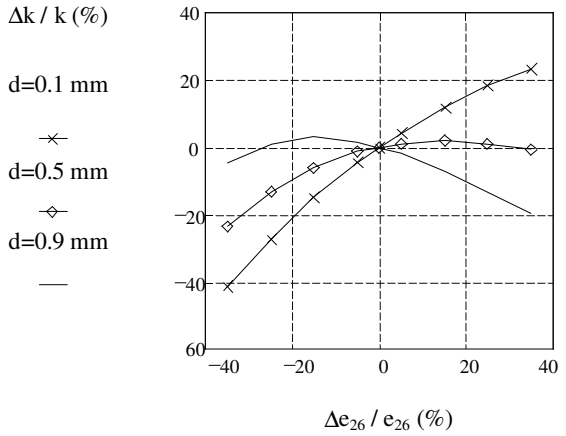


Figure 3. The relative variation of k versus relative variation of piezoelectric constant e_{26}.

For relative variation of e_{26} below zero we obtained the next values for the sensitivity coefficient 1.18 (d=0.1 mm), 0.12 (d=0.9 mm) and above zero the values of this coefficient are 0.67 (d=0.1 mm), -0.57 (d=0.9mm).
Another kind of influence has the electrode spacing on the relative variation of k with the relative variation of γ_{11} (Fig.4): for small values of d, the values of γ_{11} give small changes of the coupling coefficient. The sensitivity coefficient for negative values of $\Delta\gamma_{11}/\gamma_{11}$ is 0.39 if the electrode distance d=0.1, 0.99 if d=0.9mm.
For positive values of $\Delta\gamma_{11}/\gamma_{11}$ this coefficient is 0.22 for d=0.1 mm and 0.88 for d=0.9 mm.

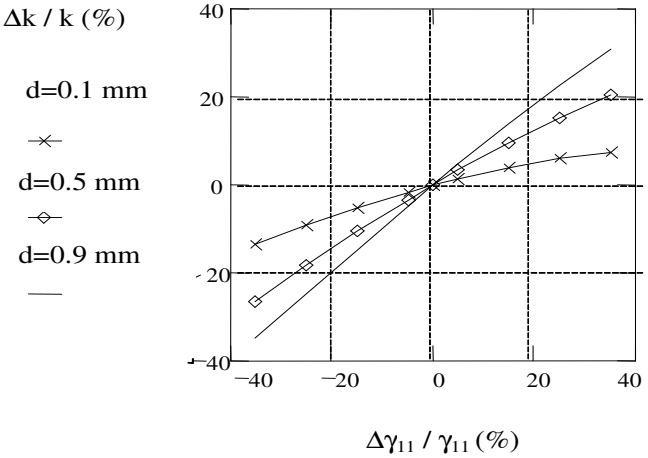


Figure 4. The relative variation of k versus relative variation of piezoelectric constant γ_{11}.

The calculations based on Tiersten's analysis, in which was retained only thickness - shear and flexure modes of vibration, shown that the influence of elastic constant c_{66}, piezoelectric constant e_{26} and Voigt's elastic constant γ_{11} is dominant on coupling coefficient behaviour.

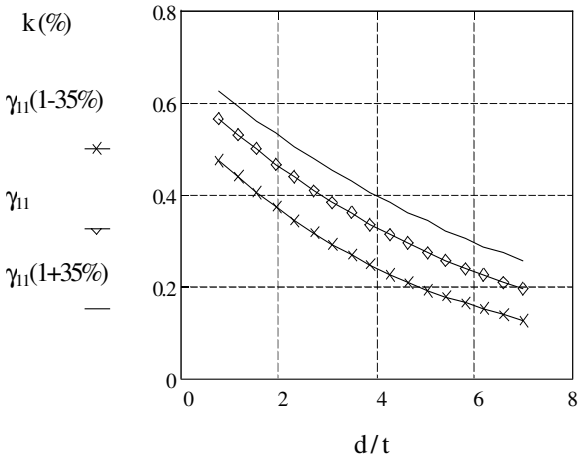


Figure 5. The coupling coefficient versus d / t for large variation of the Voigt's elastic constant γ_{11}.

From the behaviour of these coefficients we can conclude that the most sensitive material constant for filter design are the elastic constant γ_{11} and the piezoelectric constant e_{26} while the filter design is weakly dependent on the elastic constant c_{66}.The dependences of k function on d / t for an arbitrary variation (± 35 %) of material constants have been evaluated.
Fig. 5, 6, 7 show that the coupling coefficient increases with the piezoelectric constant e_{26} and with Voigt's elastic constant γ_{11} and decreases with the elastic constant c_{66}.

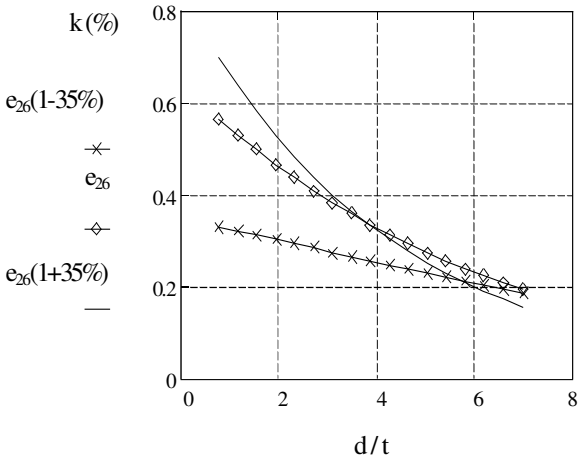


Figure 6. The coupling coefficient versus d / t for large variation of the piezoelectric constant

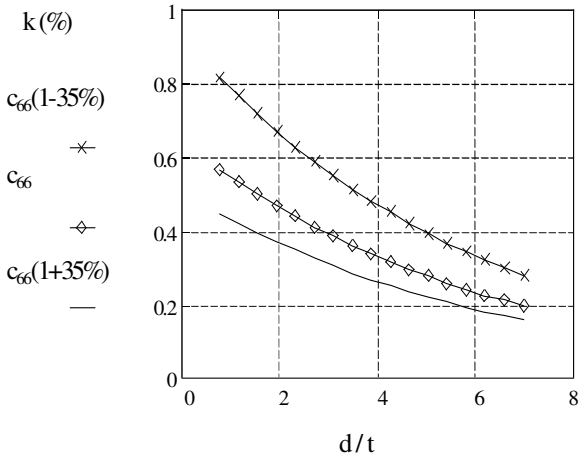


Figure 7. The coupling coefficient versus d / t for large variation of the elastic constant c_{66}.

3. Coupling coefficient dispersion for various langasite samples.

Another purpose this work was the study the influence of material constants dispersion on coupling coefficient behaviour for langasite monolithic filter.
The values of the c_{66}, e_{26} and γ_{11} constants of seven samples grown in different laboratories have been selected from literature. They are presented in Table 1.
We observe that for c_{66} the maximum difference between the minimum and maximum is ±2%, for e_{26} is 2.5% and for γ_{11} ±1.5 %.

Table 1. Material constants of several langasite samples

Sample	c_{66} (10^{10} N/m^2)	e_{26} (C/m^2)	γ_{11} (10^{10} N/m^2)
1 [7]	4.35	-0.41	12.3027
2 [8]	4.2	-0.45	12.0823
3 [9]	4.216	-0.43	12.0829
4 [8]	4.32	-0.45	12.0769
5 [10]	4.22	-0.44	12.1633
6 [11]	4.2	-0.44	12.1425
7 [6]	4.235	-0.45	12.0743

Using the accurate method [] we calculated the dependence of the coupling coefficient function of diameter and plate thickness ratio for langasite two-pole monolithic structures for the constants values of these seven samples (Fig. 8).

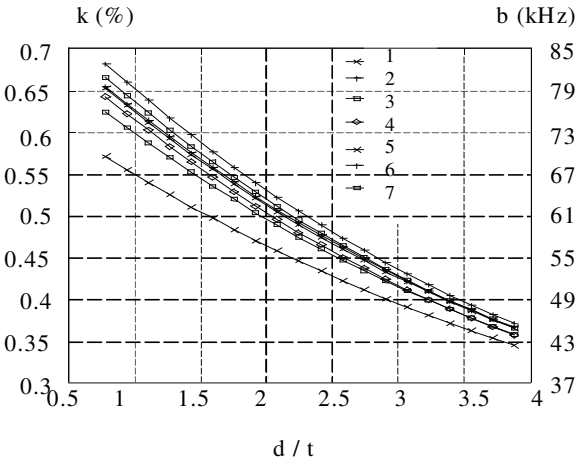


Figure 8. The variation of coupling coefficient versus d / t for seven langasite samples presented in Table 1.

Figure 8 shows that the maximum difference of the coupling coefficient (~11 %) is obtained for material constants dispersion within ±2.5% for small values of d / t (wide bandwidth filters); for narrow bandwidths (large values of d / t) the difference between coupling coefficients corresponding to minimum and maximum constant values is about 4 %.

The spread of the material constants is characteristic for langasite crystals. Langasite is an incongruent-melting material [12,13]. The solidification in the ternary system La_2O_3-Ga_2O_3-SiO_2 is a very complex process and easily can lead to the precipitation of other phases than langasite (or beside langasite). If the synthesis and crystal growing processes are correctly conducted, the pure langasite phase is obtained in a relatively narrow interval of compositions. Function of growing conditions and the composition (maintaining the composition in the allowed interval for good langasite crystals), the material parameters of the langasite crystals, such as c_{66}, e_{26} and γ_{11} may vary. The spread of the devices parameters could produce serious difficulties for design and for mass production.

4. Influence of material constants on L, C of Y-cut langasite resonators

The variation of motional inductance L and capacitance C of Y-cut langasite resonators with 5 MHz fundamental frequency, 14mm diameter, 0.275 mm plate thickness has been calculated for e_{26} and c_{66} dispersion of ±3%.

Relation between electrical characteristics and physical parameters (density, plate thickness, electrode area, e_{26}, c_{66}) are:

$$L = \rho t^3 / 8A e_{26}^2$$

$$C = 8Ae_{26}^2 / \pi^2 t c_{66}$$

The dependence of the inductance L and capacitance C on material constants e_{26} and c_{66} are presented in figure 9 and 10.

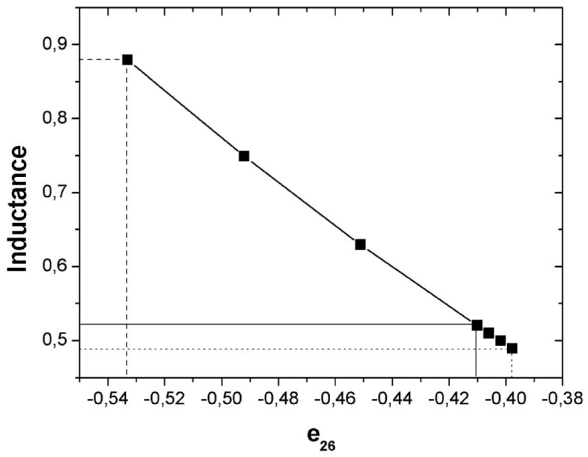


Figure 9. Inductance variation with piezoelectric constant e_{26}

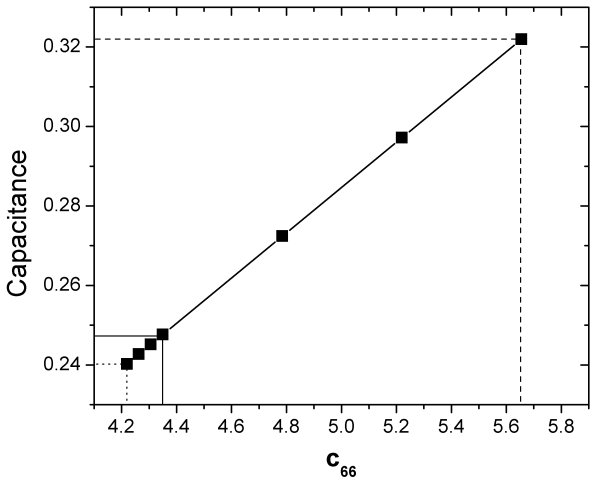


Figure 10. Capacitance variation with elastic constant c_{66}

5. Conclusions

The effect of some material constants on coupling coefficient behaviour was estimated for langasite structure by the calculation of sensitivity coefficients of elastic constant c_{66}, piezoelectric constant e_{26} and Voigt's elastic constant γ_{11}.

The influence of the elastic constant c_{66}, piezoelectric constant e_{26} and Voigt's elastic constant γ_{11} are dominant on coupling coefficient behaviour.

The most sensitive material parameter for filter design is the elastic constant γ_{11} and piezoelectric constant e_{26}, while the elastic constant c_{66} has a small influence on the filter design.

The coupling coefficient increases with e_{26} and with γ_{11} and decreases with c_{66}.
The maximum dispersion of c_{66} value of these 7 samples is ± 2%, for e_{26} is ± 2.5% and dispersion of γ_{11} is ± 1.5%.
The relation between the coupling coefficient and d/t ratio calculated using the accurate method for all LGS samples point out that the maximum variation of the coupling coefficient is obtained for small d/t values.
For 1-2.5% dispersion of langasite material constants (consistent with published results) can result in up to~20% variation of langasite monolithic filter bandwidth.
The spread of the devices parameters could produce serious difficulties for design and for mass production.

6. References

[1] H. F. Tiersten "Linear piezoelectric plate vibrations", *Plenum Press*, N.Y., (1969).

[2] I. Mateescu, J. Zelenka, A. Serbanescu, "Contribution to the design of the wave filters composed from the elastically coupled two resonators structure", *Proc.10th European Frequency and Time Forum*, pp. 132-137 (1996).

[3] I. Mateescu, G. Pop, A. Manea, M. Lazarescu, C. Ghita, "Investigation on behaviour of coupling coefficient of monolithic structure", *Proc.11th European Frequency and Time Forum*, pp. 438-442, (1998).

[4] I. Mateescu, A. Serbanescu, G. Pop, C. Ghita, "The influence of the piezoelectric crystal properties on the monolithic filters design", *Analles de Chimie*, vol. 22, nr.8, pp. 713-716, (1997).

[5] I. Mateescu , G. Pop, C. Ghita, "A new theoretical approach for the coupling coefficient in the monolithic structures", *Proc. 1998 IEEE Int. Freq. Control Symp.*, pp. 580-586, (1998).

[6] S.A. Sakharov, I.M. Larionov A.V. Medvedev., "Application of langasite crystals in monolithic filters operating in shear modes", *Proc. IEEE Int. Freq.Control Symp* , pp. 713-718, (1992).

[7] I. M Silvestrova., V. Bezdelkin, P.Senyushenkov, Yu. Pisarevsky, "Present stage of $La_3Ga_5SiO_{14}$ - research", *Proc. 1993 IEEE International Frequency Control Symposium,* pp. 348-350, (1993)

[8] I. M.Silvestrova, Yu. Pisarevsky, P.A.Senyushenkov A.I. Krupny, "Temperature dependence elastic properties $La_3Ga_5SiO_{14}$", *Sov. Phys. Solid.Stat*, v. 28, 9, pp. 2875-2878, (1986).

[9] S. Sakharov, P.A.Senyushenkov, A.Medvedev, Yu. Pisarevsky, "New data temperature stability and acoustical losses of langasite crystals", *Proc. 1995 IEEE International Frequency Control Symposium*, pp. 647-652, (1995).

[10] A.B.Ilyaev, B.S Umarov, L.A.Shabanova, M.F. Dubovik, "Temperature dependence of electromechanical properties of LGS crystals", *Phys. Stat. Solid*, v. 98, pp. 109-114, (1986).

[11]K.S. Alexandrov, B.P Sorokin, P.P Turchin, "Nonlinear piezoelectricity in $La_3Ga_5SiO_{14}$ piezoelectric crystals", *Ferroelectric Letters*, v.14, nr. 5-6, pp 115 – 125, (1992).

[12]. S.Q. Wang, S. Uda, "Phase relations around langasite ($La_3Ga_5SiO_{14}$) in the system La_2O_3-Ga_2O_3-SiO_2 in air" , *J. Crystal Growth* 250, pp.463 (2003).

[13]. S. Uda, S. Wang, N. Konishi, H. Inaba, J. Harada, "Growth technology of piezoelectric langasite single crystal", *J. Crystal Growth* (2005), (in press).

NEW RESULTS ON HIGH FREQUENCY SURFACE ACOUSTIC WAVES EXCITED ON AN ORIENTED $LiNbO_3$ SINGLE CRYSTAL LAYER TRANSFERRED ONTO SILICON.

Th. Pastureaud[2], B. Biasse[3], B. Aspar[3], W. Daniau[1], W. Steichen[2], V. Laude[1], R. Lardat[2], A. Laens[2], J.-M. Friedt[1], J.-B. Briot[2], S. Ballandras[1]

[1]Institut FEMTO-ST, UMR CNRS UFC ENSMM UTBM 6174, Département LPMO, 32 Avenue de l'Observatoire, F-25044 Besançon Cedex

[2]TEMEX, 399 Route des Crêtes, B.P. 232, 06904 Sophia Antipolis Cedex

[3]CEA-DRT-LETI, CEA/GRE, 17 rue des Martyrs 38054 Grenoble Cedex 9

Abstract

The possibility to use oriented single-crystal thin layer for the excitation of guided acoustic waves opens new perspective in the development of passive high frequency signal processing devices. The combination of Y-cut $LiNbO_3$ atop Silicon is theoretically and experimentally investigated, thanks to the development of a process at the CEA-Leti allowing for the fabrication of such compound substrates. It is shown that non lossy pure shear wave can be excited and detected, exhibiting coupling factor larger than standard bulk substrates and maximising the properties of elastic waves excited using thin piezoelectric layers.

1. Introduction

A lot of work has been dedicated to the fabrication of acoustic wave devices for radio-frequency (RF) applications on stratified media, taking advantage of piezoelectric layers deposited on high acoustic velocity substrates (Sapphire, C-diamond or Silicon). These layers usually are deposited by physical vapor deposition techniques (sputtering), providing C-oriented films exhibiting properties more or less close to those of bulk materials. For $LiNbO_3$ (as for most of highly-coupled single-crystal materials), this crystal orientation is far from the best solution for the excitation of well-coupled surface acoustic waves (K_s^2 must be larger than 5%).

A process has been developed by the CEA-Léti to transfer oriented single-crystal piezoelectric layers onto (100) Silicon, based on the association of an ion implantation and molecular bonding [1]. Combining $LiNbO_3$ and Silicon is particularly attractive for the excitation and propagation of strongly coupled transverse modes (the best solution for RF applications). Theoretical predictions as well as experimental measurements of the Rayleigh-like and shear waves on a 4" $LiNbO_3$(YX)/Silicon wafer operating at 1 GHz already have been presented and compared to the state of the art [2]. These preliminary results have shown that this kind of wave-guide represent the most efficient approach for the excitation of guided elastic wave in the RF domain.

Considering theoretical predictions, one can expect a factor of 2 for the coupling and reflection coefficients when exploiting this mode at twice the tested frequency. As a consequence, a new set of resonators has been fabricated on our substrates in order to reach the optimal working point which correspond to a frequency-thickness product equal to 1 GHz.µm (1 km.s^{-1}). In the first section of this paper, we briefly recall theoretical results obtained using a model based on the combination of a boundary integral method (BIM) and finite element analysis (FEA). We then analyse previous experimental results to try and identify the actual operating conditions of the device to try and accurately predict the wave characteristics at the tested working points. We then report the corresponding experimental results and conclude about the relevance of the proposed approach.

2. Theoretical analysis

In this section we briefly recall the BIM-FEA implemented for the prediction of guided wave characteristics in stratified layers. The model is detailed in ref. [3]. We first introduce the basic geometry of the problem, i.e. an infinite periodic grating lying along x_1 deposited atop a compound substrate composed of piezoelectric or metallic layers stacked along x_2 as shown in fig.1 (the first layer must be at least dielectric but is generally piezoelectric).

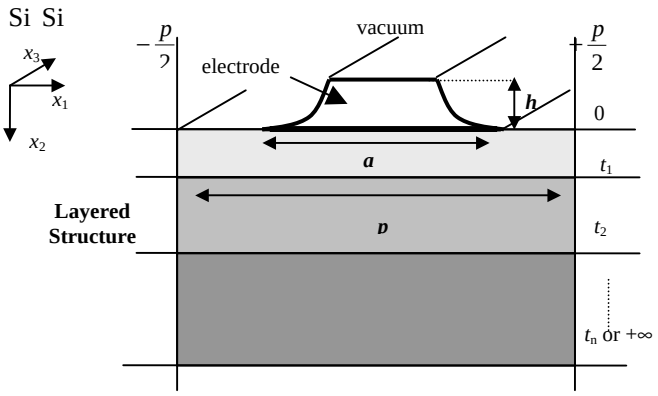


Fig.1 Typical geometry of an infinite grating of period p along x_1 atop a stratified substrate stacked along x_2. a/p is the metallization ratio and $h/2p$ the electrode relative height

Practical interdigital transducers used for the excitation of surface waves generally exhibit apertures large enough to neglect the dependence of the fields along x_3. We assume a spatial harmonic excitation driving the electrode grating, allowing for considering only one period in the formal description of the problem. Using the well established Green's function formalism, the relation between the generalized surface stress $T_{jk}.n_k$ (including the normal electrical displacement) and the generalized displacements u_i (including the electrical potential) is given by :

$$u_i(x_1) = \int_{-p/2}^{+p/2} G^p_{ijk}(x_1 - x', f)\, T_{jk}(x') n_k \, dx' \qquad (1)$$

where the periodic Green's function G^p_{ijk} is defined according to the Floquet theorem combined with the Bloch function.In the following, the normal n_k is considered normal to x_1 without any loss of generality. There is no obvious general analytical solution for eq.(1). As a consequence, the generalized displacement and stress fields are developed as orthogonal polynomial series (Chebychev in this development).

$$\left| \begin{array}{l} T_{2j}(x_1) = \dfrac{\sum_{n=-\infty}^{+\infty} \alpha_{jn} C_n(\bar{x}_1)}{\sqrt{1-\bar{x}_1^2}} \\ u_i(x_1) = \sum_{n=-\infty}^{+\infty} \beta_{in} C_n(\bar{x}_1) \end{array} \right. \quad with: \quad \left| \begin{array}{l} \bar{x}_1 = \dfrac{2a}{x_1} \\ |x_1| < \dfrac{a}{2} \end{array} \right. \qquad (2)$$

In (2), C_n are the Chebyshev polynomials, α_{in} and β_{jn} are the development coefficients which become the actual unknowns of the problem. Inserting (2) into (1) and then projecting the obtained form over the Chebyshev base provides the relation between the stress and displacement coefficients (α_{in} and β_{jn}).

$$\beta_{im} = A_{imjn}\, \alpha_{jn} \qquad (3)$$

where

$$A_{imjn} = \nu_n \frac{\pi^2}{p} j^{m-n} \sum_{l=-\infty}^{+\infty} \tilde{G}_{ij2}(\gamma+l, f) J_m\left(\frac{\pi a}{p}(\gamma+l)\right) J_n\left(\frac{\pi a}{p}(\gamma+l)\right)$$

In (3), J_n are the Bessel's functions of nth order and of first kind and ν_n accounts for the orthogonality properties of the Chebyshev polynomials [3].

Additional equations are required to solve the problem. They are provided by the boundary conditions imposed at the propagation surface. The mechanical stress and displacements are assumed continuous at the electrode/substrate boundary and the normal stresses vanish elsewhere :

$$\left| \begin{array}{ll} u_i = L_{ij}(T_{2j}) & for \quad |x_1| < \dfrac{a}{2} \\ T_{2j} = 0 & for \quad \dfrac{a}{2} < |x_1| < \dfrac{p}{2} \end{array} \right. \qquad (4)$$

The mechanical contribution of the electrode is simulated by FEA. One can establish the relation between α_{in} and β_{jn} generated by the electrode as follows :

$$\beta_{im} = D_{imjn}\, \alpha_{jn} \qquad \text{with}: \quad i, j = 1,2,3 \qquad (5)$$

and

$$D_{imjn} = \left[\int P_m C_q\right] \left[K_{ij} - \omega^2 M_{ij}\right]^{-1}_{\Gamma} \left[\int P_n C_k\right] \qquad (6)$$

where ω is the angular frequency, and the matrix $[\int P_m C_q]$ symbolically holds for the integration of the product of FEA interpolation and Chebyshev polynomials when mixing the FEA and the boundary integral formulation of (1). Finally, an electric boundary condition is imposed, considering a flat electrode assumption set to 1V., fixing then the corresponding M Chebyshev development coefficients as follows :

$$\beta_{40} = 1. \quad and \quad \beta_{4m} = 0. \; for \; m \neq 0 \qquad (7)$$

The linear system constructed with the above equations is solved using standard linear algebraic equation tools. One has access to the harmonic admittance of the structure by computing the total electrical charge Q_{total} under the electrode. The harmonic admittance YH can be normalized to be expressed in Siemens (S) as follows:

$$YH(\gamma, fp) = j2\pi\, fp \frac{\pi a}{2} \alpha_{40} \qquad (8)$$

From this harmonic admittance, one can extract general data about the excitation and guiding of any mode of the structure, i.e. the velocity, the electromechanical coupling, the propagation leakage and the wave diffraction due to massive electrodes. Thanks to FEA, no restriction has to be considered about the electrode shape and height.

3. Theory/experiment assessment

3.1 1 GHz Resonators

In the first set of experiment, synchronous resonators were built consisting in a 50 finger pair central inter-digital transducer (IDT) surrounded by two Bragg mirrors composed of 50 grounded electrodes. In ref.[2], the experimental results were compared to theoretical predictions, pointing out the fact that the phase velocity as well as the coupling coefficient did not match the expected values for the shear wave. The experimental data were respectively found equal to 4385 m.s^{-1} for the phase velocity, and 6% for the electromechanical coupling, with a 7,7% reflection coefficient on a single electrode ($h/2p$=2,5%), in comparison with theroerical values respectively equal to 4824 m.s^{-1}, 12% and 11.2%. The silicon substrate was assumed intrinsic, but the actual resistivity of the substrate was evaluated to 14 Ω.cm, which is far to correspond to an intrinsic material (more than 1000 Ω.cm). We then reconsidered these results and our theoretical predictions, simply inspecting the case of a conductive substrate. Figure 2 compares the dispersion curve corresponding to those two extreme cases.

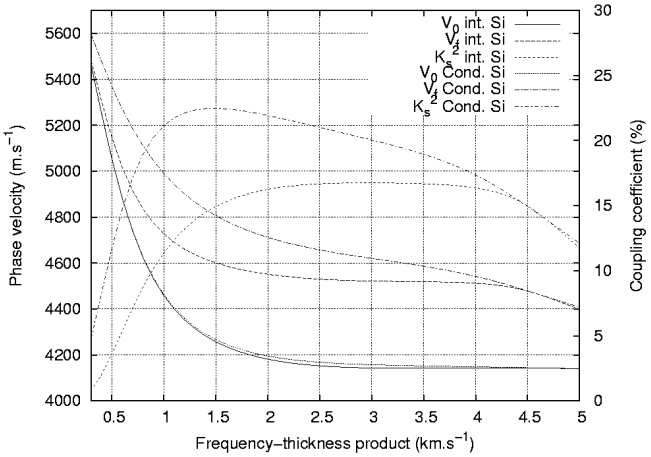


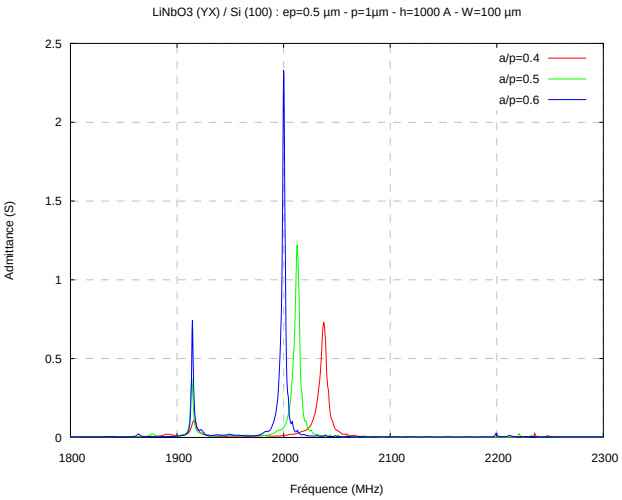
Fig.2 Comparison between dispersion behaviours of the pure shear wave on LiNbO3/Si assuming intrinsic or conductive silicon

Two obvious characteristics arise from this comparison, i.e. the coupling factor dramatically reduces when assuming conductive Si, but consequently the average velocity (between metallized and free surfaces) falls down (the metallized velocities are super-imposed). One can actually see that for a 0.5 km.s-1 frequency-thickness product, a K_s^2 smaller than 5% is predicted for the conductive Si instead of 12% for the intrinsic Si. We then have computed the harmonic admittance corresponding to conductive Si, yielding an electromechanical coupling of 6.2% and reflection coefficient equal to 9%, which fairly approaches the experimental values. There is still a problem with the velocity which remains too large compared to experiments (above 4700 $m.s^{-1}$). No improvement was obtained for the Rayleigh wave, because it is much less sensitive to the electrical properties of the substrate compared to the shear wave. However, the Rayleigh-like wave characteristics were found in good agreement with theory. Another point is that contrarily to what was expected if the silicon would behave like an insulator, we do not meet the optimal conditions for the pure shear guided wave. This aspect will be discussed in the conclusion of the paper.

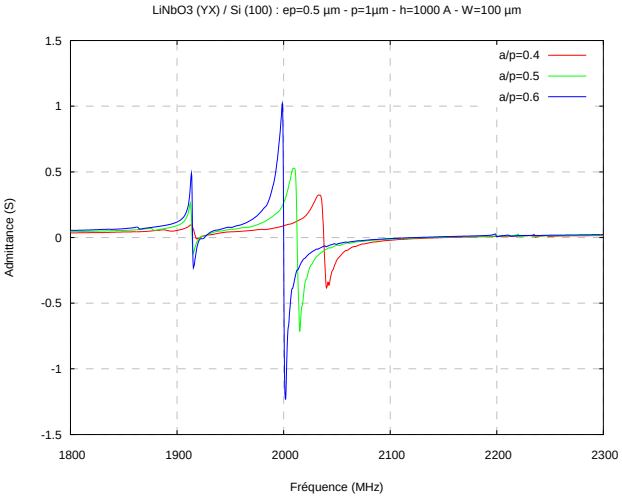
It must be specified that these experiments were conducted at a frequency-thickness working point equal to 0.5 km.s-1, which is far to be the optimum. One can actually see in fig.2 that doubling the frequency-thickness product should yield a coupling coefficient twice larger. We then have decided to perform a new set of resonators operating at a frequency close to 2 GHz to check this hypothesis.

3.2 2 GHz resonators

As in previous experiments, the tested devices consist in single-port synchronous resonators composed of a 50 finger-pair IDT surrounded by two 50-electrode Bragg mirrors. The mechanical period p was 1 µm, and 3 metallization ratio have been tested (0,4 ; 0,5 ; 0,6) to point out the influence of this parameter on the wave characteristics. The electrode height was set to 100 nm ($h/2p$ = 5%), and an acoustic aperture of 100 µm was chosen to get reed of any unwanted diffraction effect. Experimental measurements are plotted in fig.3. The wave characteristics have been extracted from these measurements and reported in Table 1. Harmonic admittance of the device also have been computed, as shown in fig.4, allowing for the characterization of the regarded mode near 2 GHz.



(a)



(b)

Fig.3 Experimental measurements of synchronous resonators operating near 2GHz
(a) conductance (b) susceptance

a/p=0.6	V (m/s)	|R| (%)	K_s^2(%)
"Rayleigh"	3777	4.2	1.3
Shear	4200	15	12.5

Table 1 Experimental characterization of the Rayleigh-like and pure shear modes

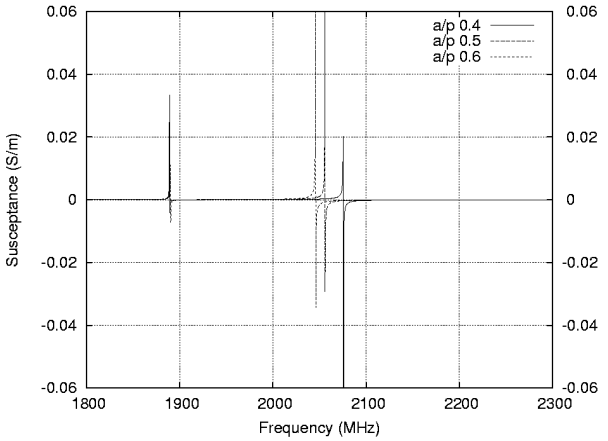


Fig.4 Theoretical harmonic admittance of infinitely periodic electrode gratings on (YX)$LiNbO_3$/conductive Si (p=1µm, $h/2p$=5%)

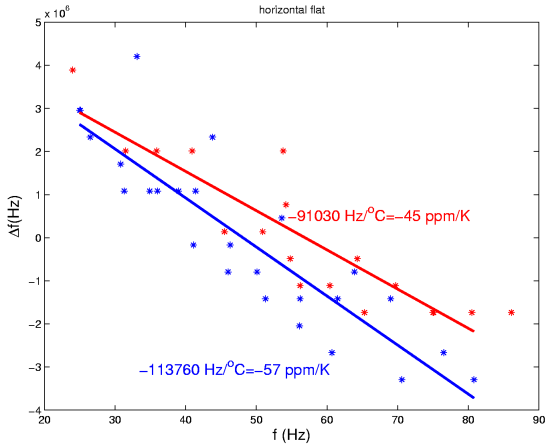


Fig.5 Temperature coefficients of frequency of the 2GHz resonators

The characteristics have been computed for the pure shear wave, yielding a theoretical velocity of 4300 m.s-1, a electromechanical coupling equal to 14.5% as well as the reflection coefficient. Even if the experimental coupling as slightly smaller than the predicted one, it is shown that the material almost behave like expected. The shear mode electromechanical coupling and reflection coefficients are actually doubled when operating at 2 GHz. Also the Rayleigh-like wave properties change significantly : the coupling clearly decreases whereas the reflection coefficient increases. Considering these new results at 2 GHz (K_s^2 equal to 12.5 % and 15% for the reflection coefficient), we definitely establish the state of the art of SAW excited via piezoelectric layers. One can also note that the frequency shift between the three metallization ration is well predicted (comparison between figs. 3 and 4). The resonance of the Rayleigh-like wave arises at the end of the frequency stop-band (negative reflection coefficient). It is found experimentally and theoretically almost insensitive to the metallization ratio.

We also have measured the temperature coefficient of frequency (TCF) of the pure shear mode. This result is reported in fig.5. The previous TCF characterisation which were performed on small $LiNbO_3$/Si parts did provide a value of –38 ppm/K, which was twice smaller as expected theoretically [2]. In the present set of experiment, the TCF is found in the range of –50 ppm/k (see fig.5). Such a value is also quite different from theoretical predictions, and tends to indicate either that the theoretical assumptions of the TCF calculation do not match the real situation, or that some thermo-elastic properties are affected by the bonding process. The second hypothesis is coherent with the fact that discrepancies still persist between theory and experiment concerning the phase velocities of the Rayleigh-like and pure shear guided waves.

4. Conclusion

In this paper, we have reconsidered the simulation of elastic wave guided in single-crystal thin layers of (YX) lithium niobate deposited atop silicon using a molecular bonding process. Considering a conductive silicon substrate instead of an intrinsic one, a good agreement is found between theoretically predicted and experimental electromechanical coupling coefficient of the pure shear wave (the second mode of the structure). A new set of resonators operating in the vicinity of 2GHz has been fabricated and characterized. The electromechanical coupling coefficient of the pure shear wave has been doubled, and a reflection factor equal to 15% has been measured. Theoretical predictions of these characteristics are in good agreement with experiments. As a consequence, improving the knowledge of the elastic properties of the substrate remains the main difficulty to enable the use of such substrates for real filtering applications. Moreover, as shown theoretically, the tested operating points are far from the optimal one, corresponding to a frequency-thickness product equal to 3 GHz. Consequently, we can expect better result if using thicker niobate layers. Another way to benefit from the optimal operating conditions would consists in reducing the grating period down to 0.25 µm. In both cases technological difficulties have to be overcome, but one should note that the obtained characteristics are quite sufficient to address modern filter requirements, yielding a very interest for this above-IC solution for integrated signal processing devices.

Acknowledgement : This work has been performed within the French National Network for Research in Communication (RNRT) under grant

References :

[1] B. Aspar, H. Moriceau, E. Jalaguier, C. Lagahe, A. Soubie, B. Biasse, A.M. Papon, A. Claverie, J. Grisolia, G. Benassayag, F. Letertre, O. Rayssac, T. Barge, C. Maleville, and B. Ghyselen, Journal of Electronic Materials, 30(7), 834 (2001)

[2] M. Solal, Th. Pastureaud, S. Ballandras, B. Aspar, B. Biasse, W. Daniau, W. Steichen, V. Laude, A. Laëns, Proc. of the IEEE Ultrasonics Symposium, München, pp. 128-131, 2002

[3] S.Ballandras, A. Reinhardt, V. Laude, A. Soufyane, S. Camou, W. Daniau, T. Pastureaud, W. Steichen, R. Lardat, M. Solal, P. Ventura, J. Appl. Phys., Vol. 96, n°12, pp. 7731-7741, 2004

CHEMICAL CONTROLLED $GaPO_4$ DISSOLUTION HIGH FREQUENCY ANTIMESA AT-RESONATORS MANUFACTURING

N. Prudhomme[1], D. Cachau-Herreillat[1], V. Flaud[2], P. Papet[1] and O. Cambon[1].
1- Laboratoire de Physicochimie de la Matière Condensée, UMR CNRS 5617,
2- Institut Gerhardt
Université Montpellier II, cc003, Place E. Bataillon,
34095 Montpellier, Cedex 5, France.

ABSTRACT

The chemical dissolution of gallium orthophosphate in phosphoric acid is reminded with the influence of different additives, in particular phosphate compounds. For more important dissolution depths than 40µm, the H_3PO_4-$GaPO_4$ mixtures solutions are not suitable, producing very important surface defects like holes. A (NH3 – K3PO4) basic solution is used to dissolve up to 250µm of material thickness. The measure of the roughness either with a profilometer or with AFM spectroscopy method gives an average roughness parameter Ra of 0.02µm without generating holes or important surface defects. The piezoelectric properties of the etched plates are measured by the air-gap method during the reducing of the thickness. No modification of the quality factor Q is observed up to 200µm etching depth.
A three dimensional dissolution process (mechanical mask) is then developed in order to manufacture AT antimesa resonators. Up to now, according to the raw material quality, the highest obtained frequency for the fundamental mode is 26.5MHz (QF= $1.1*10^{12}$) and 16.5MHz (QF= $1.6x10^{11}$) after 88µm and 150µm dissolution depth in H_3PO_4+$GaPO_4$ and NH_3+K_3PO_4 solvent respectively.

1. INTRODUCTION

Gallium orthophosphate is now a well known material in the quartz isotypes group. Due to its physical and piezoelectrical properties (coupling coefficient k=16% and thermal stability up to 950°C) $GaPO_4$ is a good candidate for high tech applications such as pressure sensors at high temperature, viscosimeters, microbalances and VCXO devices [1-3]. However, the required resonance frequencies for electronic devices are increasingly higher, thus the manufacturing process must be further developed. In the case of shear-mode BAW devices, the resonance frequency is inversely proportional to the thickness of the piezoelectric plate. Thus, the reduction of the plate thickness is an essential step in the resonator manufacturing process. The chemical dissolution is often selected to gradually reduce the thickness of the plates. In the case of $AlPO_4$ material, it has been shown [4, 5] that it is possible to reduce the thickness of the plates, improving in the same time the surface state, by using the solvent suitable for the crystal growth which contains an important solute concentration. In this work the mixture $AlPO_4$ 2.7M – H_3PO_4 9M was selected to manufacture high frequency resonators. The same way was followed for quartz chemical dissolution and the selected solvent was NaOH [6-9].
For $GaPO_4$ material, the chemical etching of AT-cut plates has been studied in orthophosphoric acid solutions containing different solute concentrations for dissolution depths up to 40 microns [10]. Anisotropy of the dissolution velocity was characterised by using different crystallographic plate orientations [11]. For very important dissolution depths (up to 250 microns of dissolved material), a new solvent is used. Moreover a masking process is developed to dissolve only the central part of the plates in order to manufacture high frequency antimesa AT-plates.

2. EXPERIMENTAL

The initial $GaPO_4$ single-crystals are grown from either the slow temperature increasing method or the vertical temperature gradient method [12, 13]. Then, the plates are cut along different crystalline orientations. The AT-cut samples are issued from crystal grown by Piezocryst Company*.
The plate thickness decrease during chemical etching is followed by measuring the resonance frequency (F = K/e) using the air gap method [14] with a HP 8753 network analyser.
The frequency constants K of the different cuts are calculated with elastic constant set of Reiter and al [15].

$$K_X = \tfrac{1}{2}\sqrt{(C_{11}/\rho)} = 2159 \text{ MHz/µm}$$

$$K_Y = \tfrac{1}{2}\sqrt{(1/2(C_{11}-C_{12})/\rho)} = 1252 \text{ MHz/µm.}$$

$$K_{AT} = \tfrac{1}{2}((C_{66}\cos^2\theta+C_{44}\sin^2\theta+2C_{14}\cos\theta\sin\theta)/\rho)^{1/2} = 1226 \text{ MHz/µm (with } \theta = -15.9°)$$

The thickness of the Z-plate is evaluated by using a micrometer because Z-cut orientation is not piezoelectrically excited.

* Piezocryst, Graz, Austria.

Author to whom correspondance should be addressed ;email : ocambon@lpmc.univ-montp2.fr

During the dissolution process, the concentration and the temperature are maintained constant.
The surface state evolution is checked by optical microscopy observations with an Olympus BX51 microscope with different magnifications and polarised light. The roughness average Ra is measured with a Mahr surface profilometer. An Autoprobe CP ThermoMicroscopes Atomic Force Microscope is also used to analyse more accurately the surface.

3. RESULTS AND DISCUSSION

Like for $AlPO_4$ material, first results have been obtained on $GaPO_4$ chemical etching in phosphoric acid solvent with different solute concentrations. The complete dissolution kinetics law has been determined for the AT-cut. Concerning the surface state evolution, it has been demonstrated that these solvent can be used for removing up to 40µm of material [10]. For deeper chemical etching, holes can appear according to the raw material quality. Very different etching rate values have been measured for the four different orientations [11]. The orthophosphoric acid solvent reveals the anisotropy of the material structure and consequently the structural defects. For increasing the etching depths, a new solvent is studied: NH_3+K_3PO_4.

3.1 Dissolution kinetics in NH_3 – K_3PO_4 solvent.

Because of the very low boiling temperatures of the solutions in the NH_3-H_2O system, it is very difficult to maintain constant the concentration of the solutions. Nevertheless by dissolving K_3PO_4 compound in the solution, the boiling point is increased. Finally the (NH_3 3.5M + K_3PO_4 1.5M) mixture is selected for $GaPO_4$ dissolution study. The dissolution rates vary from several µm/h to 160 µm/h in the 35°C-80°C temperature range for the four studied crystal orientations (fig 1). As for the study of the dissolution in the phosphoric acid solvent [11], the fitting of experimental data by least square method carry out analytical expressions for the dissolution rate versus temperature.

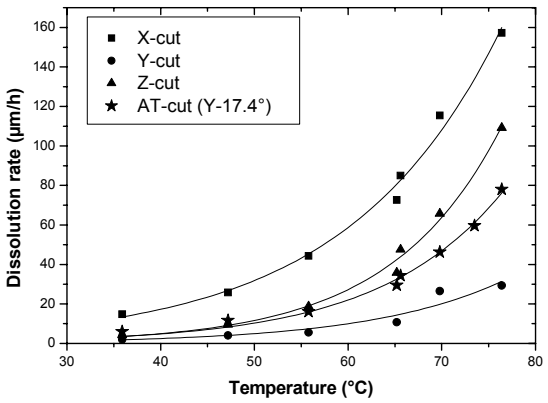


Figure 1: Dissolution rates in NH_3-K_3PO_4 solution

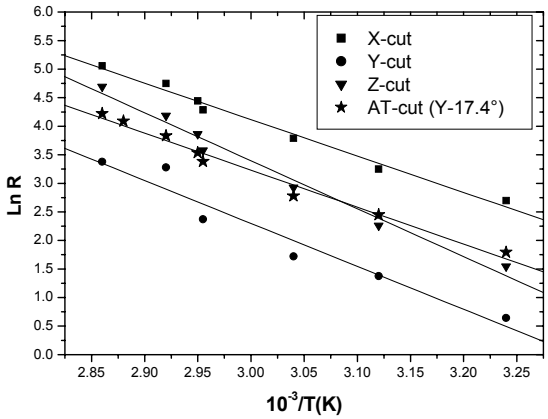


Figure 2: Determination of the parameters of the dissolution rate equations.

The general expression of the dissolution rate R is: $R= Aexp(-Ea/RT)x[solvent]^{\alpha}$. Only the temperature influence (Ea computing, fig.2 and Table I) can be determined because we have just studied one solvent concentration.

Table I: Ea calculation for different orientations.

Cut orientation	Ea (kJ x mol^{-1})
X-cut	53.1
Y-cut	62.5
Z-cut	69.9
AT-cut (Y-17.4°)	53.9

3.2 Surface state evolution in NH_3 – K_3PO_4 solvent.

The evolution of the average roughness parameter Ra depends on the plate orientation and the initial surface state (fig 3). After 300µm etching depth, the average roughness of a polished X-cut is quite constant and the initial rough surface state (Ra= 0.4µm) of a AT-cut (Y-17.4°) becomes smoother (Ra= 0.2µm). The surface states of Y and Z-cuts are deteriorated by the dissolution process; Ra parameter increases drastically up to 0.8 and 0.9µm respectively.

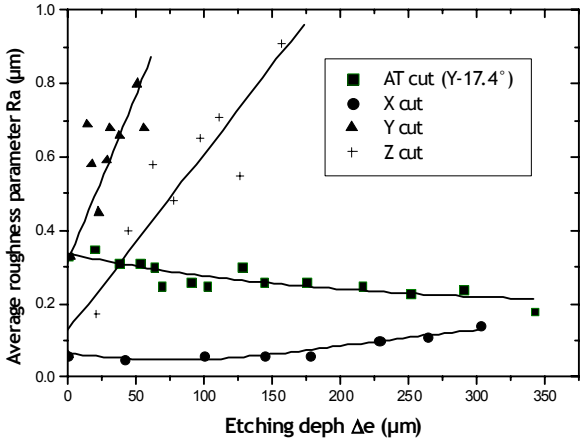
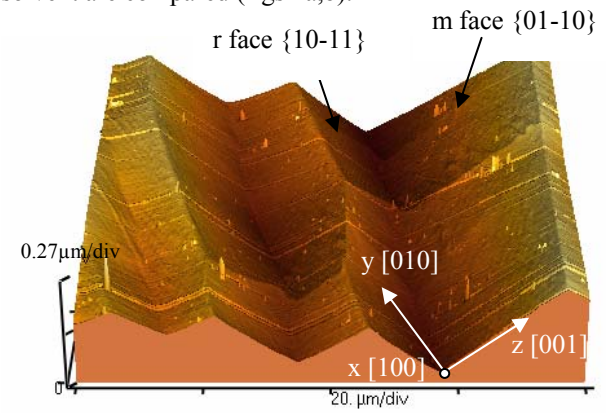
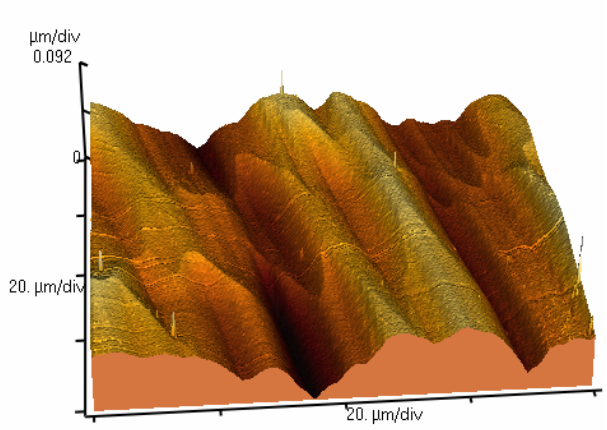


Figure 3: Roughness average evolution during the chemical dissolution in NH_3-K_3PO_4 solution.

All these roughness measurements are in a good agreement with the surface observations by using the Atomic Force Microscope instrument. The surface profiles are recorded on initially polished AT-cut plates. The effect of the phosphoric acid solvent and ammonia solvent are compared (figs 4a,b).



(a)



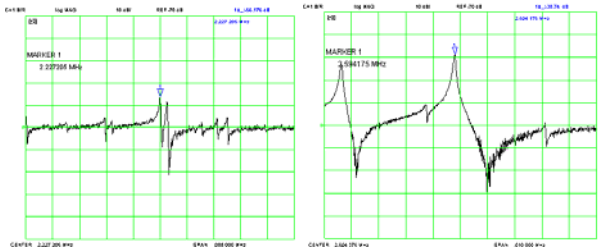
(b)

Figure 5: AFM 3D-profiles of AT-cut plates
(a) H_3PO_4 9M/$GaPO_4$ 0.4M/K_3PO_4 1M T= 82°C and 80µm etching depth (Ra= 0.1µm)
(b) NH_3 3.5M/K_3PO_4 1.5M T= 80°C and 64µm etching depth (Ra= 0.02µm)

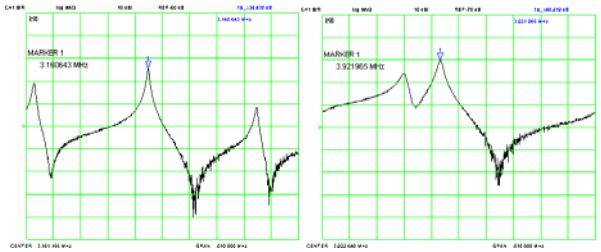
The phosphoric acid solvent tends to reveal the structure of the material with deep etch figures. The planes m {01-10} and r {10-11} appear proving that the dissolution rate is the weakest for these directions. This is also the lowest crystal growth direction. The ammonia solvent produces smoother etch patterns and the Ra parameter is lowered by using NH_3 solvent (Ra= 0.02µm) instead of H_3PO_4 solvent (Ra= 0.3µm). Moreover, no hole is observed for prolonged etching depths in ammonia solvent.

3.4 Evolution of piezoelectric properties during the dissolution in NH_3-K_3PO_4 solvent.

This part concerns the air-gap piezoelectric measurements of the previous AT (Y-17.4°) sample that the surface roughness has been measured after each step of chemical dissolution in NH_3-K_3PO_4 solution (figs 6a to 7d). The surface was initially rough with a Ra parameter value of 0.32µm (see the figure 3).
All these results are presented with the same amplitude scale of 10dB/div that allows directly appreciating the improving of the piezoelectric response. The first signal is that of a bad resonator, and then the signal is improved up to 175µm etching depth. For 250µm etching depth the amplitude begins to decrease and the bandwith enlarges.



(a) : 0µm (b) : 68µm



(c): 175µm (d): 250µm

Figures 6: Piezoelectric measurements for different etching depth of a AT-cut plate with an initial rough surface state.

The performances (Bandwidth +/-3dB and the Q factor) of this resonator have been precisely measured (figures 7 and 8). After 200µm etching depth, the bandwidth value increases and the QF factor begins to degrade.

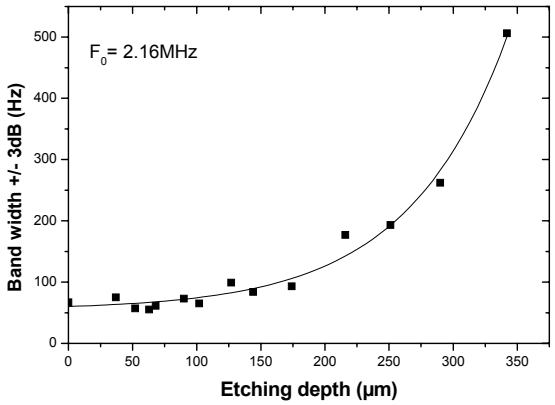


Figure 7: Bandwidth evolution during the chemical dissolution

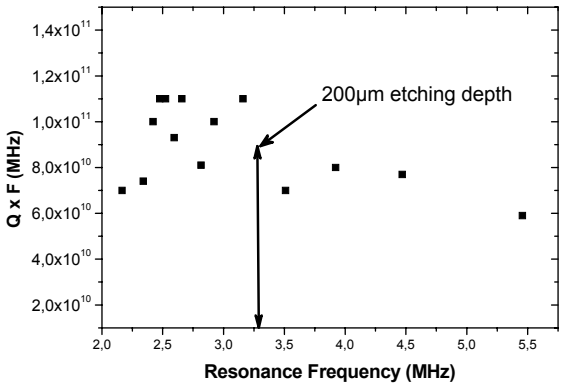


Figure 8: QF factor evolution during the chemical dissolution

3.4 Antimesa plates design.

To reach high frequency plates, the thickness of the plates have to be very thin (lower than 50µm). To ensure the mechanical resistance of the samples, it is necessary to keep a thicker crown on the peripheral zone of the plate to manufacture antimesa plates (fig 9.a). From our knowledge about chemical etching during a long time [10], it is necessary to use a mechanical mask to maintain the sample into the solution and to protect, in the same time, the peripheral crown against the dissolution (fig 9.b).

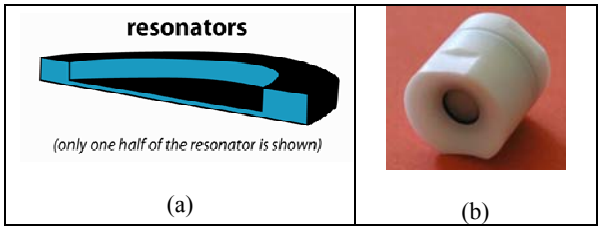


Figure 9. a) Antimesa plates
b) Mechanical mask in PTFE

All the samples are initially AT polished plates with two diameters: 5 and 10mm.

a. Plates with 10mm diameter:

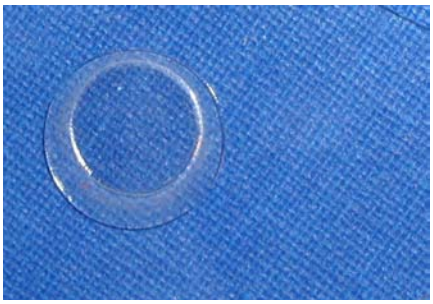
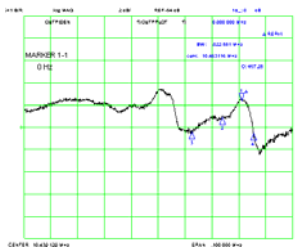
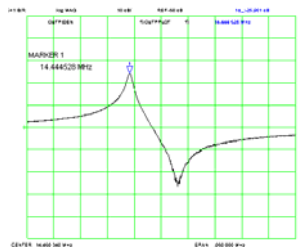
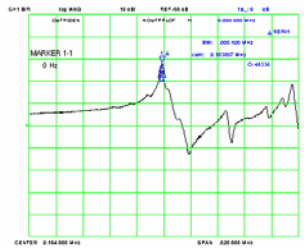
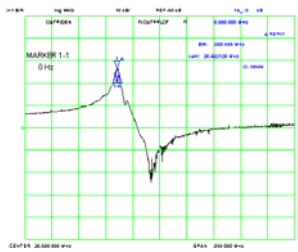


Figure 10: Antimesa plate at 18MHz manufactured in H_3PO_4/$GaPO_4$ solvent starting from 10MHz (54 µm chemical engraving)

Before etching	After chemical dissolution
	
	K_3PO_4,H_2O 2M H_3PO_4 9M
F_0= 10.43MHz (117.5µm)	F= 14.444 MHz (85µm) QF= $2.2x10^{11}$

After chemical dissolution

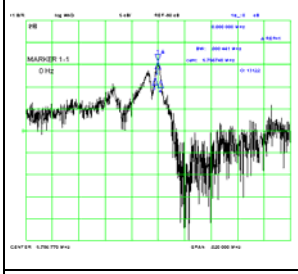
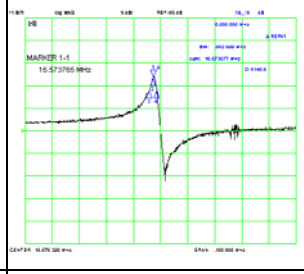


Before etching	After chemical etching
	
	K_3PO_4,H_2O 0.5M H_3PO_4 9M $GaPO_4$ 0.4M
F_0= 9.154MHz (134µm)	F= 26.4MHz (46.5µm) QF= $1.1x10^{12}$



Figures 11: Antimesa resonators manufactured in phosphoric acid solutions.

Although holes are formed in the piezoelectric membrane, good piezoelectric signals are measured on the good parts of the plates chemically etched in phosphoric acid solutions.

Before chemical etching	After chemical dissolution
	
F= 5.786MHz (230µm)	F= 16.573MHz (80µm) QF= $1.6x10^{11}$

External crown (antimesa design)



Figure 12: antimesa resonator manufacturing in NH_3+K_3PO_4 solution.

After 150µm etching depth in NH_3+K_3PO_4 (fig 12), the surface state is good (only some stripes due to initial polishing appear) and no hole is formed during the dissolution. The piezoelectric signal is improved (QF =$1.6x10^{11}$) but does not reach very high value. This could be due to the intrinsic quality of the material and particularly to the OH-content.

b. Plates with 5mm diameter:

For very high resonance frequencies, smaller plates are used (5mm diameter). Results of experiments of chemical engraving realized with NH_3-K_3PO_4 solvent are shown in figure 13. The external crown is well design and the surface state is homogeneous. Only one big defect appears due to an isolated defect in the raw material. The light diffraction bands observed in fig.13 proves that the piezoelectric membrane is not perfectly parallel. This problem could be due to the bad circulation of the solvent around the plate during the dissolution.

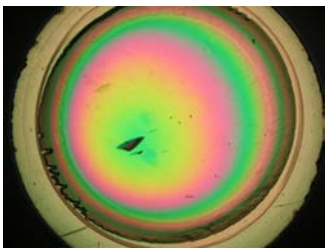


Figure 13: Antimesa chemically designing in NH_3-K_3PO_4 solution (etching depth= 76.3µm).

4. CONCLUSION

The chemical controlled dissolution of $GaPO_4$ material has been studied by using two kinds of solvent: one based on orthophosphoric acid and the second one based on the ammonia solvent. According to the intrinsic material quality, acid solutions produce some holes above 40µm etching depth while based ammonia solvent allows dissolving up to 200µm without generating any hole in the plates. However the QF factor is less than that of quartz due to the material quality. Well designed antimesa plates for high frequencies have been obtained by using a masking process but the experimental dissolution process is not yet optimised and needs to be improved to ensure a good circulation of the solvent around the plate in order to obtain very flat and parallel piezoelectric membranes.

AKNOWLEDGMENTS

The authors are grateful to European Union for his financial support (GAPOGROWTH N°G5RD-CT-2002-00709) and PIEZOCRYST for providing $GaPO_4$ plates.

REFERENCES

[1] P.W. Krempl, F. Krispel and W. Wallnöffer, Industrial development and prospects of $GaPO_4$, Ann. Chim. Sci. Mat., vol. 22, pp. 623-626, 1997.
[2] J. Detaint, B. Capelle, A. Zarka, D. Palmier and E. Philippot, Frequency control and filtering applications of the new piezoelectric materials of class 32, Ann. Chim. Sci. Mat., vol. 22, pp. 651-660, 1997.
[3] W. Wallnöffer, J. Stadler and P. W. Krempl, Temperature dependence of elastic constants of $GaPO_4$ and its influence on BAW and SAW devices. Proceeding of 7th European Frequency and Time Forum, Neuchâtel, 16-18 march 1993, FSRM Ed., pp. 653-675.
[4] O. Cambon , A. Goiffon, A. Ibanez, E. Philippot, Kinetics and dissolution mechanism of berlinite $AlPO_4$, in acid medium, European Journal of Solid State Chemistry, vol. 29, pp. 547-561, 1992.
[5] O. Cambon, A. Goiffon, A. Ibanez, E. Philippot, Controlled dissolution of crystals : application to berlinite ($AlPO_4$ α), a piezoelectric material, European Journal of Solid State Chemistry, vol. 103, pp. 240-252, 1993.
[6] M. Deleuze, A. Goiffon, A. Ibanez, E. Philippot, O. Cambon, Controlled crystal dissolution applied to quartz, J. of Materials Science, vol. 29, pp. 5390-5398, 1994.
[7] O. Cambon, M Deleuze, J.P. Michel, J.P. Aubry, E. Philippot, Controlled dissolution of quartz material. Part II. Quartz chemical etching applied to blanks industrial manufacturing, J. de Phys. IV, C2, pp.85-91, 1994.
[8] O. Cambon, JP. Michel, J. Lamboley, JP. Aubry, Y. Billon, D. Mouton, I. Lozach, An industrial quartz

manufacturing process by chemical etching, 8th European Frequency and Time Forum, pp. 256-262, 1994.
[9] O. Cambon, L. Semene, D. Cachau-Herreillat, A. Goiffon, E. Philippot, Elaboration of quartz devices by controlled dissolution, Ann. Chim. Sci. Mat, vol. 22, pp.589-594, 1997.
[10] N .Prud'homme, D. Cachau-Herreillat, P. Papet, O. Cambon, Gallium orthophosphate device manufacturing by chemical etching, 2003 IEEE International Frequency Control Symposium, Tampa, 4-8 may 2003, John R. Vig Ed., pp. 650-653.
[11] N .Prud'homme, D. Cachau-Herreillat, P. Papet, O. Cambon, Chemical etching of gallium orthophosphate: application to piezoelectric device manufacturing, J. de Phys IV, 4th EWPM, Montpellier, 21-23 july 2004, to be published.
[12] O. Cambon, P. Yot, D. Balitsky, A. Goiffon, E. Philippot, B. Capelle and J. Detaint, Crystal growth of $GaPO_4$: a very promising material for BAW devices manufacturing, Ann. Chim. Sci. Mat., vol. 26, pp. 79-84, 2001.
[13] W. Wallnöffer, P. W. Krempl and F. Krispel, Ann. Chim. Sci. Mat., vol. 26, pp. 75-78, 2001.
[14] J. Détaint, J. Schwartzel, C. Joly, E. Philippot, Determination of the cut-off frequency of unelectroded plates and effects of local thickness modifications in plane resonators, 3rd EFTF, Besançon 21-23 march 1989.
[15] C. Reiter, P.W. Krempl, H. Thanner, W. Wallnöffer, P. Worsch, Ann. Chim. Sci. Mat., vol. 26, pp. 91-94, 2001.

Posters Session 1

DESIGN AND FIRST RESULTS OF THE NEW PTB CAESIUM FOUNTAIN CSF2

S. Weyers, A. Bauch, D. Griebsch, R. Schröder, R. Wynands
Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany

ABSTRACT

At PTB a second caesium fountain clock, CSF2, is in the process of being set up. It differs from the first PTB caesium fountain standard CSF1 in a number of details, which are consecutively specified. First cold atoms have been prepared in the new apparatus.

1. Introduction

The caesium fountain frequency standard CSF1 has been operating at PTB since 1999. So far, CSF1 has contributed fifteen times to international atomic time (TAI) as a primary frequency standard. A relative short term instability of $2 \cdot 10^{-13}$ $(\tau/1\mathrm{s})^{-1/2}$ and a relative uncertainty slightly below 10^{-15} have been obtained [1,2]. However, further diagnostics and improvements are hampered by the fact that in our laboratory only hydrogen masers can reasonably be used as frequency references, which often exhibit daily relative frequency variations of several 10^{-15}. This was a major reason for the decision to start construction of a second fountain standard. CSF2 differs from CSF1 in a number of details. It will become operational in the near future.

2. Design features of CSF2

Fig. 1 shows a sketch of the new PTB fountain CSF2. Differing from CSF1, in CSF2 a (1,1,1) laser beam geometry makes it possible to use six large-diameter (22 mm at the $1/e^2$ intensity points) laser beams for cooling of the atoms in an optical molasses. Fiber output couplers with integrated collimating lenses are directly attached to six ports of the cooling chamber. The direction in which atoms are launched can be controlled by a swivel device which allows us to swivel the bottom part of the fountain including the cooling region and a 25 l/s ion pump below it around a pivot point at the center of the cooling region. The small ion pump improves the residual gas pressure in the cooling region.

A stainless steel TE_{201} state selection cavity is installed just above the cooling region, followed by the detection region, so that one of the horizontal detection laser beams can be used for pushing the remaining $F = 4$ atoms after the state selection process following the launch of the atoms. The dimensions of the cavity are chosen such that this cavity is suited as well for a state selection of rubidium atoms (^{87}Rb, 6.8 GHz), if needed for future operation.

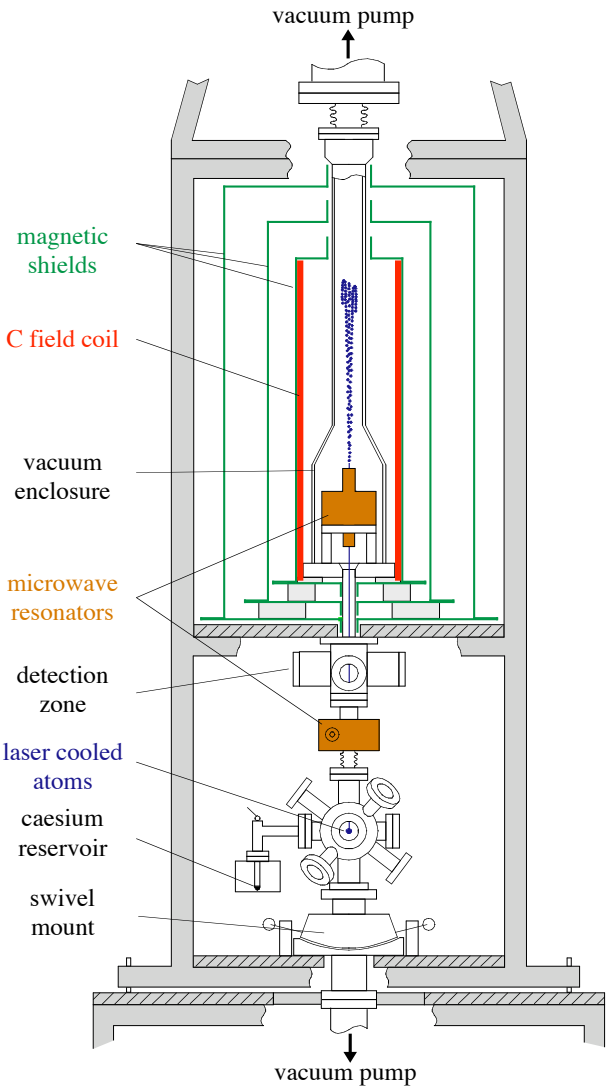


Fig. 1: Sketch of the CSF2 setup

The TE_{011} Ramsey cavity is a copy of the cavity used in CSF1 [3].

The detection region comprises two photodetectors with the respective optics for imaging of the atomic fluorescence, located at opposite sides of the fountain and aligned for the two detection zones for $F = 4$ and $F = 3$ atoms, respectively.

The main vacuum tube is made of copper. This ensures a higher thermal conductivity than for the titanium that is used in CSF1. A three-layer magnetic shield was designed with improved geometry and better contact between the shielding cylinders and their end caps.

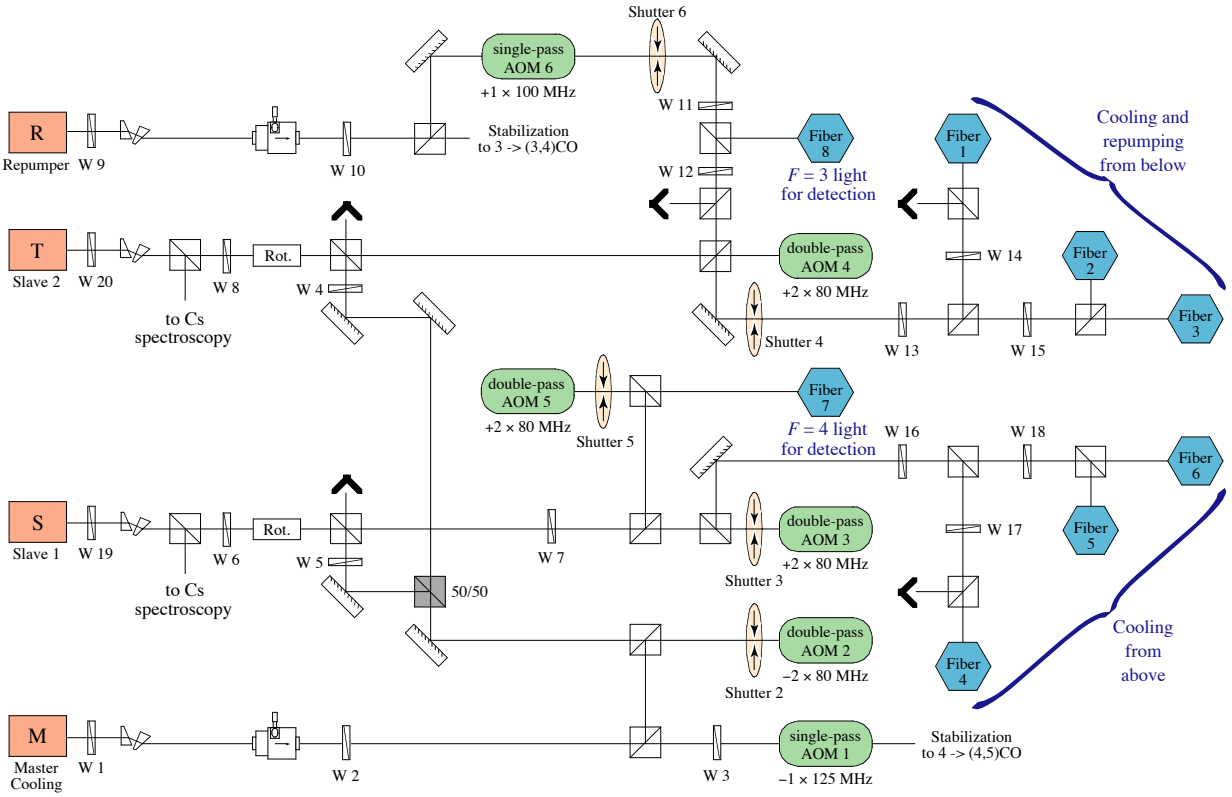


Fig. 2: Sketch of the optical system for CSF2 (W1-W20: half-wave plates)

The optical system (Fig. 2) has been redesigned and includes two commercial extended-cavity diode laser systems (DL 100, TOPTICA), which are used as master and repumping laser, respectively. The master laser injects two 150 mW slave laser diodes, which provide the laser light for cooling and detecting the atoms using the $F = 4 \rightarrow F' = 5$ transition.

A portion of the master laser beam is frequency modulated (AOM 1) and used for locking to a crossover line of the caesium saturated-absorption spectrum. Hence there is no frequency modulation on the corresponding cooling and detection light beams.

The laser light is coupled into polarization maintaining optical fibers and guided to the cooling and detection regions of the fountain, respectively.

As a last point, it should be mentioned that the electronics for the control of the optical system has undergone a complete redesign. It now uses phase-locked loop techniques instead of several arbitrary-function generators and single-sideband mixers, as it was initially the case for CSF1.

3. Present status of CSF2 and future plans

The vacuum system has been completed recently and a base pressure of less than 10^{-7} Pa is obtained. We succeeded to cool first clouds of cold caesium atoms in an optical molasses recently.

Next steps will be the implementation of the new control electronics, launching of atoms, homogenization of the C field by supplementary coils and generation of Ramsey fringes.

For the frequency synthesis initially a system built for the conventional PTB primary caesium clocks will be used, which is based on a commercial high performance BVA quartz oscillator. Recently the development of an improved frequency synthesis system in collaboration with the NPL India has been started. The new frequency synthesis system will also be capable of applying phase modulation instead of frequency modulation for the interrogation of atoms.

In the medium term, for a better control of the effects of cold collisions we also plan the implementation of the technique of population transfer by adiabatic passage [4].

4. Outlook and conclusions

The second PTB caesium fountain CSF2 will become operational in the near future. Full operational capability as a primary standard is not expected before 2006, however.

With two caesium fountains in our laboratory we expect to make effective progress in the uncertainty and instability evaluation of both fountains, as in the past the lack of a proper reference was a major obstacle. In the longer run it is envisioned that one of the fountains could contribute to TAI on a near-continuous basis while the other fountain may be used for further investigations.

5. References

[1] S. Weyers, A. Bauch, R. Schröder, Chr. Tamm, The atomic caesium fountain CSF1 of PTB, in: Proceedings of the 6th Symposium on Frequency Standards and Metrology 2001, University of St. Andrews, Fife, Scotland, World Scientific, p. 64-71.

[2] A. Bauch, R. Schröder, S. Weyers, Discussion of the uncertainty budget and of long term comparison of PTB's primary frequency standards CS1, CS2, CSF1, in: Proceedings of the 2003 IEEE FCS and the 17th EFTF, Tampa, USA, p. 217-221.

[3] R. Schröder, U. Hübner, D. Griebsch, Design and realization of the microwave cavity in the PTB caesium atomic fountain clock CSF1, IEEE Transactions Ultrason. Ferroel. Freq. Control **49** (3), p. 383-392 (2002).

[4] F. Pereira Dos Santos, H. Marion, S. Bize, Y. Sortais, A. Clairon, Controlling the cold collision shift in high precision atomic interferometry, Phys. Rev. Lett. **89** (23), 23304 (2002).

CORRESPONDING AUTHOR: S. Weyers, Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany, phone number: +49 (0)531 592 4416, fax number: +49 (0)531 592 4479, email: Stefan.Weyers@ptb.de

SMALL RF FILTERS MADE WITH THIN-FILM HIGH TEMPERATURE SUPERCONDUCTORS

S. Zhgoon, A. Barinov. O. Shteynberg, P. Mozhaev[1]
Moscow Power Engineering Institute, 14, Krasnokazarmennaja, Moscow, 111250, Russia. zhgoon@ieee.org
[1]Institute of Physics and Technology RAS, 34, Nakhimovsky, 117218, Moscow, Russia

ABSTRACT

We describe modeling and experimental results related to different variants of lumped element filters suitable for 0.5-3 GHz frequency range based on spiral inductors in combination with planar capacitors and segments of transmission lines. The inductive elements with conductor width of several tens of microns can have Q-factor exceeding 2000-3000 or even better depending on the film quality and thickness. These elements together with different designs of lumped capacitors can form a basis of high performance RF filters.

1. INTRODUCTION

Very low surface resistance of thin-film high temperature superconductors (HTSC) at 0.5-3 GHz allows making highly selective filters with very small dimensions. The filters can consist of distributed or lumped elements as well as of their combinations. Narrow but low loss HTSC strips can be quite long and they require folding in order to preserve small size. However the coupling between distributed elements limits the size reduction possibilities. Lumped elements often present more options to confine electromagnetic fields and therefore such structures can provide denser placement [1-5].

Multiple turns planar coils take little space on the substrate surface and present sufficient inductance values. The electromagnetic field of a multiple turns coil is relatively well confined around it and this feature allows reasonably close spacing of the sections of a multiple poles filter. Miniature resonators are built with such coils that can have Q>1000 at about 1 GHz [3-5].

A serious limitation of HTSC materials is that single layer designs are often preferable due to frequent degradation of the film during the double side deposition on a substrate. Another concern is related to difficulties of preparation of bridge structures connecting inner contact of a coil to an outer circuit. Thus ordinary filter configurations such as ladder or bridge type are not easily realized. We have found several approaches for building wideband and narrowband filters with good properties and we describe the layouts and the properties of such filters in detail. Our modeling results are in reasonable agreement with the experimental data.

2. PREPARATION TECHNIQUES AND THEIR LIMITATIONS

HTSC thin films for microwave and RF applications most often consist of $YBa_2Cu_3O_{7-x}$ (YBCO), however instead of Y almost all rare earths metals can be used with relatively little impact on processing and properties of the films. Bi, Tl and Hg based compositions with superior critical temperature (T_c) are less widely studied as candidates for microwaves, however they have to some extent similar properties concerning sample preparation.

As a rule in order to grow low microwave surface resistance HTSC films the choice of substrate is limited to materials with proper crystallographic structure and dielectric properties such as LaAlO3, $NdGaO_3$ and some others [6]. These substrates are presently available in sizes up to 50 mm. Another limitation is that relatively high temperature has to be used during HTSC film growth (about 600 C and higher) and very little temperature variation (several degrees C) over the wafer area can be tolerated in order to obtain reasonable deviation of HTSC film properties.

Most laboratory-scale processes of HTSC epitaxial films deposition are based on single side deposition (the double side large area deposition although reported [7] being not widely available to date) and on very small deposition area: quite often the substrate size does not exceed 1 cm^2. Besides the robustness of the film against water and acids is very low, while oxygen is easily released at elevated temperatures (>300 C). Any additional processing, such as deposition of films under and on top of HTSC films (especially on YBCO) irreversibly damage their properties by solid state chemical reaction even at room temperature quite fast (in several days or hours). This is the reason why this technology finds way to applications very slowly and requires very special approaches to the design strategies.

The most established fabrication process uses single side YBCO deposition with film thickness under 300-500 nm with subsequent wet or dry etch patterning (with conventional photoresist) and silver/gold contact pads formation. A selective etch for gold or silver (I-KI water solution) fortunately exists that does not attack YBCO and other HTSC compositions [8, 9]. We prefer to use for etching YBCO films a water free solution of concentrated HCl 5% of volume and H_3PO_4 5% of

volume in glycerol. In order to reduce undercut we add NH_4F into this solution sequentially by little amounts and check the etching speed of dummy samples. The etch is ready for use when the film of required thickness is completely removed in about 30 seconds.

The procedure of noble metal deposition on contact pads does not permit the use of adhesion layers (they slowly react with YBCO later) and therefore usually requires techniques for adhesion improvement and oxygen content restoration. The latter procedure is performed by annealing in flowing oxygen at more than 300 C. Additional dielectric layers on top of HTSC films can in principle be formed [10], however their choice is limited and if another HTSC layer is required the processing is led at high temperature (over 600 C) [11]. Such processing inevitably affects the lower layers and that is the reason why single layer microwave devices without ground plane are most often preferred in present designs.

3. PROPERTIES OF PLANAR COILS

In order to test different films properties in about 1 GHz we use an originally designed resonator with a single layer spiral HTSC coil that has capacitive coupling to input and output lines. The resulting resonator equivalent circuit is shown in Fig. 1.

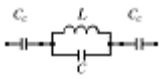


Fig. 1. Equivalent circuit of a single resonator

The corresponding layout variant in Fig. 2 has maximum size less than 4 mm in order to fit on the surface of a 5*5*0.5 mm^3 $NdGaO_3$ substrate. This design provides the resonance frequency below 1 GHz on this substrate.

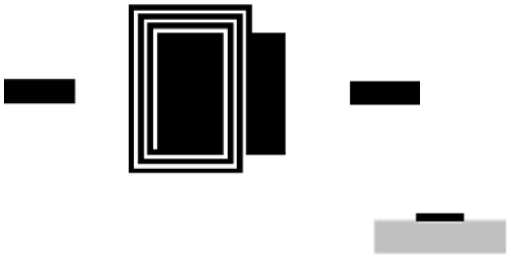


Fig. 2. Single layer resonator layout

With this designs electromagnetic properties of HTSC films are estimated from resonators parameters and compared to their other physical properties related to the preparation process and to the choice of substrates. Fig. 3 shows a typical data processing result obtained from measurement of the resonator frequency response versus temperature. This film had a critical temperature of about 85 K and a relatively good quality in terms of crystalline properties. The limitation of the Q factor at the level of 1500 is most probably due to $NdGaO_3$ substrate properties in the present case.

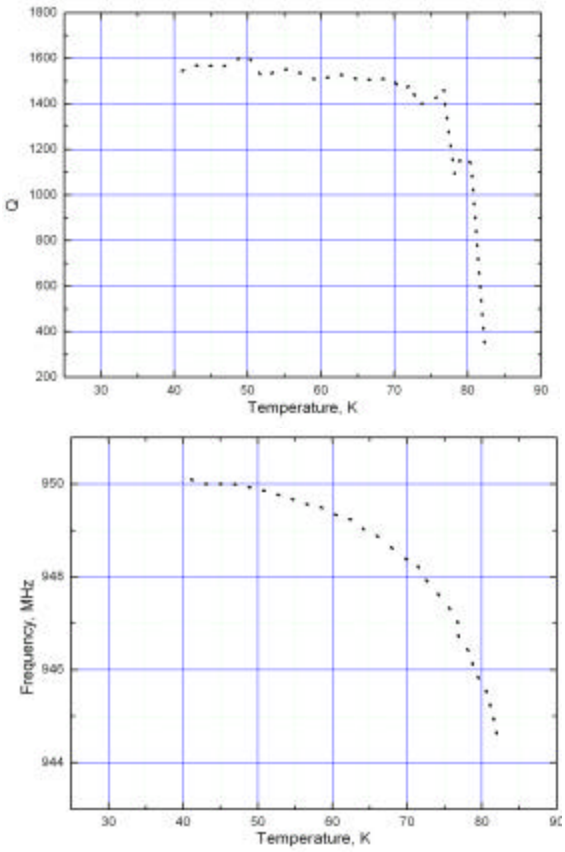


Fig. 3. Q-factor and central frequency behavior of a planar coil based resonator

The behavior of the resonator frequency reflects the inherent properties of superconductors. This change of about 0.5-1% is due to kinetic inductance together with the change of dielectric properties of the substrate and the dimensions of the structure. As for HTSC devices one would expect relatively high working temperatures (close to 70-77 K) special care should be taken for efficient temperature stabilization of the device.

4. FILTER DESIGNS

As was mentioned above there are many interesting designs of HTS lumped elements filters, which were offered by different researchers. However only a small number of published designs include spiral inductors. In fact the spiral coil gives the largest inductance value and the lowest resonator frequency of a lumped element inductor when the inner and the outer ends of the coil are connected to a lumped element capacitor. All distributed element types of resonators require larger substrate area and present more problems to close

placing because of the wider electromagnetic field distribution. The challenge of the filter design with spiral multi-turn planar coils is related in overcoming the difficulties of connection to the spiral inner end. In the printed circuits technology the most evident type of connections of the inner end of the coil to the rest of the structure is implemented with different kinds of bridges. Known examples [1,2,12-15] include deposition of thin-film conductive bridges over dielectric insulation (or an air gap) in multi-layer structures, formation of bonded wire bridges, or fabrication of conductive via bridges through the substrate. All these types of connections are ordinary to the normal metal technology.

However realization of the bridge connection between two high temperature superconducting pads is complicated due to technological difficulties.

In this paper we discuss 3 design concepts of filters, which evolved in attempts to step away from a complicated fabrication procedure to a simplified one.

4.1 Design with wire bridges

The first filter design has been realized on a substrate with both sides coated by superconducting thin films [3]. The equivalent circuit of third order filter is shown in Fig.4. We have chosen as the inductive element of the resonator a 3 turn 20 nH planar coil with very narrow 30 μm wide conductors.

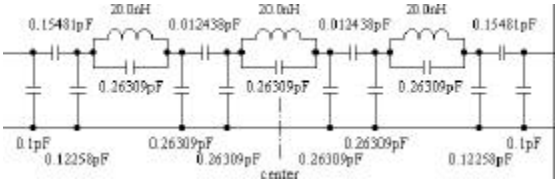


Fig. 4. Equivalent circuit of a passband filter

The connection between inner ends of inductive coils with the remaining circuit has been fabricated by using ultrasonic bonding of 40-microns gold wires.

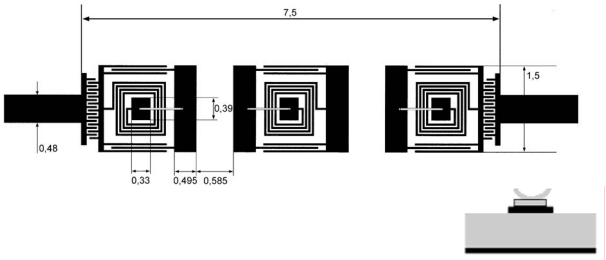


Fig. 5. Three-pole filter design

These wires were bonded to the inner pads of the coil covered by a silver film (annealed at 300 C in flowing oxygen) in order to ensure good adhesion for wire bonding and low conductivity loss in contacts. The use of normal metal coating in the region of wide contact pads and short wires has resulted in relatively high value of single resonator's unloaded quality factor (about 1000) that could be further increased by duplicating the number of gold wires. The application of normal metal ribbons could in principle further reduce the influence of this critical element. This filter design provides the smallest area the possible occupied by the multiple turns inductive element with a small inner contact pad area, while interdigital capacitive elements give a reliable way to obtain sufficiently large capacitors with high enough Q.A filter layout for narrow band operation is shown in Fig. 5.

The measured response of this filter is shown in Fig. 6. The filter response corresponds quite well to modeling results (Fig. 7) and the filter size is among the smallest achievable with lumped element techniques for a strip width wider that 50 microns.

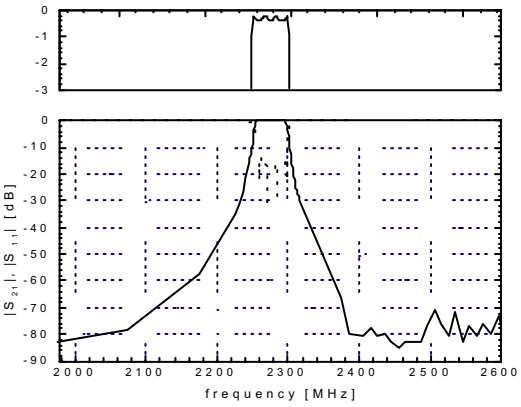


Fig. 6. Measured frequency response of realized lumped element filter ($|S_{21}|$ - solid line, $|S_{11}|$ - dashed line) at 20K

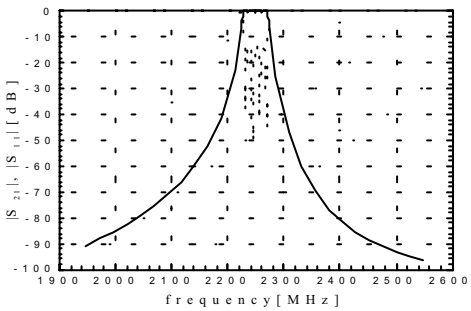


Fig.7. Calculated frequency response of the lumped element filter ($|S_{21}|$ - solid line, $|S_{11}|$ - dashed line)

4.2 Design without bridges

In the second filter design we have suggested a concept that avoids bridge connections while preserving the advantages of planar spiral coils [4,5]. In this filter wide inner and wide outer pads form relatively large

capacitors together with the ground plane thus resulting in a small area occupied by each resonator. The equivalent circuit of a 4pole filter for this design is shown in Fig. 8, and the layout variant is shown in Fig. 9.

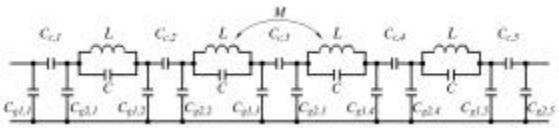


Fig. 8. Equivalent circuit of a passband filter

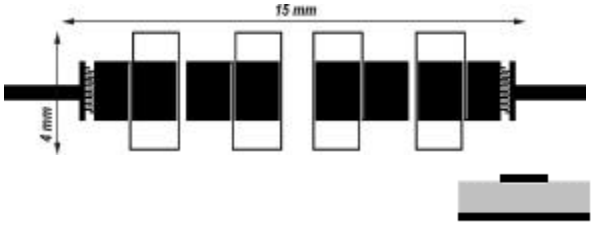


Fig. 9. Layout of the four-pole lumped element bandpass filter

Another variant of such filter that presents more flexibility for response synthesis is shown in Fig. 10. The shift of the sections allows reducing the coupling of resonators in this area.

However as the size of the coil itself is usually larger than the substrate thickness, the influence of the ground plane conductivity on the resonator Q-factor is significant, that is why a thin-film superconducting ground plane is still preferred for both of these designs (for the previously discussed design and for this one).

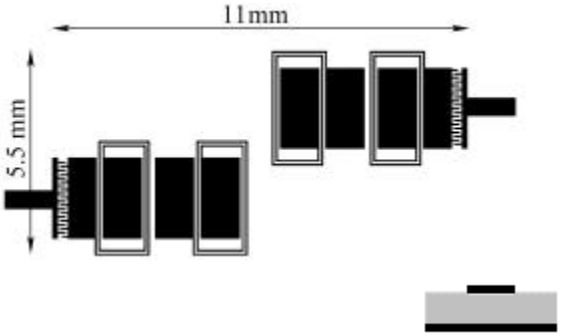


Fig. 10 A design variant with reduced coupling between central sections.

We have realized all these filters (the first and the second types) with Al ground planes and we have obtained a quite reasonable filter response, shown in Fig. 11,13.

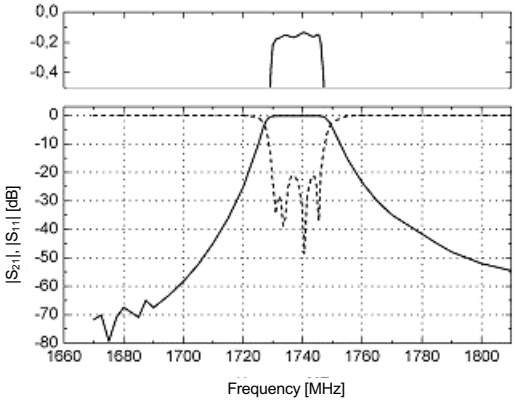


Fig. 11. Measured 4-poles 1-turn filter frequency response with Al ground plane

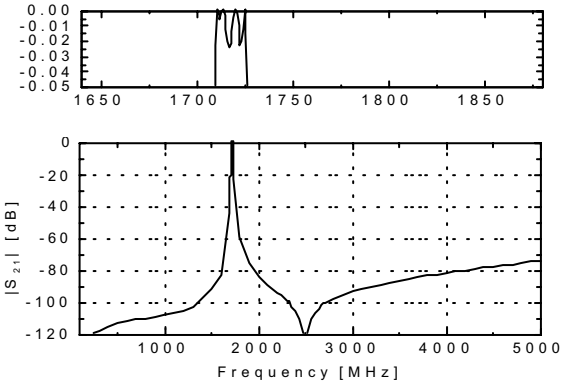


Fig.12. Calculated frequency response of the lumped element filter $1^3/_4$ turns.

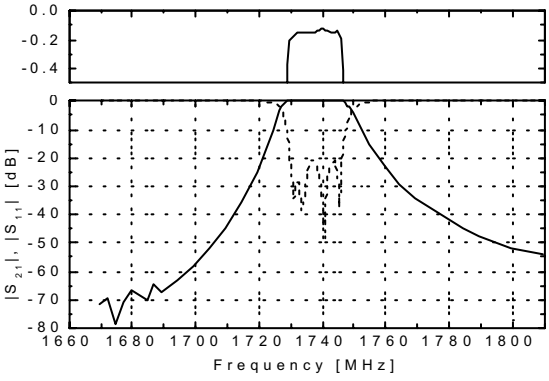


Fig.13. Measured frequency response of realized lumped element filter $1^3/_4$ turns at 20 K temperature (|S21| - solid line, |S11| - dash line).

It is well known, that the growth of HTSC films on both sides of a substrate and further processing significantly increase the difficulties of filter fabrication and quite often the HTSC film on the backside of the substrate degrades by the end of preparation operations.

4.3 Design without bridges and ground plane

In the third design concept we tried to avoid this weak point of the previous filters designs. Now we present a new variant of filter design with 2? turns planar coils as inductance elements, which do not require a ground plane (Fig. 14).

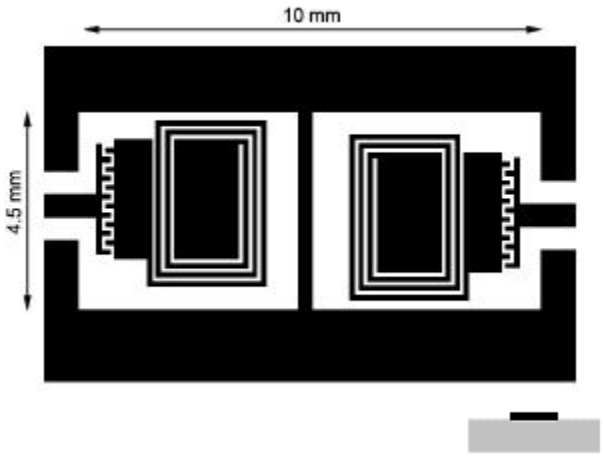


Fig. 14. Layout of the two-pole lumped element bandpass filter realized in one layer.

Filters fabricated with this design do not require substrates with both sides coated by superconducting thin films. The whole structure of this filter is located in a single HTS layer. Obviously these types of design require more sophisticated software and more skilled designers, however this is a further very important simplification of the technological process that increases the reliability of fabrication and repeatability of filter characteristics.

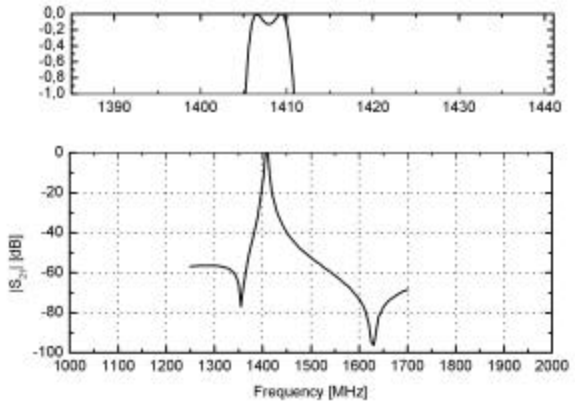


Fig. 15. Modeled frequency response of the lumped element filter with 2? turns coils.

The modeled response of this filter is shown in Fig. 15. As a rule connections to outside circuits require bonding or soldering (or gluing with silver paint) and in all cases the HTSC film has to be covered with a contact metal layer of Ag or Au/Ag. In this case coated superconductor kinetic inductance gets short-circuited by metal layers at RF and this part of layout behaves almost as normal metal and not as a superconductor. The simulation of the same design with contact elements partly consisting of normal metal (Ag) at 20 K is presented in Fig. 16.

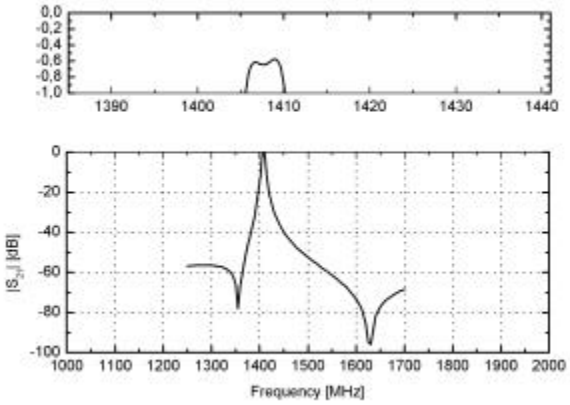


Fig.16. Simulation of the same design with contact elements partly consisting of normal metal (Ag) at 20 K frequency response of the lumped element filter $2^3/_4$ turns.

As could be expected some loss (0.5 dB) is present however it is relatively small because main elements – the turns of the coils are still superconducting. This loss can be further improved by removing metal from unnecessary regions and leaving it only at the contact pads regions.

5. DISCUSSION

Most of filter designs described in this study are narrowband. Nevertheless relatively wideband filters with steep slopes can also be made on the basis of planar coils.

One important feature has to be kept in mind, when designing and applying such filters. As soon as the dimensions decrease the width of the conductor in the spiral decreases. At the same transmitted power level the current density becomes higher. In high Q resonators the current in the coil may be quite strong and thus it may easily surpass the critical current density of the HTSC film. In this case the frequency response of a resonator behaves in a nonlinear way. At high power level the region around the resonance frequency flattens and Q drops to very small values. A filter response of the filter types described above drops in the vicinity of

the central frequency and the filter serves as a power-limiting element.
While this feature is acceptable and sometimes useful in receiver front ends, it may not be tolerated in more powerful circuits, and this is one of the limitations of the spiral coils designs in comparison to larger distributed elements designs.

Another major limitation is cryocooling. As these filters require cooling to at least 77 K in order to become operative the final dimensions of the system may not drop below several tens of centiliters and the power consumption below Watts level. This limits the application area to either relatively large moving systems, or stationary systems or to systems where this temperature is obtained just because it is ambient in some places as in the shadow of a satellite in space.

The operation costs are also relatively high as the excessively multiple cooling-heating cycles with a temperature difference of 200 C are slowly damaging most of ordinary materials forming the hermetic package as well as the filter itself due to the difference in expansions of the HTSC film and of the substrate. This fact means that keeping a device cool for all its working life is a good but expensive way to ensure the correct long-term operation.

6. CONCLUSIONS

Due to inevitable use of cryocooling such filters do not compete with mass-production solutions based on SAW, on FBARs, or on dielectric resonators. However they may occupy a relatively narrow niche of exclusively low loss devices in exceptionally demanding applications, such as in astronomy, base stations for mobile communications and some others.

7. ACKNOWLEDGEMENTS

This work is supported by INTAS, grant 01-0249.

REFERENCES

[1] M. Reppel, H. Chaloupka, S. Kolesov. Electron. Lett. 34 (1998) pp. 929-930.
[2] M. J. Lancaster, F. Huang, A. Porch, B Avenhaus, J. S. Hong, D. Hung. IEEE trans. on MTT v.44 no.7 (1993) pp.1339-1346.
[3] A.E. Barinov, S.A. Zhgoon, V.A. Sukhov, H.T. Su, M.J. Lancaster, F. Huang. Superconducting lumped element filter with spiral inductors, Microwave and Optical Technology Letters, Vol.29, Iss.2, 20 Apr 2001, pp.94-95.
[4] A.E. Barinov, S.A. Zhgoon, V.A. Sukhov. Planar superconducting lumped element bandpass filter with spiral inductors, Physica C: Superconductivity and its Applications, 2001 Jun, Vol.355/3-4, pp.257-259.
[5] A.E. Barinov and S.A. Zhgoon. Planar superconducting lumped element bandpass filter with spiral inductors, Supercond. Sci. Technol. 15 (2002), 21 May 2002, Insitute of Physics Publishing, pp.1040-1042.
[6] E.K.Hollmann, O.G.Vendik, A.G.Zaitsev, B.T.Melekh, Substrates for high-T_c superconductor microwave integrated circuits, Supercond. Sci. Technol., v. 7, No 9, pp. 609-622 (1994)
[7] P. Bodin, J.L. Skov, A. Kuhle, M. Hagensen, T. Clausen, I. Rasmussen, S. Hjorth, J. Bindslev Hansen, Double-sided $YBa_2Cu_3O_{7-x}$ thin films for microwave applications, Supercond. Sci. Technol., v. 7, iss. 10, pp. 717-720 (1994)
[8] W.Eidelloth, R.L.Sandstrom. Wet etching of gold films compatible with high Tc superconducting films. Appl. Phys. Lett. Vol. 59 (13), pp. 1632-1634, 1991.
[9] S. Zhgoon, V. Yakunin, V. Ananchenko, J.Joffrin, J.-Y. Prieur. Processing of HTcS Antenna Microbolometers. OE-LASE 94, SPIE Proceedings, Vol. 2159 pp.180-185, Los Angeles, CA, USA.
[10] S.A.Zhgoon, G.D.Lobov, A.R.Kuzhakhmetov, V.A.Suhov, Yu.V.Kislinsky, E.A.Stepantzov. Some Features of Phase Locking up to 2,5 THz in Grain Boundary Josephson Junctions at 77 K. Proceedings of EUCAS-95, pp.1311-1314.
[11] F.C. Wellstood, J.J. Kingston, J. Clarke, Thin-film multilayer interconnect technology for $YBa_2Cu_3O_{7-x}$, J. Appl. Phys., v. 75, iss. 2, pp. 683-702 (1994).
[12] D. Zhang, G.-C. Liang, C. F. Shih, M. E. Johansson, R. S. Withers. IEEE trans. on MTT v.43 no.12, pp. 3030-3036, 1995.
[13] K. Siato, N. Sakakibara, Y. Ueno, Y. Kobayashi, D. Yamaguchi, K. Sato, T. Mimura. IEICE trans. Electron. E83-C no.1, pp. 15-19, 2000.
[14] C.K. Ong, L. Chen, J. Lu, C.Y. Tan, B.T.G.Tan. IEEE Microwave and Guided Wave Letters v.9, no.10, pp. 407-409, 1999.
[15] D. G. Swanson, R. Forse, B. J. L. Nilsson. IEEE MTT-S Int. Microwave Symp. Dig., pp. 192-1193, 1992.

Corresponding author: Dr. Sergei Zhgoon
Moscow Power Engineering Institute,
14, Krasnokazarmennaja, Moscow, 111250, Russia.
Tel 7(095)362 7212, Fax 7(095)362 8938
e-mail zhgoon@ieee.org

CONTROLLABLE TRIMMING OF SAW FILTERS AND RESONATORS ON LANGASITE AND LITHIUM TANTALATE

S. A. Zhgoon, A. S. Shvetsov
Moscow Power Engineering Institute, 14, Krasnokazarmennaja, Moscow, 111250, Russia. zhgoon@ieee.org

ABSTRACT

We have studied the magnetron reactive sputtering of Yb_2O_3 as concept feasibility verification process, for controllable trimming of SAW devices and we have proven that dense oxides may be quite efficient for fast and repeatable processing of SAW (LSAW) resonators and filters. Our technological equipment is provided with capabilities of frequency response acquisition *in situ* during the deposition of the trimming film and continuous response monitoring is carried out until the response reaches desired shape specifications.

In this paper we extend the choice of substrates by including experiments with trimming of relatively low frequency LGS resonators and we provide recently obtained data on ageing during storage. We discuss the strategies of film deposition and result monitoring for individual trimming of already connected and coverless packaged devices taking into account the heating during processing. We discuss the influence of this procedure on temperature stability of the devices and on power handling capability.

This process seems to have all features necessary for a promising candidate in attempts to introduce on-wafer controllable trimming into volume production of SAW devices.

1. INTRODUCTION

The yield of SAW devices depends on processing conditions and in most cases of device failure the frequency specifications are not met due to substrate, film thickness or metallization ratio inhomogeneity over the wafer area. We describe an efficient process that allows decreasing the excessively high device working frequency with possibilities to individually monitor the trimming results during the process. This process is based on deposition of ytterbium oxide (Yb_2O_3) dielectric thin film on ready devices.

SAW and LSAW velocities are altered by deposition of thin-film coatings on the substrate surface. Some coatings may efficiently increase the SAW velocity in case of Rayleigh waves, however application of such coatings to trimming of LSAW based devices is very limited due to loss increase related to acoustic fields spreading in this case. When SAW or LSAW velocity decreases the loss related to energy leakage is generally not high and the loss in the layer material becomes more meaningful. That is why proper choice of such materials is important.

Using thin layers of gold for SAW resonators trimming has been presented in [1]. The application of this established for BAW devices process is limited in the case of SAW devices to relatively small frequency shifts, defined by coalescence of isolated islands during the film thickness growth and the formation of conducting paths. On the other hand, due to the high density of some materials, notably of langasite and of lithium tantalate, the choice of dielectric materials that provide noticeable velocity decrease by deposition of very thin layers is limited. One has to choose among oxides, nitrides, carbides and some other chemical compounds of rather high atomic mass elements. It is preferable that these compounds are denser than the substrates in order to provide efficient mass loading of the surface.

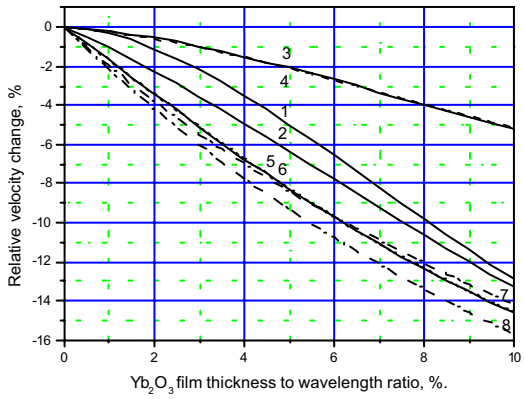


Fig. 1. Modeling of SAW velocity dependence on Yb_2O_3 film thickness to wavelength ratio. The curves are numbered in accordance with the substrates: 1 – LT-42 open surface, 2 – short circuited surface; 3 – LGS open, 4 – short; 5 – ST-quartz open, 6 – short; 7 – LN-128 open, 8 – short.

Dielectric oxides are materials that are among the easiest to deposit. As a rule they present good dielectric properties, and quite often their acoustic properties are good enough for trimming application. The selection of oxides with densities higher than 9 g/cm^3 is restricted. We find in the list of such oxides tantalum and tungsten oxides, ytterbium, lutetium and hafnium oxides and

some more exotic materials that fall into this category. Tungsten and tantalum oxides can form several different compositions with different oxygen content and this may influence the stability of the process, lutetium is very expensive, hafnium was not available and that is why we have stopped on Yb_2O_3 [2]. The Yb_2O_3 film slows down the velocity on lithium tantalate very efficiently. The films thickness required for similar trimming results is only about 2-3 times higher than the required thickness of gold or platinum. This choice has shown to be representative and successful, and that is why we continue to discuss it in this contribution. However, we suppose that other prospective materials that did not enter in our brief list above deserve careful investigation for different processing and performance benefits that they may bring.

Modeling of SAW propagation on the surface of several most important substrates covered with Yb_2O_3 thin film (elastic data taken from [3]) is presented in Fig. 1. The cases of the film deposition on an electrically open surface and on electrically short-circuited surface are considered for each of the 4 substrates: lithium tantalate with orientation yxl/42° (LT-42), langasite with orientation yxlt/48.5°/26.6° (LGS), quartz with orientation yxl/42.75° (ST-quartz) and lithium niobate with orientation yxl/128° (LN-128). As it could be expected, the trimming is more efficient for substrates with higher SAW velocity (LN-128 and LT-42) or with lower density (ST-quartz). The velocity change on LGS is about 2-3 times less pronounced than on LT-42 and about 3-5 times less pronounced than on LN-128 and ST-quartz. However our estimation of Yb_2O_3 mechanical parameters is based on the data from [3] and it may differ from the parameters of a real Yb_2O_3 thin film. That is one of the reasons why experimental verification of trimming possibility on LGS was undertaken in this work.

2. EXPERIMENTAL

The DC magnetron sputtering of ytterbium oxide from an Yb metal target was carried out in a mixture of Ar with about 1% of oxygen. Transparent non-conductive film formed with deposition rates 0.3-1.0 microns per hour. This film has good adhesion to the $LiTaO_3$ and to the langasite substrates and it also has good adhesion to the Al film. The process was reproducible and very stable.



Fig.2. Lidless packaged LGS resonator on a holder ready for trimming.

The device under test was packaged and mounted lidless on a holder with SMA-type connectors. The holder was placed at a distance of 8-10 cm from the target, in a way that the open substrate surface with electrodes faced the target.

Our vacuum chamber was fitted with two hermetic RF connectors that allowed connecting to a two-terminal SAW device. The S_{11} or S_{21} response was registered continuously with an HP8505 network analyzer during the deposition. Although the device and the transducers electrodes were to some extent immersed in plasma, the concentration (and consequently the conductivity) of the latter was small enough to influence the response in 50-Ohm track very little. No impedance matching elements were used in our present tests. The change in the device response during plasma on and off states was in the range of a tenth of a dB.

Initially we tended to disconnect the leads from the network analyzed in order to preserve it from damage, when we were switching the plasma on. However we finally understood that the voltage induced by plasma in our particular arrangement into the RF circuit was negligible and our model of network analyzer would not be damaged by it. This feature allowed us to keep the response monitoring and registering continuously and also to observe continuous changes in the response without losing meaningful details. In order to avoid the equipment damage we would not recommend other researchers to rely on our experience and to connect any type of network analyzers to any device in any other arrangement without prior careful verification of plasma induced voltages by proper means with proper equipment.

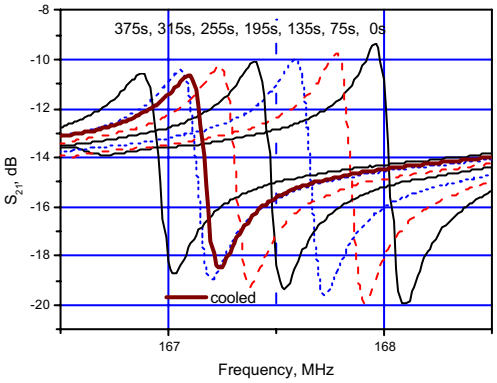


Fig. 3. Gradual change of an LGS resonator frequency response during continuous trimming.

A typical behavior during the continuous trimming process is presented in Fig. 3, where a relatively low frequency single-port LGS resonator was series connected between two ports of a network analyzer. The response recordings (thin lines) were taken each minute

(except the first trimming stage duration of 75 seconds). The total of 6 trimming stages were carried out. Then the sample was left to cool down back to room temperature. The thick line in Fig. 3 corresponds to the final state after cooling down to room temperature.

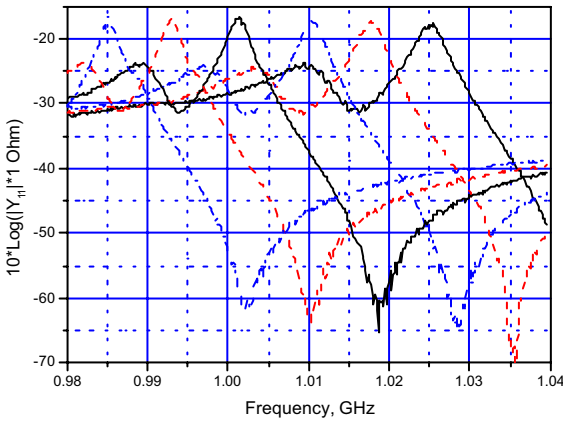


Fig. 4. Gradual change of an LT-42 resonator frequency response during trimming.

The same type of process, with about the same deposition rate has been applied to a resonator on LT42. The behavior of the LT-42 resonator response during the trimming process is shown in Fig. 4. Recordings are taken after cooling down intervals following each 60-seconds deposition run. More detailed data for this process was published in [2]. If we compare the effect of trimming in Fig. 3 and 4 the relative fractional frequency change for LGS (in cooled down state) is about 0.08% and for LT-42 it is about 0.8%. If we take into account the wavelengths ratio (about 4) – we obtain that the trimming efficiency for LGS is about 2.5 times smaller than that for LT-42, this ratio is quite close to our modeling results.

3. INFLUENCE OF SUBSTRATE HEATING

As the deposition rate for the LGS sample was rather high strong heating of the substrate was observed. Depending on heat transfer between the substrate and the sample holder the temperature could in principle rise up to about 150-180°C in a continuous deposition process. This temperature change has resulted in additional decrease of the observed response central frequency and thus obscured to some extent the monitoring of the trimming process. As shown in Fig.3, the heating of the substrate in the case of LGS has resulted in about 0.2 MHz of additional frequency shift that has to be extracted from the total shift (including heating effect) of about 1 MHz. So the heating in the present case was creating an about 20% additional frequency shift during processing.

The influence of heating on the frequency response shift is essentially a nonlinear effect. It accumulates with time and it is not only different for different substrates with different types of temperature coefficients (close to linear for LT-42 and LN-128 and parabolic for ST-quartz and LGS), it changes with sample mounting conditions, heat transfer, plasma torch shape and it is also dispersive as the series and parallel resonances react in a different way on substrate temperature change.

The above considerations mean that seemingly equidistant shifts in response recordings in Fig.3 after equal time intervals are in fact a combination of many factors, so that the prediction of exact frequency after trimming by the readings of monitored frequency response is not an easy task. In order to transfer this approach to volume production and to on-wafer controlled individual device trimming process several strategies could to be adopted for reduction of temperature influence on the final trimming results.

For relatively short trimming times for this particular sample and process we would probably estimate the final frequency shift to be about 10-20% lower than the monitored reading. And if the trimming accuracy of this range is acceptable – then the expected result should simply be corrected by this value. For example, for a particular type of samples a restricted number of runs with complete cooling of test structures could be used to give a relation between the monitored response and the response at a particular temperature after a particular time of trimming (or after achievement of particular frequency shifts). The average wafer (or sample holder) temperature should also be monitored in the case when many devices are subjected to the heat spreading from a processed sample or wafer region.

In this approach after establishing a particular trimming rate one can approximately trim each sample for certain time following to the pre-trimming wafer mapping without connecting to each sample during the trimming process. This deposition time-based strategy seems to be the easiest for wafer stage trimming process through a pre-programmed moving window because introduction of individual connectors to an on-wafer device under processing is a very complicated design challenge. However, as the processing is still targeted to be individual for each sample, such connectors, if ever designed, will be very beneficial for the reliability of final trimming results, as they will *in-situ* monitor any unwanted change in the process rate and produce an early warning of any failure.

For lithium tantalate and especially for lithium niobate the temperature coefficient is quite high and intermediate cooling steps could be relatively efficient as in Fig. 4, where after each 60 seconds of processing

the sample was allowed to cool down for 3 minutes before the response was recorded.

Efficient heat transfer sample holder design should facilitate the cooling and speed up the process in this strategy. Water-cooled (fixed temperature) wafer holder with means for reliable wafer fixation with heat-transfer grease (or other means) could probably release this problem completely.

Careful temperature monitoring of processed regions or samples is a good way for temperature effect prediction. It is not a trivial technique either because the required accuracy of data is high, while it is difficult to place temperature sensors in identical conditions with the processed sample. When lidless packaged samples are trimmed instead of a temperature sensor one can keep a closed-lid sample near the processing position and extract the temperature related frequency shift from the result observed. Interchanging a number of such lidded samples simultaneously with changing of trimmed samples may reduce the cooling stage time for these "temperature sensors".

Finally, the heating of the processed sample can be substantially reduced by modifying of the magnetron gun design, introducing a larger distance from the target to the sample and by careful filtering of higher energy particles that re-sputter the substrate and the growing film without contributing to the film growth. In case of oxides sputtering bombardment by neutral oxygen atoms is often the case and as our experience shows facing targets magnetrons arrangements and other known solutions are quite efficient in these goals. These facts mean that further improvement of the process seems feasible.

4. AGEING

Initial data on ageing of trimmed samples is shown in Fig. 5. The sample was subjected to an excessively long processing run. The final frequency shift was about 7% and the corresponding ytterbium oxide film thickness estimate was in the range of 5% of the wavelength (or about 0.7 microns). So this example corresponds to an extreme case and ageing effects in more realistic trimming ranges are expected to be even less pronounced. The after-storage frequency shift for a lidless sample shown in Fig. 5 is less then 0.1% of the central frequency.

While this value is in the range of our measurement accuracy for this particular experiment (different cables used, room temperature not monitored) it still can be estimated as being below 1.5% of the total trimming value (about 7% of the central frequency), that seems to be reasonably small for many SAW applications. For a more realistic trimming range of about 1% the ageing related frequency shift is expected to be lower than about 150 ppm. The latter figure is comparable to frequency shift related to a 4°C temperature decrease on this substrate (with a realistic temperature specifications range of about –30°C to +85°C) and it can certainly be tolerated in RF-filters.

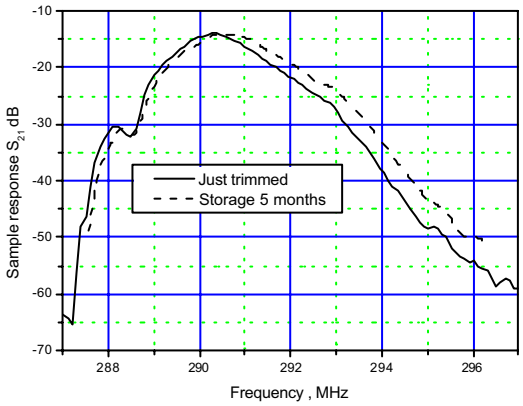


Fig. 5. Response of a device on LT-42 with very thick Yb_2O_3 coating just after deposition and after 5-months of storage.

If the trimming is applied to devices with tightened requirements to frequency stability (based on ST-quartz, for example) one has to be more careful about changing material properties too much with increased thickness of additional coatings. However these devices are narrowband as a rule, and they use relatively low Al thickness, so that the trimming range required might be much smaller – about 0.1% or similar. In this case the ageing and other Yb_2O_3 film related drawbacks are reduced in accordance with the film thickness decrease.

We would also speculate that accelerated ageing processing could be applied as well as the deposition regime optimization in order to reduce these effects.

5. DISCUSSION ON POWER HANDLING AND TEMPERATURE STABILITY

Following our past experience the deposition of different oxide films on top of Al electrodes in devices on LT-42 as a rule increased the speed of the electrodes damage (often referred as acoustical migration, or as stress migration, for example [4]) in our power handling tests. The reason for this behavior seems to be in the presence of additional stress induced by such films on electrodes surface. A film with 1% relative thickness required for about 1% frequency trimming should already be quite perceptible in this sense. We would expect that the power handling capability might deteriorate with increasing the amount of frequency trimming.

The temperature behavior of thin-film covered SAW devices depends strongly on the layer material elastic constants and their temperature behavior. Among known materials Al and is the best studied and we can expect that some useful estimates for Yb_2O_3 behavior could be deduced from the comparison with this material.

The data available for temperature behavior of Young modulus and Poisson ratio [3] are given for ceramic material with some porosity, this means that the estimates based on these data should be used only as a zero order approximation. Our modeling of SAW velocity temperature behavior versus film thickness has shown that the Yb_2O_3 film changes the TCD on LT-42 and on LN-128 significantly less than the Al film of the same thickness and the TCD can both increases and decreases. Therefore we can expect that the trimming influence on the temperature behavior is quite weak on these substrates. The temperature behavior of SAW devices on LGS and on ST-quartz changes significantly with the deposition of an Yb_2O_3 film. According to our estimation a frequency shift by 1% resulting from trimming shifts down the turnover temperature by 40°C for both substrates. For comparison aluminum electrodes producing this amount of turnover temperature shift will have the thickness equal to about 2.5% of the wavelength for LGS and about 1.5% of the wavelength for ST-quartz.

6. CONCLUSIONS

This process seems to have many features necessary for a promising candidate in attempts to introduce on-wafer individual controllable trimming into large-scale production of SAW devices.

7. ACKNOWLEDGEMENTS

Clarisay Inc. (Dallas, TX) has supported the initial stages of this work. Fomos-Technology (Moscow, Russia) has provided high quality 4" LGS wafers and SAW resonators on LGS for this work. The use of Professor K. Hashimoto (Chiba University, Japan) free software is gratefully appreciated.

REFERENCES

[1] S.M. Wang et al., "In-Situ Frequency Trimming of SAW Resonator Using Conventional Crystal Resonator Fine Tuning Method with Gold Thin film Addition" Proc. IEEE Ultrason. Symp. 2003, pp. 1730-1733.

[2] S. Zhgoon, A. Shvetsov, O. Shteynberg, K. Bhattacharjee, J. Flowers. Approach to On-wafer Controllable Trimming of SAW Filters. Proc. IEEE Ultrason. Symp. 2004.

[3] Physical-chemical properties of oxides, Moscow, “Metallurgiya”, 1978 (in Russian).

[4] M. Hikita, N. Shibagaki, K. Asai, K. Sakiyama, A. Sumioka. New Miniature SAW Duplexer Used in GHz Band Digital Mobile Cellular Radios. Proc. IEEE Ultrason. Symp. 1995, p.33.

Corresponding author: Dr. Sergei Zhgoon
Moscow Power Engineering Institute,
14, Krasnokazarmennaja, Moscow, 111250, Russia.
zhgoon@ieee.org

STATISTICALLY LIMITING ERRORS OF PASSIVE REMOTE SAW SENSING EMPLOYING DIFFERENTIAL PHASE MEASUREMENTS

Yuriy S. Shmaliy and Oscar Ibarra-Manzano
FIMEE, Guanajuato University, Salamanca, Gto, 36730, Mexico, shmaliy@salamanca.ugto.mx

ABSTRACT

A statistical analysis is given of the limiting estimate errors of the time delay (phase difference) between two reflectors of the passive wireless surface acoustic wave (SAW) sensor with single differential phase measurement (DPM). Rigorous and approximate solutions for the mean and mean square errors along with the error probability for the estimate to exceed a threshold are studied in detail. The errors are estimated in a wide range of signal-to-noise ratios and thresholds. Practical recommendations for designers of the remote SAW sensor interrogating systems with DPM are also given.

1. INTRODUCTION

Passive remote wireless surface acoustic wave (SAW) sensing has gained currency during a couple of decades to measure temperature, force (pressure, torque, and stress), velocity, direction of motion, *etc.* with a resolution of about 1% [1]. Several passive wireless SAW devices were recently manufactured to measure temperature [1], identify the railway vehicle at high speed [2], and pressure and torque [3].

The information bearer in such sensors is primarily the time delay of the SAW. Therefore, most passive SAW sensors are designed as reflective delay lines, which basic operation principle is sketched in Fig. 1. The interrogator transmits the radio frequency (RF) pulse that is received by the antenna of the SAW sensor at some time instant t_0. The interdigital transducer (IDT) converts the electric signal to SAW and about a half of its energy distributes to the reflector. The SAW propagates on the piezoelectric crystal surface with a velocity v through double distances ($2L_1$ and $2L_2$), attenuates (6 dB per μs delay time [3]), reflects partly from the reflectors (R_1 and R_2), and comes back to IDT. Inherently, the SAW phase delays on the piezoelectric surface. Afterwards, the SAW is reconverted by IDT to the electric signal, and retransmitted to the interrogator. Finally, two noisy RF pulses appear at the coherent receiver at the moments t_1 and t_2, in which an inter distance time delay $\Delta\tau_{21} = 2(L_2 - L_1)/v$ bears information about the measured quantity.

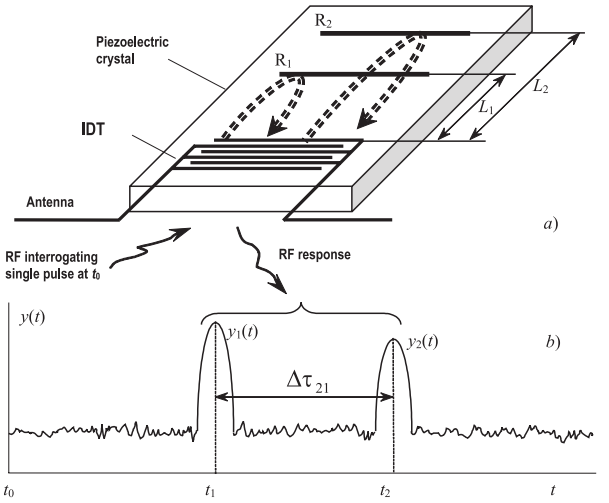


Figure 1: Operational principle of remote SAW sensing with DPM: (a) basic design of passive SAW sensors, (b) two reflected RF pulses at the coherent receiver detector

The most widely used method to estimate $\Delta\tau_{21}$ is the differential phase measurement (DPM) at coherent receiver [4] employing the maximum likelihood function approach. Here the orthogonal components of the reference signal are formed, and the detector yields the inphase (I) and quadrature phase (Q) components of the received RF pulses. The phase of the RF pulse relative to the reference is then estimated by

$$\hat{\theta} = \arctan\frac{Q}{I} \tag{1}$$

to range from $-\pi/2$ to $\pi/2$ or by

$$\hat{\theta} = \begin{cases} \arctan(Q/I), & I \geq 0 \\ \arctan(Q/I) \pm \pi, & I < 0, \end{cases} \begin{cases} Q \geq 0 \\ Q < 0 \end{cases} \tag{2}$$

to lie between $-\pi$ and π. Finally, the phase difference between two pulses is calculated as

$$\hat{\Theta}_0 = \hat{\theta}_2 - \hat{\theta}_1. \tag{3}$$

To obtain the estimate (3) with a reasonable accuracy, the interrogating signal must bear a sufficient energy, obtaining a large enough value of the signal-to-noise ratio (SNR) in a RF pulse. In this paper we

bring a mathematical tool to evaluate errors of such sensors at receiver with single DPM.

2. STATISTICAL SIGNAL MODEL

The transmitted interrogating RF pulse very often has a near Gaussian waveform [1] and is modeled as a real pulse,

$$x(t) = \sqrt{2S}e^{-q^2(t-t_0)^2}\cos(2\pi f_0 t + \vartheta_0), \quad (4)$$

where $2S$ and θ_0 are its peak power and initial phase, respectively, f_0 is a carrier frequency, and q is a constant. Note that noise does not perturb (4) substantially at the SAW sensor.

At the receiver, the reflected RF pulse attains the form of

$$s_1(t) = \sqrt{2S}\alpha_1(t)e^{-q^2(t-t_1)^2}\cos[2\pi f_0 t + \vartheta_1(t)], \quad (5)$$

where $\alpha_1^2(t)$ is a normalized instantaneous energy of the RF pulse caused by its propagation. The phase shift relative to the carrier is $\vartheta_1(t) = \phi(t) - \psi_1 + \theta_0$, where $\psi_1 = 2\pi f_0(2L_1/v)$ and $\phi(t)$ is a phase shift caused by wave propagation, Doppler effect, frequency instability, *etc.* At the receiver, (5) is contaminated by an additive stationary Gaussian noise $n(t)$ with a known variance σ^2. A narrowband and mean zero, $\langle n(t)\rangle = 0$, model of this noise is

$$n(t) = A(t)\cos[2\pi f_0 t + \varphi(t)], \quad (6)$$

where $A(t) \geq 0$ is the random amplitude distributed by Rayleigh and $|\varphi(t)| \leq \pi$ is the uniformly distributed random phase mod 2π. A mixture

$$y_1(t) = s_1(t) + n(t) \quad (7)$$

represents a noisy signal at the receiver,

$$y_1(t) = V_1(t)\cos[2\pi f_0 t + \theta_1(t)], \quad (8)$$

in which $V(t) \geq 0$ is a positive valued random envelope distributed by Rice and $|\theta_1(t)| \leq \pi$ is a modulo 2π random phase. The second reflector produces the response $y_2(t)$ of the same performance (5) and (8), in which the subscript 1 is merely substituted by 2.

For a constant Doppler frequency, the actual phase difference between two reflectors R_2 and R_1,

$$\Theta_0 = \psi_1 - \psi_2 \quad (9)$$

$$= -4\pi f_0 \frac{L_2 - L_1}{v} = -2\pi f_0 \Delta\tau_{21}\ , \quad (10)$$

is estimated at coherent receiver via

$$\Theta(t) = \theta_2(t) - \theta_1(t). \quad (11)$$

as (3), using (1) or (2). It is important that at the ideal receiver (having uniform amplitude response and linear phase) distributions of (11) and its estimate (3) are equal.

3. DISTRIBUTION OF THE DIFFERENTIAL PHASE

Distribution of (11) for the fixed Θ_0 and peak-instantaneous SNRs in the pulses,

$$\gamma_1 = \frac{S\alpha_1^2(t_1)}{\sigma^2} \text{ and } \gamma_2 = \frac{S\alpha_2^2(t_2)}{\sigma^2}\ ,$$

were pioneered by Tsvetnov in [5] and [6]. In [7], we showed that these distributions are also valid for the linearly varying amplitude and phase. Because SNRs in the received pulses may or may not be equal, we distinguish below (and through out the paper) two cases plausible in passive remote SAW sensing.

3.1. Different SNRs in the RF pulses

Here we suppose that the SNR in the second reflected pulse is lower, $\gamma_2 < \gamma_1$, owing to the SAW attenuation with distance. The modified form of Tsvetnov's pdf for this case was offered in [8]; that is,

$$p(\Theta|\gamma_1, \gamma_2, \Theta_0)$$

$$= \frac{1}{4\pi}\int_{-\pi/2}^{\pi/2}(1 + 2\gamma_+ - E)\, e^{-E}\cos z dz, \quad (12)$$

where $E = \gamma_+ - \gamma_- \sin z - \lambda\cos z$, $\gamma_+ = (\gamma_1 + \gamma_2)/2$, $\gamma_- = (\gamma_2 - \gamma_1)/2$, $\lambda = \gamma_\times \cos(\Theta - \Theta_0)$, and $\gamma_\times = \sqrt{\gamma_1\gamma_2}$. For the computational purposes, Tsvetnov proposed in [6] to extend (12) to the Fourier series,

$$p(\Theta|\gamma_1, \gamma_2, \Theta_0)$$

$$= \frac{1}{2\pi} + \frac{1}{\pi}\sum_{n=1}^{\infty} c_n(\gamma_1, \gamma_2)\cos n(\Theta - \Theta_0), \quad (13)$$

where $c_n(\gamma_1, \gamma_2) = c_n(\gamma_1)c_n(\gamma_2)$ is combined with the coefficients

$$c_n(\gamma) = \frac{\sqrt{\pi\gamma}}{2}e^{-\gamma/2}\left[I_{\frac{n+1}{2}}\left(\frac{\gamma}{2}\right) + I_{\frac{n-1}{2}}\left(\frac{\gamma}{2}\right)\right], \quad (14)$$

in which $I_v(x)$ is a modified Bessel function of the first kind and arbitrary v-order.

3.2. Equal SNRs in the RF pulses

A situation with $\gamma = \gamma_1 = \gamma_2$ is that when RF pulses at the receiver have equal amplitudes. Tsvetnov's pdf for this case is

$$p(\Theta|\gamma,\Theta_0)=\frac{e^{-\gamma}}{2\pi}+\frac{e^{-\gamma}}{2\pi}\int_0^{\pi/2}(\gamma\cos z+\lambda)\,e^{\lambda\cos z}dz \tag{15}$$

that may also be computed by (13).

3.3. Von Mises/Thikhonov approximation

We showed in [9] that, with an inaccuracy of less than 0.6% corresponding to a case of equal SNRs, (15) may be modelled using the von Mises/Tikhonov pdf,

$$p(\Theta|\gamma,\Theta_0)=\frac{1}{2\pi I_0(a)}e^{a\cos(\Theta-\Theta_0)}\ ,$$

where a SNR sensitive coefficient $a(\gamma)$ is

$$a(\gamma)=\gamma(1+0.525e^{-1.1503\gamma})$$

and $\gamma=2\gamma_1\gamma_2/(\gamma_1+\gamma_2)$. For a large difference between the SNRs, the error of such an approximation increases up to several percent.

4. LIMITING ERRORS WITH DPM

The error of passive remote SAW sensing with DPM $\varepsilon=\Theta_0-\hat{\Theta}_0$ is estimated by the factors such as the mean error and MSE [7], respectively,

$$\langle\varepsilon\rangle=\Theta_0-\left\langle\hat{\Theta}_0\right\rangle\ , \tag{16}$$

$$\left\langle\varepsilon^2\right\rangle=\Theta_0^2-2\Theta_0\left\langle\hat{\Theta}_0\right\rangle+\left\langle\hat{\Theta}_0^2\right\rangle\ , \tag{17}$$

or root MSE (RMSE) $\sqrt{\langle\varepsilon^2\rangle}$. In regard to (16) and (17), we have two feasible cases to analyze.

4.1. Estimate error for different SNRs

It was shown in [7] that the estimate bias (16) is calculated for (3) by

$$\langle\varepsilon(\gamma_1,\gamma_2)\rangle=\Theta_0-2\sum_{n=1}^{N}\frac{(-1)^{n+1}}{n}c_n(\gamma_1,\gamma_2)\sin n\Theta_0, \tag{18}$$

where the series length may be limited by $N\geq 2\max\gamma$. It follows that (18) reaches zero if $\Theta_0=0$ irrespective of the SNRs in the pulses. Next, for $\Theta_0\neq 0$, the bias rises dramatically if one of the SNRs tends toward zero (lost of pulse). Finally, the bias vanishes almost fully and independently on Θ_0 as the minimum SNR exceeds 10 dB.

The MSE is calculated, by (17), to be

$$\left\langle\varepsilon^2(\gamma_1,\gamma_2)\right\rangle=\frac{\pi^2}{3}+\Theta_0^2+4\sum_{n=1}^{N}\frac{(-1)^n}{n}\beta c_n(\gamma_1,\gamma_2) \tag{19}$$

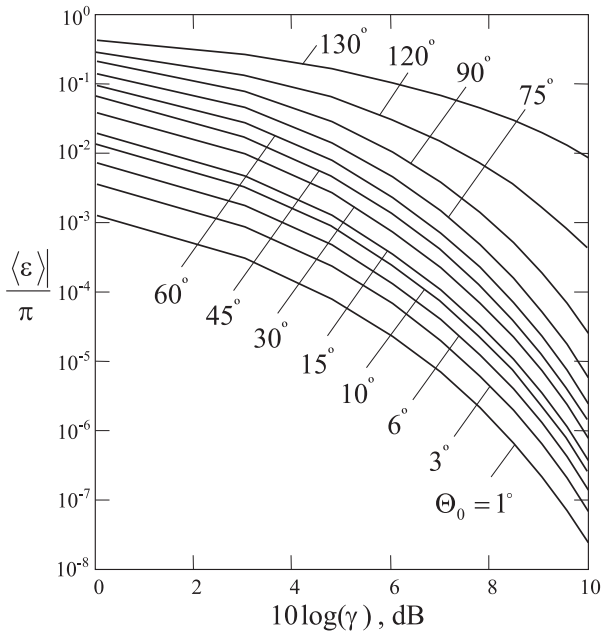


Figure 2: Limiting mean error of remote passive SAW sensing employing single DPM at the ideal coherent receiver.

where

$$\beta=\frac{1}{n}\cos n\Theta_0+\Theta_0\sin n\Theta_0\ .$$

Its maximum value $\pi^2/3+\Theta_0^2$ is kept by zero SNR in one of the pulses ignoring the SNR in the other one. The MSE diminishes to zero asymptotically and very slowly by increasing both SNRs. The following asymptotic solution fits properly the actual data with large SNRs, namely

$$\left\langle\varepsilon^2(\gamma_1,\gamma_2)\right\rangle\big|_{\min\gamma>13\mathrm{dB}}\cong\frac{\gamma_1+\gamma_2}{2\gamma_1\gamma_2}\ . \tag{20}$$

4.1. Estimate error for equal SNRs

In this particular case of $\gamma_1=\gamma_2=\gamma$, (18) degenerates to

$$\langle\varepsilon(\gamma)\rangle=\Theta_0-2\sum_{n=1}^{N}\frac{(-1)^{n+1}}{n}c_n^2(\gamma)\sin n\Theta_0\ , \tag{21}$$

having the same common properties. Figure 2 illustrates (21) for a variety of actual angles Θ_0. It follows that the maximum bias of $<1.8°$ is achieved if $|\Theta_0|<10°$ and it rises by increasing Θ_0. Yet, the value $\gamma>10$ dB makes the bias to be $<1.8°$ for the angles $|\Theta_0|$ up to 130°. On the whole, we notice, the SAW sensor operation with angles Θ_0 lying in the narrow range around zero is preferable from the standpoint of small regular error obtained and low SNR values allowed.

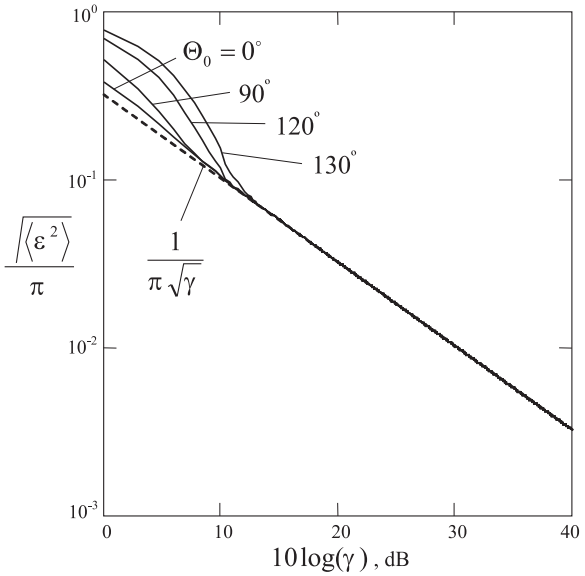


Figure 3: Limiting RMSE of remote passive SAW sensing employing single DPM at the ideal coherent receiver.

The MSE of the estimate is calculated by

$$\langle \varepsilon^2(\gamma) \rangle = \frac{\pi^2}{3} + \Theta_0^2 + 4\sum_{n=1}^{N} \frac{(-1)^n}{n} \beta c_n^2(\gamma) \,. \quad (22)$$

Figure 3 sketches the RMSE associated with (22) in a broad area of angles Θ_0. Observing Fig. 3, we resume that

- The maximum value $\pi^2/3 + \Theta_0^2$ is kept by zero SNR in one of the pulses ignoring SNR in the other one,
- Having large angles $|\Theta_0|$ results in substantial RMSE with low SNR that in the limiting case of $\gamma = 0$dB and $|\Theta_0| = \pi/2$ yields $\sqrt{\langle \varepsilon^2 \rangle} \cong \pi/2$,
- With $\gamma > 13$dB the RMSE changes linearly in the logarithmic measure and is invariant to Θ_0. With large SNRs, one may consider instead an asymptotic solution

$$\langle \varepsilon^2(\gamma) \rangle \big|_{\gamma > 13\mathrm{dB}} \cong \frac{1}{\gamma} \,. \quad (23)$$

On the whole, to get the RMSE with a single DPM to be less than 1.8°, the interrogating pulse must provide the SNR of more than 30 dB at the receiver detector and this is irrespective of the value of Θ_0.

5. LIMITING ERROR PROBABILITY

Another measure of errors in the passive SAW sensing employing DPM is the error probability for the estimate $\hat{\Theta}_0$ to exceed a threshold ζ. Traditionally, we consider two cases.

5.1. Error probability for different SNRs

For this case, the error probability is specified in two equivalent forms [7]

$$P_E(\zeta|\gamma_1,\gamma_2) = \frac{\gamma_\times \sin\zeta}{2\pi} \int_{-\pi/2}^{\pi/2} \frac{e^{-(\gamma_+ - \gamma_- \sin t - \gamma_\times \cos\zeta \cos t)} dt}{\gamma_+ - \gamma_- \sin t - \gamma_\times \cos\zeta \cos t} \,, \quad (24)$$

$$P_E(\zeta|\gamma_1,\gamma_2) = 1 - \frac{\zeta}{\pi} - \frac{2}{\pi}\sum_{n=1}^{N} \frac{1}{n} c_n(\gamma_1,\gamma_2) \sin n\zeta \,. \quad (25)$$

It follows that the error probability reduces as a reciprocal of γ_1, γ_2, and ζ. Furthermore, degrading amplitude in one of the pulses increases error so that with small thresholds and near zero SNR in one of the pulses (lost of signal) the error probability reaches almost unity.

With large SNRs, the following asymptotic is available

$$P_E(\zeta|\gamma_1,\gamma_2)\big|_{\gamma_1,\gamma_2>10\mathrm{dB}} \cong \mathrm{erfc}\left(\zeta\sqrt{\frac{\gamma_1\gamma_2}{\gamma_1+\gamma_2}}\right) \quad (26)$$

$$= 2Q\left(\zeta\sqrt{\frac{2\gamma_1\gamma_2}{\gamma_1+\gamma_2}}\right), \quad (27)$$

where

$$Q(x) = \frac{1}{2}\mathrm{erfc}\left(\frac{x}{\sqrt{2}}\right)$$

is the Gauss Q-function and $\mathrm{erfc}(x)$ is the complimentary error function.

5.2. Error probability for equal SNRs

By $\gamma = \gamma_1 = \gamma_2$, (24) and (25) produce

$$P_E(\zeta|\gamma) = 1 - \frac{\zeta}{\pi} - \frac{2}{\pi}\sum_{n=1}^{N} \frac{1}{n} c_n^2(\gamma) \sin n\zeta \quad (28)$$

$$= \frac{e^{-\gamma}}{\pi} \sin\zeta \int_0^{\pi/2} \frac{e^{\gamma\cos\zeta\cos t}}{1-\cos\zeta\cos t} dt \,, \quad (29)$$

having, for $\gamma > 10$dB, the asymptotic solution

$$P_E(\zeta|\gamma)\big|_{\gamma>>10\mathrm{dB}} \cong \mathrm{erfc}\left(\zeta\sqrt{\frac{\gamma}{2}}\right) \quad (30)$$

$$= 2Q\left(\zeta\sqrt{\gamma}\right). \quad (31)$$

Figure 4 illustrates the error probability in a wide range of SNRs and thresholds.

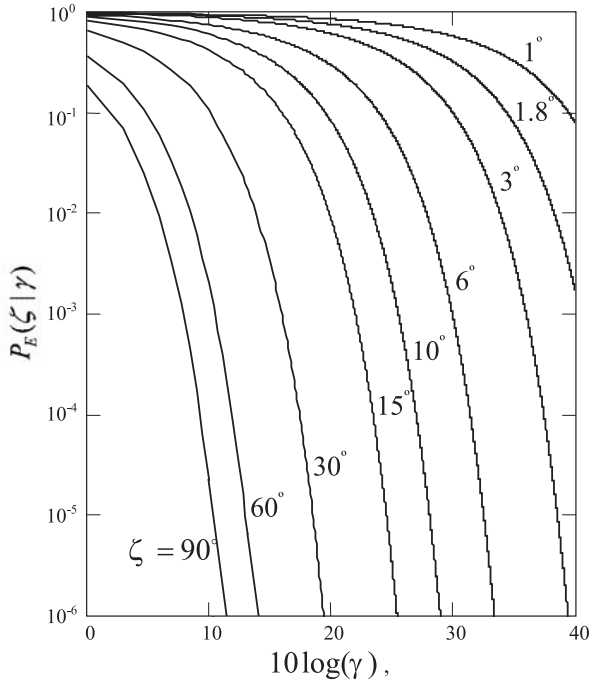


Figure 4: Limiting error probability for the estimate to exceed a threshold.

6. CONCLUSION

In this paper we observed limiting estimate errors of the time delay (phase difference) between two reflectors placed on the piezoelectric crystal surface of the SAW sensor. All estimates were obtained for a single remote wireless DPM at the ideal coherent receiver. We investigated the regular error (bias), the MSE, and the error probability to exceed a threshold. Several useful and reasonably accurate approximate solutions were also discusses. The results achieved were simulated by generating the quadrature components of the narrowband Gaussian processes and then processing the histograms. We notice that, with large number of points and high resolution in phase, the simulated curves fit the theoretical estimates.

The brief guideline for designers exploiting single DPM in the "ideal channel" is as follows:

- Having $\gamma > 10$ dB makes the bias less than 1.8° for the angles $|\Theta_0|$ up to 130°.
- The RMSE of the estimate diminishes by $1/\sqrt{\gamma}$ for arbitrary Θ_0 if γ exceeds 13 dB; the RMSE becomes less than 1.8° if $\gamma \geq 30$ dB and 0.18° if $\gamma \geq 50$ dB irrespective of the value of Θ_0.
- The error probability of $P_E < 0.01$ for the threshold of $\zeta = 1.8°$ and arbitrary Θ_0 is achieved with $\gamma > 40$ dB.

REFERENCES

[1] W.-E. Bulst, G. Fischerauer, and L. Reindl, State of the art in wireless sensing with surface acoustic waves, IEEE Trans. on Industrial Electronics, vol. 48, pp. 265-271, Apr 2001.

[2] F. Schmidt, O. Sczesny, C. Ruppel, and V. Mágoni, Wireless interrogator system for SAW-Identification-Marks and SAW-Sensor components, in Proceedings IEEE Int. Freq. Contr. Symp., 1996, pp. 208-215.

[3] G. Scholl, F. Schmidt, T. Ostertag, L. Reindl, H. Scherr, and U. Wolff, Wireless passive SAW sensor systems for industrial and domestic applications, in Proceedings IEEE Int. Freq. Contr. Symp., 1998, pp. 595-601.

[4] A. Pohl, A review of wireless SAW sensors, IEEE Trans. on Ultrason, Ferroel. and Freq. Contr., vol. 47, pp. 317-332, Mar 2000.

[5] V.V. Tsvetnov, Statistical properties of signals and noises in two-channel phase systems, Radiotekhnika, vol. 12, pp. 12-29, May 1957.

[6] V.V. Tsvetnov, Unconditional statistical characteristics of signals and uncorrelated Gaussian noises in two-channel phase systems, Radiotekhika i Elektronika, vol. XIV, pp. 2147-2159, Dec 1969.

[7] Yu.S. Shmaliy, Limiting Phase Errors of Passive Wireless SAW Sensing with Differential Measurement, IEEE Sensors J., vol. 4, pp. 819-827, Dec 2004.

[8] R.F. Pawula, S.O. Rice, and J.H. Roberts, Distribution of the phase angle between two vectors perturbed by Gaussian noise, IEEE Trans. Comm., vol. COM-30, pp. 1828-1841, Aug 1982.

[9] Yu.S. Shmaliy, Von Mises/Tikhonov distributions for systems with differential phase measurement, Signal Processing, vol. 85, pp. 693-703, Apr 2005.

LOW-LOSS NSPUDT SAW FILTERS USING INCLINED REFLECTORS

S. A. Doberstein, A. Shevtsov, S. Zhgoon
ONIIP, Russia, 644009, Omsk-9, 231 Maslennikov str., e-mail: info@oniip.ru, fax: 7-381-2-51-49-87

ABSTRACT

This paper presents the low-loss natural single phase unidirectional transducer (NSPUDT) SAW filters, which use SAW propagation direction reversal without the reversal of NSPUDT directivity. In this design the input and output NSPUDTs are placed in parallel acoustic tracks. SAW transmission between these tracks and propagation direction reversal are provided by a double reflection from the inclined reflectors. An asymmetric Quartz crystal orientation was chosen where conventional IDTs with the λ/4 electrodes have SAW unidirectional radiation (λ is the SAW wavelength on the central frequency of the filter). The 56 MHz and 172 MHz experimental filters with the LC-matching showed an insertion loss of 8-10 dB and low ripple within the passband.

1. INTRODUCTION

For development of the low-loss NSPUDT SAW filters the reversing a directivity of one of the NSPUDT is commonly used [1-4]. This complicates the manufacture technology and reduces design flexibility. Another way of obtaining the low-loss proposes the reversal of NSPUDT directivity. In this case inclined NSPUDTs are used and SAW transmission between them is provided by a single reflection of SAW from a polished edge of a piezoelectric substrate [5]. However this solution can be used only when the IDTs are placed symmetrically at specific angles equal to ±25° in relation to the X-axis on ST-Quartz. These features limit the design flexibility and application of NSPUDT SAW filters. This paper presents another design concept of low-loss NSPUDT SAW filters which uses SAW propagation direction reversal without the reversal of NSPUDT directivity. The design uses a simple single-level technology and increases the design flexibility of low-loss NSPUDT SAW filters.

2. OPERATION PRINCIPLE AND DESIGN CONCEPT

A filter design similar to the widely known structure with a double reflection [6] but an asymmetric ST-Quartz crystal orientation with SAW propagation +25° relative to the X-axis is used as a piezoelectric substrate. A SAW unidirectional radiation and detection is possible by a conventional IDT with an electrode width of λ/4 along this direction [1]. In the suggested design the input and output NSPUDTs are placed in parallel acoustic tracks. The inclined reflectors are placed on the side of the primary SAW radiation direction from the NSPUDTs.

SAW transmission between the parallel tracks is provided by a double reflection of SAW from inclined reflectors (Fig. 1). As seen in Fig. 1 in this design the low-loss NSPUDT SAW filter is realized by the SAW propagation direction reversal without the reversal of NSPUDT directivity. To obtain the low-loss SAW filter with inclined reflectors it is necessary to provide a 90°

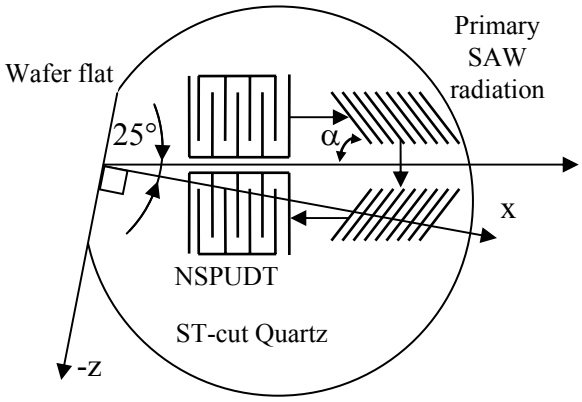


Fig. 1: Design concept of the low-loss NSPUDT SAW filter with inclined reflectors arrays

SAW reflection in the inclined reflectors. For this purpose we must choose an optimal angle β between a perpendicular to reflector electrode and SAW propagation direction on ST (X+25°) Quartz. A solution of this problem is illustrated in Fig. 2. The inclined reflectors are characterized by an angle β and pitch d. The angle β is determined from a condition when a phase shift of an incident SAW (on segment L_{inc}) equals to a phase shift of a reflected SAW (on segment L_{ref}):

$$(\omega/V_{inc})L_{inc}=(\omega/V_{ref})L_{ref}, \quad L_{inc}/L_{ref}= V_{inc}/V_{ref},$$

$$\beta = \mathrm{arctg}(L_{inc}/L_{ref})=\mathrm{arctg}(V_{inc}/V_{ref}).$$

Here $\omega = 2\pi f_o$, V_{inc}, V_{ref} is a velocity of the incident and reflected SAW, respectively; f_o is a synchronism frequency of the IDT.

The pitch d is determined from a condition when a phase shift of the incident SAW on a segment L is equal to 2π:

$$(\omega \cdot V_{inc}) \cdot L=2\pi, \quad L= V_{inc}/f_o,$$
$$d=L\cdot\cos(\beta)=(V_{inc}/f_o)\cdot\cos(\beta)=\lambda\cdot\cos(\beta).$$

Here $\lambda = V_{inc} / f_o$ is an IDT period.

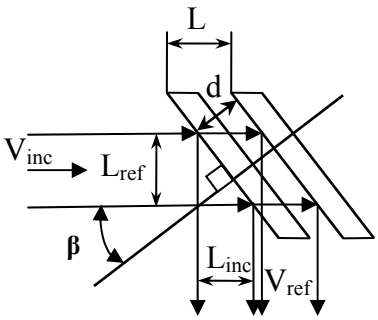


Fig. 2: The incident and 90° reflected SAW in the inclined reflectors

The calculated data for the incident and 90° reflected SAW on ST(X+25°) Quartz are presented in Table 1 (here k^2 is electromechanical coupling coefficient).

Table 1: Calculated data for the incident and 90° reflected SAW on ST(X+25°) Quartz

Parameter	Incident SAW	Reflected SAW
Euler angles	(0°,132.75°,25°)	(0°,132.75°,115°)
Phase velocity	3249 m/s	3455 m/s
k^2	0.135%	0.03%
β	43.2°	
d	0.73 λ	

3. EXPERIMENTAL RESULTS

The experimental samples of the filters were developed on two central frequencies f_o: 56 MHz and 172 MHz. The design parameters for these samples are given in Table 2.

Table 2: Design parameters of the filter samples

Parameter	f_0 =56 MHz	f_0 =172 MHz
IDT period (λ)	58 μm	18.7 μm
Number of IDT electrodes	402	402
Number of the reflector electrodes	300	300
Aperture of IDT	5800 μm	1870 μm
Aluminum thickness	0.4 μm 1.7 μm	0.4 μm 0.6 μm
Gap between IDTs	2900 μm	936 μm
α	46.8°	46.8°

The measured frequency response of the 172 MHz unmatched filter in a 50-Ω system has shown in Fig. 3. An aluminum thickness h was 0.4 μm. The insertion

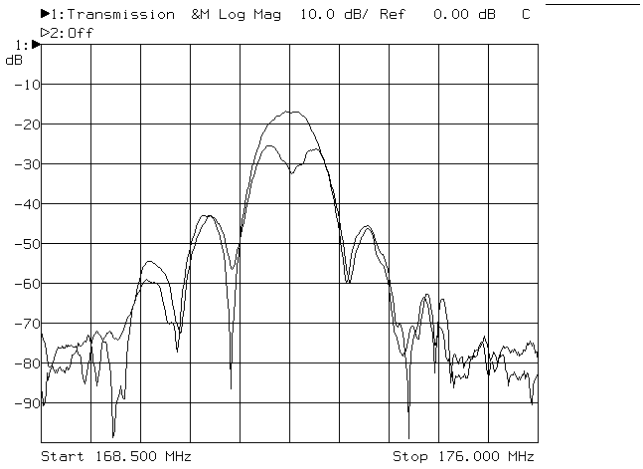


Fig. 3: Measured frequency response of the 172 MHz unmatched low-loss NSPUDT filter (h=0.4 μm)

loss for the 56 MHz filter and the 172 MHz filter was 16 and 17 dB respectively. To demonstrate that the NSPUDTs used the asymmetric crystal orientation for achieving the unidirectivity the device was fabricated on ST-Quartz substrate that had been rotated on 180°. In this case the inclined reflectors will be placed on the opposite side of the primary SAW radiation direction from the NSPUDTs (Fig. 1). The frequency response of the filter has a typical dip of 15 dB in a passband (Fig. 3). This characterizes the directivity of the NSPUDT for ST (X+25°) Quartz with an aluminum thickness of 0.4 μm.

To decrease the insertion loss it is necessary to match the filters with a 50-Ω system. Matching was provided by the LC elements (Fig. 4). Fig 5,6 show the measured input/output impedance characteristics of the 172 MHz filter after matching. As will be seen from Fig. 5,6 at the center frequency of 172 MHz the input/output impedances of the filter are close to real value of 50 Ω.

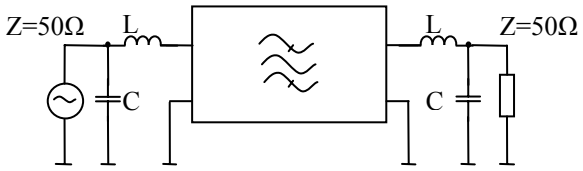


Fig. 4: Matching circuit of the filter

Fig. 7 shows the measured frequency response of the filter matched with a 50-Ω system. The insertion loss was reduced to 11.5 for the 56 MHz filter and to 12.5 dB for the 172 MHz filter. The ripple in the passband became very low. The 3-dB bandwidth was 178 kHz for the 56 MHz filter and 550 kHz for the 172 MHz filter. The sidelobe suppression was about 20 dB. Evidently the directivity of the NSPUDTs and SAW reflection from the inclined reflectors depend on an aluminum thickness. The measured frequency responses of the

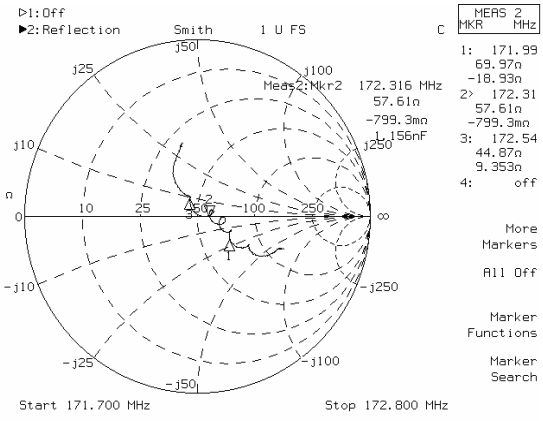


Fig. 5: Measured input impedance of the 172 MHz filter after matching

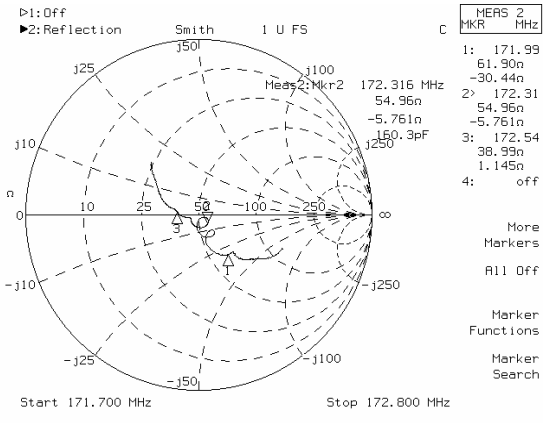


Fig. 6: Measured output impedance of the 172 MHz filter after matching

55.5 MHz and 172 MHz matched filters with an aluminum thickness of 1.7 µm and 0.6 µm, respectively, have shown in Fig. 8, 9. The insertion loss was reduced to 8.2 for the 55.5 MHz filter and 10.5 dB for the 172 MHz filter. Fig. 9 shows measured frequency response of the 172 MHz filter with an opposite directivity. The typical dip in the passband came to 45 dB, 30 dB more than in the case when the thickness was 0.4 µm (Fig. 3). As seen in Fig. 8, 9 the sidelobe suppression was about 20 dB. The calculations show that increasing sidelobe suppression is achieved by decreasing the number of the electrodes in one of IDTs. The measured frequency response of the 56 MHz filter (h=0.4 µm) with the number of the electrodes reduced to 160 in an output IDT has shown in Fig. 10. The filter provides the sidelobe suppression over 30 dB. The insertion loss was increased to 12 dB by means of the degradation of the NSPUDT directivity with the reduced number of electrodes.

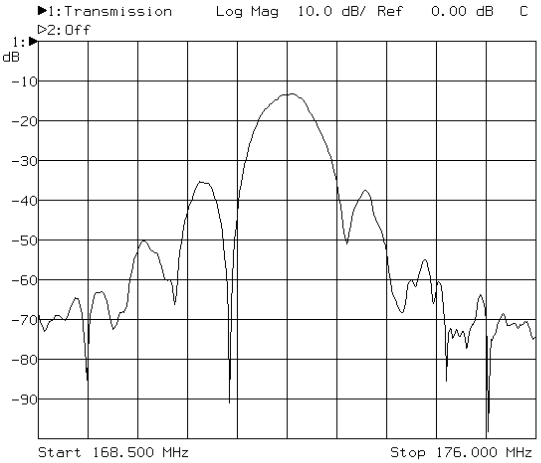


Fig. 7: Measured frequency response of the 172 MHz low-loss NSPUDT filter with matching (h=0.4 µm)

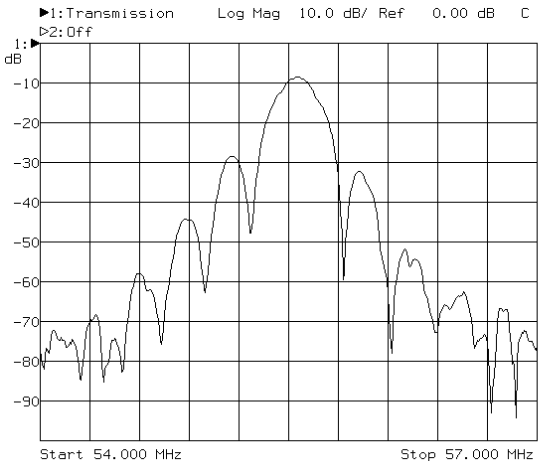


Fig. 8: Measured frequency response of the 55.5 MHz low-loss NSPUDT filter with matching (h=1.7 µm)

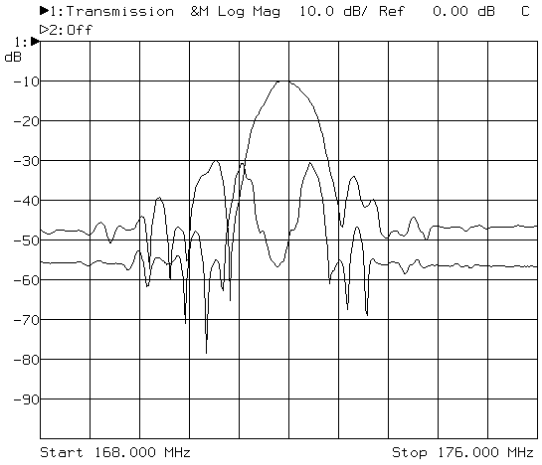


Fig. 9: Measured frequency response of the 172 MHz low-loss NSPUDT filter with matching (h=0.6 µm)

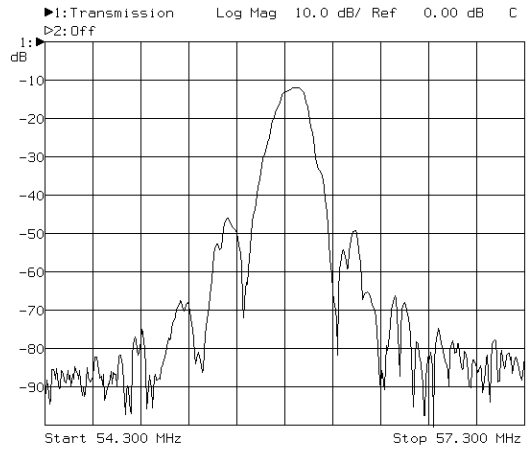


Fig. 10: Measured frequency response of the filter with reduced number of the electrodes in one of IDTs

4. TEMPERATURE CHARACTERISTICS

The temperature characteristics were investigated on the 172 MHz filter with an aluminum thickness of 0.4 μm. We measured a fractional change of the center frequency $\Delta f/f_o$ and change of the insertion loss versus temperature. The range of the temperature variation was from -60°C to +60°C (Fig. 11). The measured temperature coefficient of frequency (TCF) was ≈ -8.3 ppm/°C, which is close to the calculated TCF=-11.9 ppm/°C for ST(X+25°) Quartz [2]. The insertion loss was changed to ≈ ±4 dB. This is associated with the temperature dependence of a power flow angle of the NSPUDTs. This effect degrades the reflection of the inclined reflectors. To improve the temperature stability of the NSPUDT filters we can increase an angle Θ of the rotated Y-cut Quartz and simultaneously decrease a chosen angle of SAW propagation direction in relation to the X-axis [2]. It is necessary for NSPUDT operation. This will lead to a re-calculation of the angle β for the inclined reflectors.

5. CONCLUSION

We have presented a design of the low-loss NSPUDT SAW filter with SAW propagation direction reversal but without the reversal of NSPUDT directivity. The conventional IDTs with the λ/4 electrodes are used in the filter. These input and output IDTs are placed in parallel acoustic tracks on the ST(X+25°) Quartz. Unidirectional SAW transmission between the parallel tracks is provided by a double reflection of SAW from the inclined reflectors. The presented concept of the low-loss NSPUST SAW filter essentially simplifies the filter design and increases the development flexibility of the similar filters. The 56 MHz and 172 MHz experimental samples of the filters had an insertion loss of 8-10 dB and very low ripple with matching. The temperature characteristics were shown. Evidently the

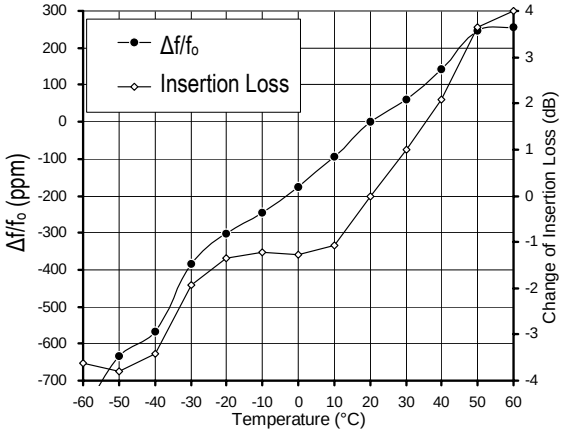


Fig. 11: Temperature characteristics of the low-loss NSPUDT SAW filter

presented results are not limiting values for the investigated design. The achieved level of the insertion loss is still high, sidelobe suppression is low and temperature stability is not enough. The further studies of the suggested design of the low-loss NSPUDT SAW filter will be devoted to the above indicated problems.

ACKNOWLEDGEMENTS

We would like to thank Mr. D. M. Gilfand for editing and refining manuscript of the paper.

REFERENCES

[1] P.V. Wright, "The Natural Single-Phase Unidirectional Transducer: A New Low-Loss SAW Transducer", in Proceedings of the IEEE Ultrasonics Symposium, 1985, pp. 58-63.

[2] C.S. Lam and D. Gunes, "A Low-Loss SAW Filter Using Two Finger Per Wavelength Electrodes on the NSPUDT Orientation of Quartz", in Proceedings of the IEEE Ultrasonics Symposium, 1993, pp. 185-188.

[3] K. Yamanouchi, M. Takeuchi, H. Odagawa and M. Tanaka, "Low-Loss SAW Filter Using Thickness Difference Type of IDT on the NSPUDT Orientation Substrate", in Proceedings of the IEEE Frequency Control Symposium, 1995, pp. 537-541.

[4] M. Takeuchi, H. Odagawa, M. Tanaka and K. Yamanouchi, "SAW Transducer Configurations for Reversing the Directivity of NSPUDT Substrates", in Proceedings of the IEEE Ultrasonics Symposium, 1995, pp. 17-22.

[5] S. Doberstein, "Low-Loss NSPUDT SAW Filter without Reversing the Directivity of NSPUDT", in Proceedings of the IEEE Ultrasonics Symposium, 1999, pp. 29-32.

[6] Herbert Matthews, Surface Wave Filters. Design, Construction, and Use, 2nd ed., John Willey & Sons, New York, 1977, ch. 9.

COHERENT POPULATION TRAPPING RESONANCES IN THERMAL ^{87}RB VAPOR

G. Kazakov[1*], I. Mazets[2], Yu. Rozhdestvensky[3], G. Mileti[4], J. Delporte[5], and B. Matisov[1]

[1]St. Petersburg State Polytechnic University, Politekhnicheskaya 29, St. Petersburg195251, Russia

[2]A.F. loffe Physics-Technical Institute, Politekhnicheskaya 26, St. Petersburg 194021, Russia

[3]Institute for Laser Physics, Birzhevaya liniya 12, St. Petersburg 199034, Russia

[4]Observatoire Cantonal Neuchatel, rue de l'Observatoire 58, CH-2000, Neuchatel, Switzerland

[5]Centre National d'Etudes Spatiales (CNES), avenue Belin 18, Toulouse 31401, France

ABSTRACT

Three different all-optical repumping schemes aimed to improve the contrast of the dark resonance in thermal ^{87}Rb vapor are considered: excitation by a σ^+-polarized bichromatic laser radiation and $\sigma^+ - \sigma^-$ bichromatic co-propagating fields, and a novel excitation method with transversal laser pumping. Qualitative considerations are confirmed by numerical calculations

1. INTRODUCTION

At present, quantum frequency standards are becoming important for a variety of applications. There are two main problems in this area: 1) the maximum obtainable stability and 2) the size of a standard, which should occupy a reasonable volume. E.g., on-board atomic clocks for space systems must have the stability of about 10^{-14} and the volume in the range $1 \div 100$ cm^3. These parameters, in principle, can be achieved for frequency standards using a gas cell with ^{87}Rb atoms as the main element. In this paper we consider all-optical frequency standards based on coherent population trapping (CPT) effect.

The CPT effect of atoms interacting with a resonant electromagnetic field is well known (see reviews [1,2] and references therein) and widely used in various fields of atomic and laser physics. Efforts to design and build CPT-based atomic clocks were undertaken by several scientific groups [3-11]. The main advantage of quantum frequency standard based on the CPT effect is that the corresponding RF resonance (so-called dark resonance) is excited by all-optical methods. However, practical use of the CPT phenomenon requires thorough optimization of the two-photon resonance parameters (width, amplitude, contrast). We analyze three different excitation schemes for a quantum frequency standard based on CPT in the ^{87}Rb gas cell and compare the dark-resonance parameters for these schemes.

2. BASIC ASSUMPTIONS

In this work we use the density matrix approach. We take into account the real hyperfine (hf) and Zeeman structure of the states involved as well as the exact values of the probabilities of the optically-induced one-photon transitions and spontaneous relaxation of the optically excited states. Effects of the thermal motion of atoms (Doppler broadening) also have been taken into account.

For ^{87}Rb at 300 K the Doppler widths of the D_1 and D_2 lines are about 400 MHz, what is more then the hf splitting of the excited state with $J' = 3/2$, but less then the hf splitting of the excited state with $J' = 1/2$ (about 800 MHz)

We assume that the collisional quenching of the excited state is negligible. Therefore optical relaxation (i.e. population relaxation from excited state to the different sublevels of the ground state) is provided by spontaneous transitions with the rate $\gamma = \gamma_{sp} \approx 3.6 \cdot 10^7\ s^{-1}$ both for the D_1- and D_2-line. Relaxation of coherences between ground and excited states is assumed to be equal to $\gamma_{sp}/2$.

Also we suppose, when considering the counter-propagating $\sigma^+ - \sigma^-$ laser beam configuration, that the vapor cell contains a buffer gas that provides Dicke narrowing of the hf transition [8].

The relaxation rate Γ of coherences between the different hyperfine- and Zeeman sublevels of the ground state was supposed to be equal to $\Gamma \approx 200\ s^{-1}$. The depolarization rate in ground state is supposed to be equal to Γ.

* E-Mail: kazakov@quark.stu.neva.ru

3. EXCITATION BY A σ^+-POLARIZED BICHROMATIC LASER FIELD

In CPT based clocks, the Zeeman sublevels $|1\rangle = |F=1, m=0\rangle$ and $|2\rangle = |F=2, m=0\rangle$ of two ground-states hyperfine components of an alkali metal atom are the working energy levels. The coherence between these transitions is created by a two-photon Raman transition induced by a circularly-polarized laser field containing two components with the frequencies ω_1 and ω_2

$$\vec{E}(z,t) = \frac{\vec{e}_{+1}}{2}(E_1 e^{-i(\omega_1 t - k_1 z)} + E_2 e^{-i(\omega_2 t - k_2 z)}) + c.c.$$

in the presence of a level-splitting static magnetic field oriented along the z axis. Here $\vec{e}_{+1} = -(\vec{e}_x + i\vec{e}_y)/\sqrt{2}$ is the unit vector of the right-handed circular polarization and $E_{1,2}$ are the amplitudes of corresponding frequency components. Let us introduce the two-photon detuning Ω as

$$\Omega = (\Delta_{hfs} - (\omega_1 - \omega_2))/2 .$$

When Ω is scanned across zero, a narrow dip (dark resonance) is observed in the absorption spectrum.

The dark resonance width is determined by the rate of the atomic coherence relaxation and the laser field parameters [1,2].

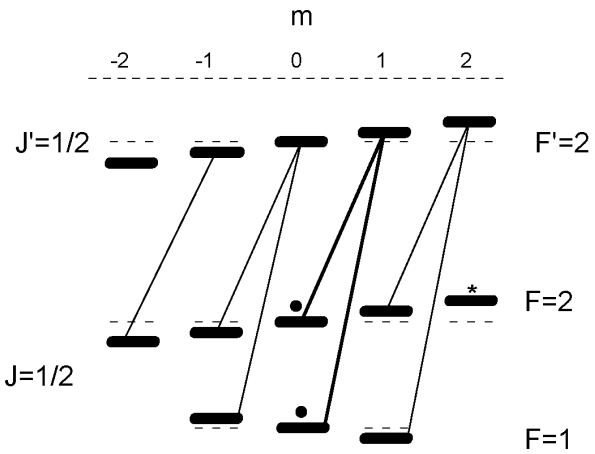


Fig.1. Scheme of optically-induced transitions in ^{87}Rb atom in σ^+-fields. Optical fields are tuned in resonance with the $|F=2\rangle \rightarrow |F'=2\rangle$ and $|F=1\rangle \rightarrow |F'=2\rangle$ transitions. The working states $|1\rangle, |2\rangle$ are marked by black circles, and the "pocket" is marked by an asterisk.

However, in the case of circularly-polarized fields there is always a trapping state of an alkaline atom (so-called "pocket", see Fig. 1). Atoms are accumulated in the pocket and do not contribute any more to the signal formation on the working transition. It leads to a drastic decrease of the amplitude and contrast of the signal.

4. EXCITATION BY $\sigma^+ - \sigma^-$ BICHROMATIC CO-PROPAGATING FIELDS

One of the possible solution to this problem is to use both σ^+- and σ^--fields (see Fig. 2.). This scheme contains no "pockets". However, such an excitation scheme belongs to the co-called closed loop class. It leads to sufficient dependency of dark-line parameters on total phase of laser fields [9].

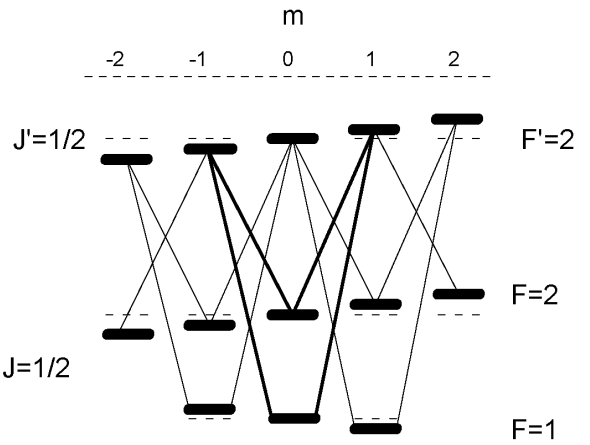


Fig.2. Scheme of optically-induced transitions in ^{87}Rb atom in $\sigma^+ - \sigma^-$ fields. Optical fields are tuned in resonance with the $|F=2\rangle \rightarrow |F'=2\rangle$ and $|F=1\rangle \rightarrow |F'=2\rangle$ transitions. «Closed loop» involving the working transitions is shown by bold black lines

A scheme involving counter-propagating light waves with orthogonal polarizations has been proposed [5], but in that case, a significant periodic variation of the resonance amplitude emerges. The period $\pi/(k_1 - k_2)$, where $k_j = 2\pi/\lambda_j$ is the wave number of the laser field of the frequency ω_j, is due to the difference of the wavelengths of the two frequency components of the laser radiation. For example, this period is equal to approximately 2 cm for ^{87}Rb atoms. Therefore spatial averaging over the cell decreases the dark resonance amplitude for the cell length $L \geq \pi/(k_1 - k_2)$ [6]. On the other hand, small cells of the length $L << \pi/(k_1 - k_2)$ are characterized by large values of the atomic relaxation at cell walls. To decrease the adverse influence of the wall relaxation, one needs to apply high buffer gas pressure (150 Torr in a cell of 12 mm^3 volume [7]). However, there are also inelastic collisions of the alkali atoms with the buffer gas destroying the coherence on the working transition and thus resulting in decrease of the amplitude and significant broadening of the dark resonance. Thus the scheme proposed in [5] yielded practically only a small (1.4-fold) increase of the amplitude [7]

compared to the usual co-propagating light waves configuration.

We considered theoretically the case of co-propagating bichromatic σ^+ and σ^- laser fields, which is shown to be free of aforementioned disadvantages. These fields are supposed to be the decompositions (in the terms of the circular polarizations) of two bichromatic linear polarized fields with orthogonal polarization planes:

$$\vec{E} = \frac{\vec{e}_x}{2} E_1 \exp\left[i(k_1 z - \omega_1 t)\right] + \frac{\vec{e}_y}{2} E_2 \exp\left[i(k_2 z - \omega_2 t)\right] + c.c.$$

In such a configuration shown in Fig. 3, the optimum total phase of excitation over the closed loop is maintained through the whole cell, resulting in overcoming the size limitations inherent to the scheme of Ref. [5]. Similar schemes were proposed in [10,11].

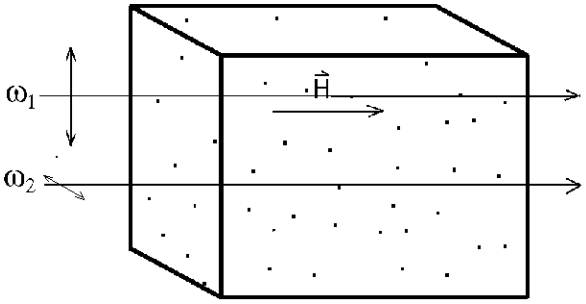


Fig.3. Spatial configuration of the atomic vapor cell and the laser beams for σ^+ - σ^- co-propagating excitation scheme

5. EXITATION SCHEME BASED ON THE PERPENDICULAR PUMP-PROBE LASER BEAMS

In order to preserve atoms from leaving the working sublevels and thus enhance the amplitude and contrast, we propose a new method, based on the use of a two-frequency optical pumping, as shown in Fig.4. (pumping scheme with a similar spatial configuration was considered in a different context [12])

Let us suppose that ^{87}Rb atoms in a cell interact with the linearly-polarized laser field (the pumping field)

$$\vec{E}_{pump} = \frac{\vec{e}_z}{2}\left(E_3 \exp\left[i(k_3 x - \omega_3 t)\right] + E_4 \exp\left[i(k_4 x - \omega_4 t)\right]\right) + c.c.$$

One of the components of the pumping field is tuned in resonance with the $|F=2\rangle \rightarrow |F'=2\rangle$ transition and the second component is tuned in resonance with the $|F=1\rangle \rightarrow |F'=1\rangle$ transition, both these transitions belonging to the D_1-line (see Fig.5)

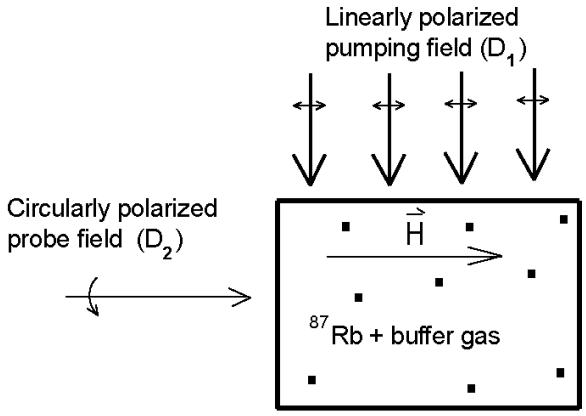


Fig.4. The proposed Rb Gas Cell Frequency Standard for the novel transversal pumping scheme

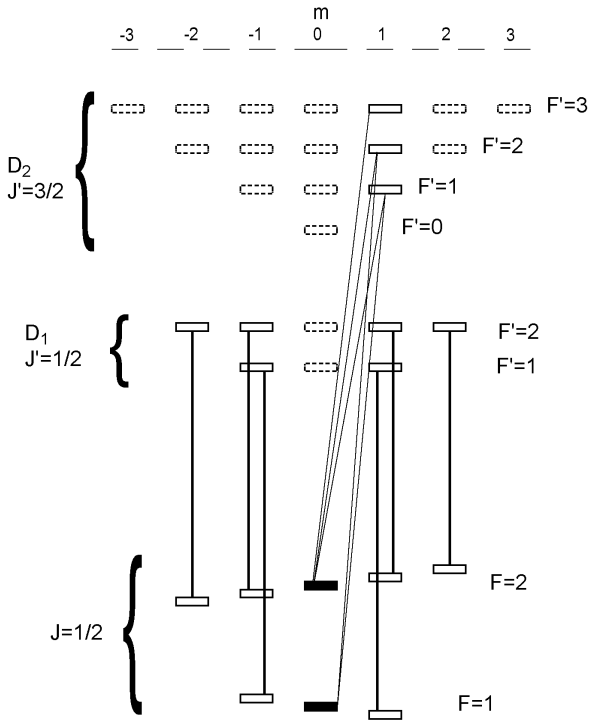


Fig.5. The scheme of optically-induced transitions between the energy levels of ^{87}Rb. Atoms are accumulated in the two working levels (black rectangles) due to optical pumping via the transitions indicated by bold vertical lines and spontaneous relaxation. The transitions excited by the probe field are shown by thin inclined lines.

In this case, atoms are not pumped out from the working levels $|1\rangle = |F=1, m=0\rangle$ and $|2\rangle = |F=2, m=0\rangle$, since the corresponding Clebsch-

Gordan coefficients and, hence, transition dipole matrix elements are equal to zero [13]. As a result, after few optical pumping cycles (coherent excitation followed by spontaneous relaxation of the optically excited state) all the atoms will be accumulated on the working sublevels $|1\rangle$ and $|2\rangle$. Note that the two frequency components of the pumping field don't need to be correlated.

Suppose now we switch on the weak (compared to the pumping field) probe field

$$\vec{E}_{probe} = \frac{\vec{e}_{+1}}{2}\left(E_1 \exp\left[i(k_1 z - \omega_1 t)\right] + E_2 \exp\left[i(k_2 z - \omega_2 t)\right]\right) + c.c.,$$

which is circularly polarized and consists of the two correlated frequency components tuned in resonance with the transitions from the ground state components to the optically excited state with $J' = 3/2$ (see Fig. 5). Since all the atoms are accumulated in the working sublevels, the contrast and amplitude in the probe beam absorption signal is greatly enhanced.

6. NUMERICAL CALCULATIONS

We have performed numerical calculations of the absorption in a ^{87}Rb vapor cell at room temperature. We propose that the vapor cell is filled with a buffer gas and contains N alkali atoms in the evaporated phase at the temperature of about 300 K. Let us denote the total population in excited states (for the σ^+ and σ^+-σ^- excitation schemes), or the total population on all the sublevels of the $J' = 3/2$ excited state (for the scheme with transversal pumping) as ρ_e.

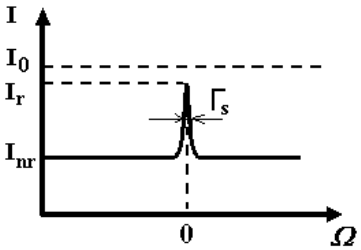


Fig.6. Qualitative picture showing the dark resonance line shape. Γ_S is a width of dark resonance

We find ρ_e from the density matrix equations. Then the absorbed power of probe field δU_{probe} is equal to the power emitted by spontaneous transition:

$$\delta U_{probe} = \gamma \cdot \hbar\omega \cdot N \cdot \rho_e,$$

where $\hbar\omega$ is the optical transition energy. The "absorbed photodetector current" is equal to the difference between the photodetector current without (I_0) and with (I) the presence of the gas cell, see Fig.6:

$$I_0 - I = \delta U_{probe} \cdot \frac{e}{\hbar\omega} \cdot \kappa = \gamma \cdot e \cdot N \cdot \rho_e \cdot \kappa$$

where e is elementary charge, $\kappa \approx 1$ — effective quantum efficiency. In our calculations we suppose that the vapor cell contains $N = 10^{11}$ Rb atoms, that is typical for a cell of the volume ~10 cm^3 at 300 K [14].

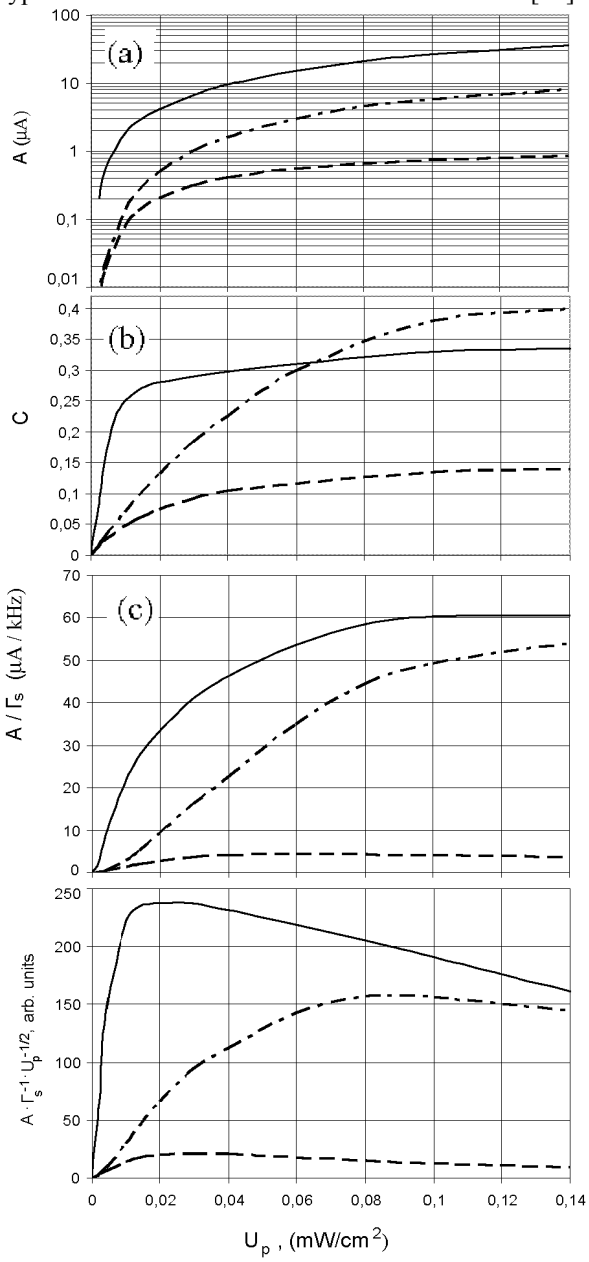


Fig.7. (a) Amplitude, (b) contrast, (c) amplitude-to-width ratio (d) amplitude-to to-width ratio, divided by the square root of the probe field intensity of the dark resonance versus the probe field intensity. Solid line is for the novel scheme involving perpendicular pumping beam with the intensity $U_0 = 10\,mW/cm^2$ (in this case about 90% of atoms are accumulated on the working levels). Dashed line is for the traditional scheme involving the σ^+-polarized probe beam only. Dot-dashed line is for the co-propagating σ^+- and σ^--probe beams.

The amplitude A of the dark resonance is defined as the difference between the photodetector current I_r on under the two-photon resonance condition and the same current I_{nr} far from this resonance. *The contrast* C is the ratio of the amplitude to the "absorbed current" $I_0 - I_{nr}$ far from the resonance.

The results of numerical calculations are shown in Fig. 7. It shows that the amplitude-to-width ratio, divided by the square root of intensity (that determines the figure of merit of quantum frequency standard [14] in the case when the noise is determined by photodetector shot noise) for transversal pumping scheme and σ^+ - σ^- scheme are one order of magnitude larger than for the σ^+-polarized beam only. So we can conclude that schemes with σ^+ - σ^- co-propagating laser beams and with transversal laser pumping can give much better stability than σ^+-polarized beam only. Standard stability estimations [14] demonstrate that the best short-term stability attainable by these schemes is $\sigma_y \le 10^{-13}/\sqrt{\tau}$.

This research is supported by the INTAS-CNES, project 03-53-5175, by the Ministry of Education and Science of Russia, project UR.01.01.287 and by the Russian Foundation for Basic Research, project 05-02-26523-z.

REFERENCES

[1] B.D. Agap'ev, M.B Gorny, B.G Matisov, and Yu.V. Rozhdestvensky, Coherent Population Trapping in Quantum Systems,@ Physics – Uspekhi, vol. **163**, № 9, pp.1-36, September 1993.

[2] E. Arimondo, Coherent Population Trapping in Quantum Systems, Progress in Optics, Ed. E. Wolf, Amsterdam: North-Holland, 1996, vol. **35**, pp. 257-354.

[3] J.Kitching, S.Knappe, N.Vukicevic, et. al., A Microwave Frequence Reference Based on VCSEL-Driven Dark Line Resonances in Cs Vapor, @ IEEE Trans. Instrum. Meas. vol. **49**, № 6, pp.1313-1317, December 2000.

[4] M. Merimaa, T. Lindvall, I. Tittonen, et. al., All-optical Atomic Clock Based on Coherent Population Trapping in ^{85}Rb,@ JOSA B, vol. **20**, № 2, pp. 273-279, February 2003.

[5] A.V. Taichenachev, V.I. Yudin, V.L. Velichansky, et. al., High-Contrast Dark Resonances on the D_1 Line of Alkali Metals in the Field of Counterpropagating Waves,@ JETP Letters, vol. **80**, № 4, pp. 236-240, 2004.

[6] C. Affolderbach, S. Knappe, R. Wynands, et. al., Electromagnetically Induced Transparency and Absorption in a Standing Wave,@ Phys. Rev. A, vol. **65**, 043810, pp.1-10, 2002.

[7] S.V. Kargapoltsev, J.Kitching, L.Hollberg, et. al., High-Contrast Dark Resonances in $\sigma_+ - \sigma_-$ Optical Field,@ Laser Phys. Lett, vol. **1**, № 10, pp. 495-499, 2004.

[8] S. Brandt, A. Nagel, R. Wynands, and D. Meshede, Buffer-gas-induced Linewidth Reduction of Coherent Dark Resonances to below 50 Hz,@ Phys. Rev. A, vol. **56**, № 2, pp. R1063-RR066, August 1997.

[9] D. V. Kosachev, B. G. Matisov, Yu. V. Rozhdestvensky, Coherence phenomena in Multilevel Systems with Closed Interaction Contour,@ J. Phys. B, vol. **25,** pp. 2473-2488, 1992.

[10] Y.-Y. Jau, E. Miron, W. Happer, et.al., Push-pull optical pumping of Pure Superposition States,@ Phys. Rev. Lett., vol. **93**, № 16, pp. 160802_1-160802_4, October 2004.

[11] T. Zanon, S. Tremine, S. Guerandel, et.al., Observation of Raman-Ramsey fringes with optical CPT pulses,@ IEEE Trans. Instrum. Meas., vol. **54**, № 2, pp. 776-779, April 2005.

[12] N.D. Bhaskar, Potential for Improving the Rubidium Frequency Standard with a Novel Optical Pumping Scheme Using Diode Lasers,@ IEEE Trans. Ultrason. Ferroelectr. Frequency control, vol. **42**, № 1, pp. 15-20, January 1995.

[13] J. Deng, Light Shift Compensation in a Rb Gas Cell Frequency Standard with Two-Laser Pumping,@ IEEE Trans. Ultrason. Ferroelectr. Frequency control, vol. **48**, № 6, pp. 1657-1661, November 2001.

[14] J.Vanier, C.Audoin, The Quantum Physics of Atomic Frequency Standards, Bristol: Adam Hilger, UK, 1989.

DESIGN OF GaAs RESONANT STRUCTURES FOR SENSORS APPLICATIONS

T.G, LEBLOIS, C.R. TELLIER
Institut FEMTO-ST
Département LCEP, ENSMM
26 chemin de l'épitaphe, 25030 Besançon cedex, France
email: tleblois@ens2m.fr Fax: +33 3 81 88 57 14

ABSTRACT

This paper deals with GaAs resonant structures. An analytical model is used to design the resonator vibrating in flexural mode. Theoretical cross sectional shapes are derived from the simulator TENSOSIM. Frequency changes and modifications in cross sectional shape induced by small misalignment remain negligible.

1. INTRODUCTION

Recently new mechanical sensors based on piezoelectric effects in GaAs have been investigated [1, 2]. The possibility of fabricating mechanical devices in GaAs by bulk micromachining sustains the interest in these resonant sensor applications. This paper gives more complete information on the metrological performances and on the micromachining of 3D GaAs structures in order to design new resonators on (hhl) GaAs substrates.

2. THEORY OF PIEZOELECTRIC RESONATOR

Some previous works on GaAs resonant structures deal with shear, longitudinal and flexural modes of vibration [1, 2]. But, these studies are always limited to (100) and (110) wafers. We propose in this paper to give some metrological performances of GaAs resonators fabricated on (hhl) planes and vibrating in flexural mode. The study is performed using an analytical modelisation. The possibility of piezoelectrically activating different modes of vibration is determined by the rotated piezoelectric coefficient d'_{ik}. In the crystallographic coordinates system, only one independent coefficient differs from zero: d_{14}= -2.69 pC/N.
We remark that this coefficient is comparable to the d_{11} constant of the commonly used quartz crystal for piezoelectric applications (d_{11}= -2.3 pC/N). The resonator is a built-in-free beam whose cross section orientation is shown in figure 1.

2.1 Piezoelectric excitation:

Let us recall that the flexural mode can be split up into traction and compression on both sides of the neutral fibre. Consequently, the flexural vibration can be activated by the components of $[d'_{ik}]$ that act in the longitudinal mode: d'_{21} and d'_{31}. Suitable electrode geometries are given in figure 1. A large value of d'_{21} gives an in-plane flexural vibration in the second configuration of electrode (Fig. 1a) and an out-of–plane vibration in the first configuration (Fig. 1b).

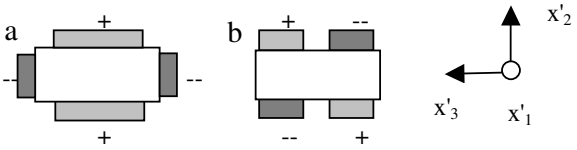


Fig.1 Two configurations of electrodes

In the same way, a large value of d'_{31} induces an out-of-plane vibration in the case 1b and an in-plane vibration in the case 1a. In order to obtain pure flexural mode, the piezoelectric coefficients d'_{2k} and d'_{3k} with $k\neq1$ must remain low with respect to d'_{21} and d'_{31} respectively. Figure 2 shows polar plots of d'_{21} and d'_{31} (d'_{i1} versus the in-plane azimuthal angle ψ) on (115) and (441) planes.

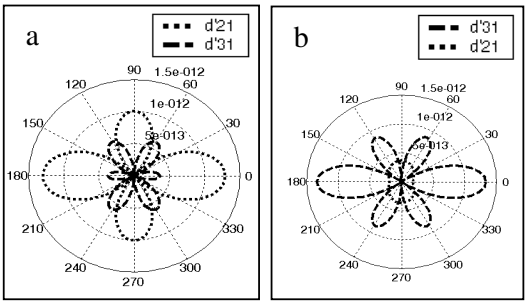


Fig.2 d'_{ij} polar plots on (115) plane (case a) and (441) plane (case b)

Figure 2a shows that the coefficient d'_{21} is large for ψ=0 (d'_{21}=1.26 pC/N) and for ψ=90° (d'_{21}=1 pC/N). These alignment directions are particularly suitable for exciting flexural vibrations. Moreover, for both alignment angles d'_{31} exhibits low values and d'_{25}, d'_{35}, d'_{26} and d'_{36} are equal to zero. Figure 2b shows that d'_{31} firstly takes greater values than d'_{21} and secondly presents a maximum for ψ=0 (d'_{31}=1.3 pC/N). For this orientation, calculations give for d'_{25}, d'_{35}, d'_{26} and d'_{36} values equal to zero which means that torsional mode is not excited. According to above results, we retain the following orientations:

(i) plane (115) and alignment ψ=0 (direction <110>)

(ii) plane (441) and alignment ψ=0 (direction <110>)

2.2 Resonance frequency:

The size of the beam allows us to use Bernoulli-Euler model. The equations of equilibrium lead to the following relation:

$$\frac{\iint x_{2(3)}^2 dA}{\rho A s_{11}} u_{2(3),1111} + \ddot{u}_{2(3)} = 0 \qquad (1)$$

where A represents the surface of the cross-section and ρ the density.

The displacement $u_{2(3)}$ can be expressed as a product of two functions:

$$u_{2(3)}(x_1, t) = g(x_1) f(t) \qquad (2)$$

Introducing equation (2) in equation (1) and taking account the boundary conditions of the beam, we determine the resonant frequency f_R:

$$f_{Rn} = (\beta_n \lambda)^2 \frac{b(a)}{2\pi\lambda^2} \sqrt{\frac{1}{3\rho s'_{11}}} \qquad (3)$$

where λ is the length, 2b the thickness, 2a the width of the beam and n the partial rank.

It may be of interest to follow the evolution of resonance frequency firstly with the temperature T and secondly with a misalignment Δψ of the beam. In the expression of f_R, three terms depend on T: the compliance, the density and the geometrical parameters. The dilatation tensor is isotropic. Consequently the angles of cut do not change with T. Three different values of temperature coefficients Ts_{ij} are introduced in the frequency response. The dependence of resonance frequency is shown in figure 3 in the case of (115) and (441) planes.

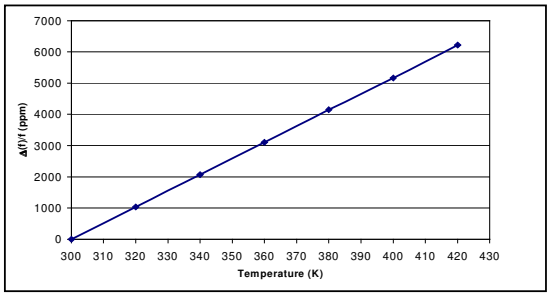


Fig.3 Effect of the temperature on the resonance frequency

Figure 3 reveals a linear evolution of $\Delta f_R/f_R$ with T. The slope of the linear plot is estimated at 51 ppm/K. In conclusion, we can say that the variation of f_R is only slightly dependent on temperature (deviation less than 0.1% for ΔT=20K).

A misalignment Δψ of the beam often appears during the process of photolithography. So, it is of interest to know the deviation of f_R with Δψ. Formulation (3) is used to obtain the numerical results presented in figure 4.

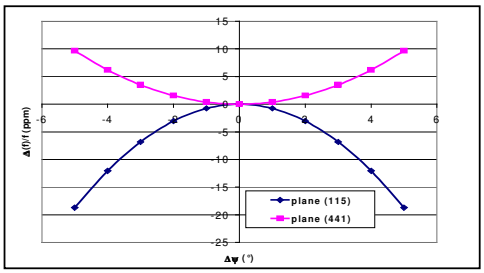


Fig.4 Effects on the resonance frequency of a misalignment Δψ

Curves of figure 4 exhibit non linear variations with a deviation of f_R with Δψ that is twice greater for (115) beam than for (441) beam. But deviations remain negligible for any orientations.

2.3 Metrological performances:

A main parameter of the resonator is the electromechanical coupling coefficient k. It is determined using the expression of elastic U_m, piezoelectric U_p and electric U_e density of energy [3]. U_e and Up are given by:

$$U_e = \frac{\varepsilon V_0^2 C_0}{8ab}$$

$$U_p = \frac{1}{8ab} \iint_A d'_{2(3)1} T_1 \frac{\partial V}{\partial x_{2(3)}} dx_2 dx_3 \qquad (4)$$

In these expressions, V is the electrostatic potential of the beam, V_o the applied potential at the electrodes and C_0 the normalized static capacitance. Moreover an equivalent circuit is generally used to describe the piezoelectric performances of resonator near the resonance frequency. It is well known that for low coupling materials such as GaAs the dynamic electromechanical coupling factor is nearly equal to the ratio of the motional capacitance (C_m) to the static one (C_0). The ratio V/V_0 is determined using conformal transformations [4]. The microfabrication process is more easier when we adopt an in-plane electrodes configuration. We suppose that the electrode configuration is symmetric and we call g the gap between the in plane electrodes and e the width of one electrode. The size parameters of the beam are supposed to be: 2b= 200µm and 2a=400µm. So, figures 5 and 6 display the evolution of the electromechanical coefficient and the motional capacitance in the case of in plane electrodes configuration.

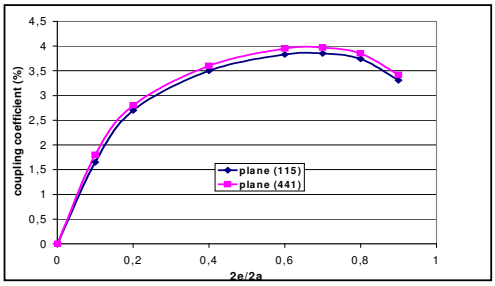


Fig.5 Evolution of k with the width of electrodes

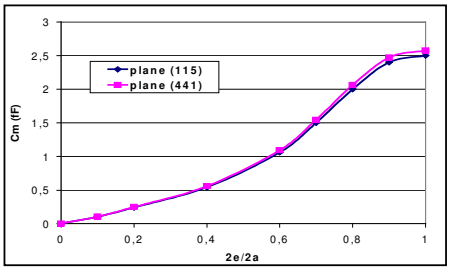


Fig.6 Evolution of the motional capacitance with the width of electrodes

The coupling coefficient passes through a maximum value for a width of electrode 2e=280µm. We observe that this value increases with the ratio b/a. The motional capacitance increases with 2e and equal 1.4 fF in the case of 2e=280µm. According to the above results, an optimisation of the resonator design can be performed.

3. ETCHED SHAPES

3.1 Micromachining

Anisotropic etching of crystals is a technique for bulk micromachining of mechanical devices such as micro-resonators. Silicon is the most used crystal. But etchings of GaAs crystal in H_2SO_4: H_2O_2: H_2O systems are also found [5, 6] to be anisotropic. Recently [5] the $1H_2SO_4$:$1H_2O_2$:$8H_2O$ etchant has been extensively studied and a database composed of dissolution constants was determined from experiments. This study shows that effectively GaAs microstructures can be fabricated by wet etching even if the anisotropy ratio k is small (k≈2.4). But the non-centrosymmetric nature of GaAs complicates fabrication and analysis of {hhl} structures because the two sides of structures must have different etching shapes. New software applications are added to the simulator TENSOSIM [5] in order to derive 3D etching shapes for (hhl) cantilevers micromachined with double sided masks. The present study focuses on (441) and (115) cantilevers aligned along the above metrological directions. Let us first examine predictions for (441) cantilevers. Fig. 7 gives the cross-sectional theoretical shape of a (441) cantilever aligned along the $[1\bar{1}0]$ direction (x'_1 axis) just at the opening of the structure. As expected the structure exhibits no symmetry. Moreover the cross-section is composed of two protrusions that can affect the frequency response.

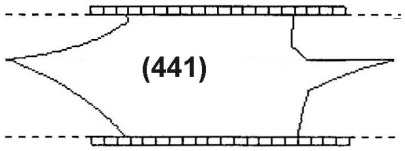


Fig.7 The theoretical shape of the (441) section

The theoretical cross-section of the other interesting orientation for resonant GaAs cantilever (i.e. (115) orientation) is shown in Fig. 8. Here again the simulator generates a cross-section with two protuberances. So that at this point the choice between the two possible orientations is not easy. Consequently it may be interesting to investigate effect of misalignment or of disorientation on the etching shape of cantilevers.

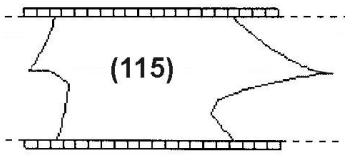


Fig.8 The theoretical shape of the (115) section

3.2. Effect of misalignment and of disorientation

Taking into account the complexity of the cross-sectional etching shape we have limited our study of the influence of disorientation and of misalignment on cross-sectional shapes at the opening to change in the width of cantilever as measured just beneath the mask. Let us first focus on (441) cantilevers and to effect of a misalignment $\Delta\Psi$ of the cantilever with respect to <110> direction. Fig. 9 illustrates changes in cross-sectional shape caused by misalignment of the mask or by small uncertainty $\Delta\theta$ in the orientation of the GaAs substrate. At this point it should be mentioned that owing to the mirror symmetry (Fig. 10) related to {110} planes misalignment of $\pm\,\Delta\Psi$ about the <110> direction lead to similar changes in shapes and in width. Consequently only cross-sections related to positive values of $\Delta\Psi$ are drawn in Figs 9A and 9B. We observe that in fact the shape of (441) cross-section varies slowly with increasing misalignment or disorientation.

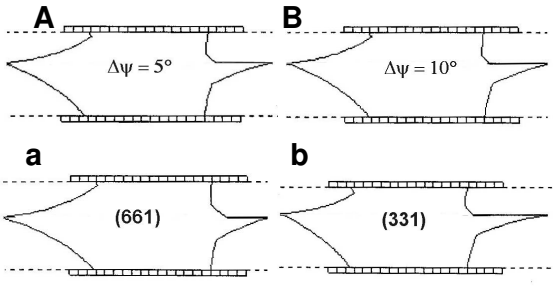


Fig. 9 Effects on etching shape of a misalignment $\Delta\Psi$ (A,B) and of a small disorientation of substrate ($\Delta\theta$<5°)

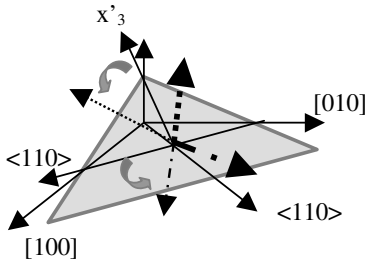


Fig. 10 The mirror symmetry associated to {110} plane

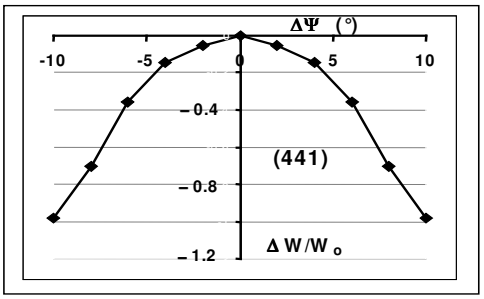


Fig. 11 Effects on the width W of a misalignment ΔΨ

The symmetric plot in the relative changes $\Delta W/W_o$ in the (441) cantilever versus the misalignment ΔΨ is displayed in Fig. 11. Apparently until the misalignment remains in the range ±5° the departure does not exceed 0.5%. Changes in the width induced by small deviations in the angle of cut θ can be depicted in Fig. 12. It may be seen that the changes $\Delta W/W_o$ are also very moderate (typically less than 2%).

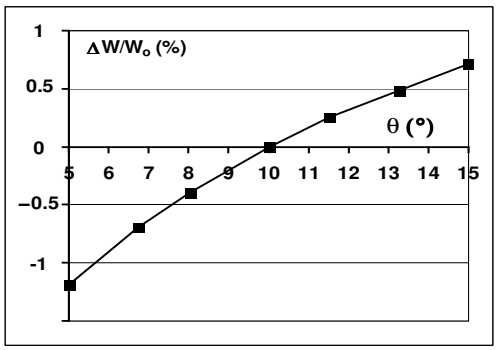


Fig. 12 The width W of the cantilever as a function of the angle of cut θ

Let us now turn attention to the currently technical problem of misalignment that can be encountered when we micromachine (115) cantilevers. Figs. 13A and 13B reveal that very slight modifications in cross-sectional shapes result from misalignment. In particular the two protrusions are still present when the misalignment reaches 10°. At this time plot of $\Delta W/W_o$ versus ΔΨ (Fig. 14) reveals that $\Delta W/W_o$ increases approximately by 2.5% when the misalignment increases from 0° to 10°.

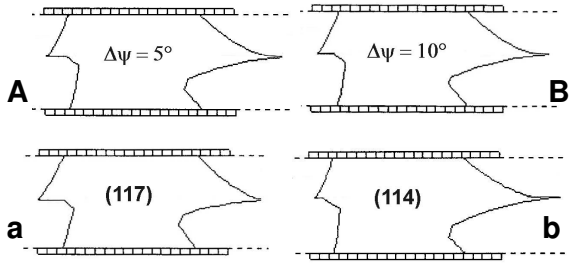


Fig. 13 Effects on etching shape of a misalignment ΔΨ (A,B) and of a small disorientation of substrate (Δθ<5°)

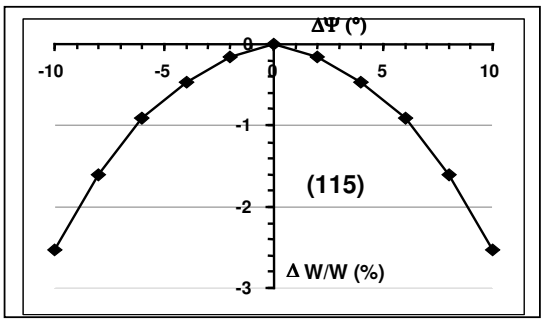


Fig. 14 Effects on the width W of a misalignment ΔΨ

Results for disorientation of the GaAs substrate about the (115) orientation are displayed in Figs. 13a, 13b and 15. Here again we observe that cross-sectional (117) and (114) etching shapes resemble to (114) shape. Consequently (Fig. 15) the width of cantilever is only slightly dependent on the angle of cut θ.

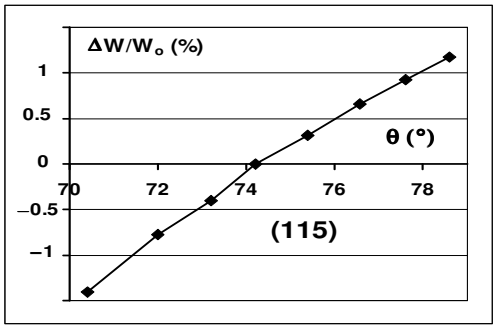


Fig. 15 The width W of the cantilever as a function of θ

4. CONCLUSION

The analytical model of piezoelectric resonator allows us to predict metrological performances of beams vibrating in flexural mode. Two orientations of substrates are retained; these orientations furnish favourable metrological performances such as electromechanical coupling coefficient and equivalent electric parameters when the cantilever is aligned along <110> direction. The model allows us also to optimise the geometry of the resonator and the associated electrodes.

Experimental and theoretical investigations on the micromachining of GaAs crystals call for the possible fabrication of (hhl) cantilevers. Effects of a misalignment or of a disorientation of substrate on structure shape and dimensions are very weak as expected for an etching process characterized by a small anisotropy ratio. Another consequence of this small anisotropy is that compensation masks are not needed. effectively micromachined resonant GaAs cantilevers are free of significant convex undercuts. However in return it not possible to affirm that protrusions formed at the opening of the structure can easily disappear with prolonged etching. So in the future we plan to develop new software to have after opening a detailed evolution

of cross-sections of (hhl) cantilevers as a function of etching time.

REFERENCES

[1] K. Fricke, Piezoelectric properties of GaAs for application in stress transducers, *J. Appl. Phys.,* **70**, pp 914-918, 1991

[2] J. Söderkvist and K. Hjort, Flexural vibrations in piezoelectric semi-insulating GaAs, *Sensors and Actuators,* **39**, pp 133-139, 1993.

[3] T. Leblois, G. Huve and C. Tellier, GaAs micro-beams for mechanical sensors, *Proc. 15th Micromechanics Europe Workshop*, Leuven, Belgium, 5-7 october 2004, pp 9-12

[4] J. Hermann, Determination of electromechanical coupling factor of quartz bars vibrating in flexure or length-extension, *Proc. 29th Annual Symposium on Frequency Control,* pp 26-34, 1975

[5] C.R Tellier, G. Huve and T.G. Leblois, GaAs micromachining : 3D etching shapes, database and Simulations, *Proc. 15th Micromechanics Europe Workshop*, Leuven, Belgium, 5-7 october 2004, pp 111-114

[6] Don W. Shaw, Localized GaAs etching with acidic hydrogen peroxide solutions, *J. Electrochem. Soc.: Solid State Sci. Technol., 128* (1981) 874-880

THE EFFECT OF DIFFERENT ELECTRODE FILMS ON THE RESONATOR PARAMETERS

V.Georgieva, L.Spassov, Ts.Yordanov
Institute of Solid State Physics, Bulgarian Academy of Sciences,
72 Tzarigradsko Chaussee Blvd, Sofia 1784, Bulgaria
Lazarova@issp.bas.bg

ABSTRACT

The aim of the investigation was defining the influence of different kinds and thickness of thin metal electrodes on the resonators' equivalent dynamic parameters and their amplitude frequency characteristics (AFCs) as well as estimating the changes of thin films' properties.

Scanning Electron Microscopy (SEM) and Transmission Electron Microscopy (TEM) carried out the morphology investigations of thin metal films. The surface morphology of Au, Al and Ag films were compared. The grain size changes of all studied films as a function of the electrodes thickness were estimated too.

The main parameters as the equivalent dynamic resistance (R_q), the quality factor (Q) and AFCs were studied as a function of the different kinds of electrodes and thickness and at the same mass loading. The best Rq and Q – factor (9,09Ω and 57583) were obtained for resonators with Ag electrodes (1200Å).

The measured resonators AFCs showed dependence from electrode thickness, but AFCs retained identical for all electrode metals at the same mass loading.

As a result of the investigations the conclusion is that, the selection of electrodes material and its thickness are of a great extent responsible for the resonators acoustic parameters. They could be affected significantly by changing the kind and the thickness of the electrodes.

1. INTRODUCTION

The thin one-component films of Al, Au and Ag were commonly used as electrodes for quartz resonators. Their properties depend on numerous conditions as: the purity of the film material, the temperature and the rate of deposition, the kind of deposition process. Thin films of various materials can be produced by evaporation of a source material in vacuum with low energy atoms (0, 1 eV). The high vacuum conditions cause a little surface damage and less residual gas.

Each electrode mentioned above has specific characteristics. Aluminium is a highly reactive metal and wasn't the preferred material for quartz resonator, because of the changes at metal-quartz interfaces. Gold is the preferred electrode material: it isn't highly reactive and doesn't form an oxide. Au adheres weakly to quartz, but enough for electrodes used. The porous gold electrode can replace commonly used electrodes [1]. Silver is known with catalytic reaction at the electrode surface. The electrode affects the quartz crystal resonator to a great extent. Surface morphology and structure partly defines modifying the physical properties of the films used in quartz crystal resonators. Scanning Electron Microscopy (SEM) and Atomic Force Microscopy (AFM) are used for morphology thin film studying and modification after deposition [2-5]. The influence of electrode changes on resonators usually is registered by measuring equivalent dynamic parameters [6-8].

The aim of the investigation is defining the influence of different kinds and thickness of thin metal electrodes on the resonator parameters as well as estimating the morphology changes of the films.

2. EXPERIMENT

The resonator structures are created by a standard technology. AT – cut quartz plate with 120μm thickness and 8mm diameter are used. After standard chemical treatment the plates are divided into 5 groups for forming Ag, Au and Al electrodes with different thickness. All electrodes are deposited by thermal evaporation using sources with purity as follow: Au and Al – 99,995%, Ag and Cr – 99, 9995%. The plates are maintained at temperature 260°C ±2°C to obtain high quality films. Cr is used as sub-layer for improving weakly adhering Au film. The deposition rate of 65Å – 70Å is achieved for all metal films. Resonators with electrode diameters 4mm and thickness Al-1200Å and 4800Å, Ag-1200Å and Au- (200Å Cr–580Å Au) and Au-(200Å Cr–1200Å Au) are created. The film thickness during the deposition is measured with high accuracy by quartz crystal monitor and after deposition by Talystep "Taylor - Hobson".

The main studied thin layer parameter is the morphology. The investigations are carried out by Scanning Electron Microscopy (SEM) on JSM 5300 with magnification of 20000 and electron beam energy of 29 kV. The resolution of the method is increased by evaporation of carbon and gold subsequently on all samples. In aim of more precisely investigation of the electrodes, all samples are analyzed by transmission electron spectroscopy on EM Philips-400. The samples are prepared for observation by method of two steps replica gelatine-carbon-platinum.

The motional parameters of resonator' structures are measured by a Selective Level Meter PSM-14 from Wander and Golderman. The equivalent dynamic resistance (R_q), and the quality factor (Q) are defined as the main for our investigation.

3. RESULTS AND DISCUSSION

3.1. Morphology and structure

The uncoated quartz surface is smooth. The metal layers' surface, determined by SEM, shows an homogenization when the Al layers are in the range: 1200 Å – 4800 Å. Such a change is not observed at Au layers – a possible reason could be the presence of Cr sub-layer. Comparing the three kinds of metal at one and the same thickness (1200 Å), it is obvious, that Au has the smoothest surface, while Ag - the roughest one. The Al surface takes the middle position between the other two. The results can be explained with highest mobility of Au atoms, followed by those of Al and Ag atoms.

TEM studies showed that Au layers are amorphous with rarely encountered micro regions with defects. These results do not depend on the layer thickness (580 Å –1200 Å) - Fig.1.

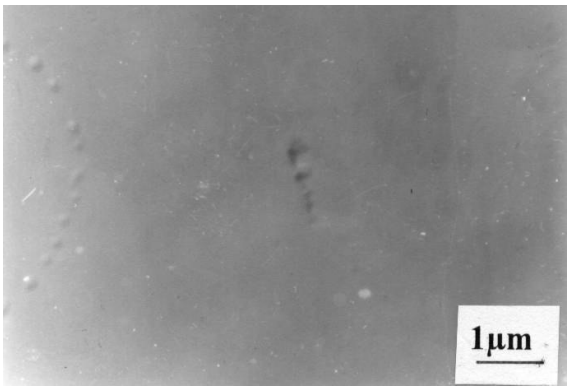


Fig. 1. Transmission electron micrograph (TEM) of the 1200 Å Au thin layer.

The structure studies of Ag electrodes show a two-phase growth. The layers are amorphous with micro crystals continuously distributed throughout the films. The size of the greater part of them is between 0.1 µm and 0.2 µm but micro crystals in the range 0.5 µm up to 2 µm are observed, too (Fig.2).

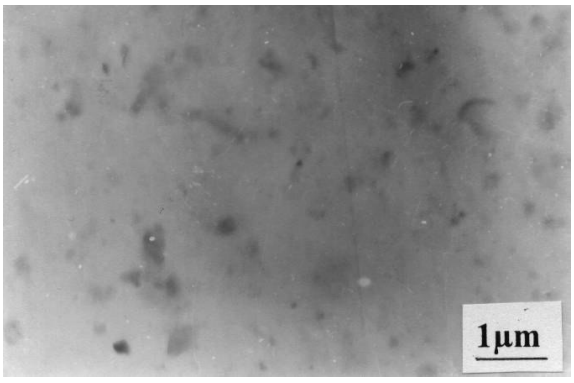


Fig. 2. Transmission electron micrograph (TEM) of the 1200 Å Ag thin layer.

Figure 3 shows a typical structure of Al layers with a thickness of 4800 Å.

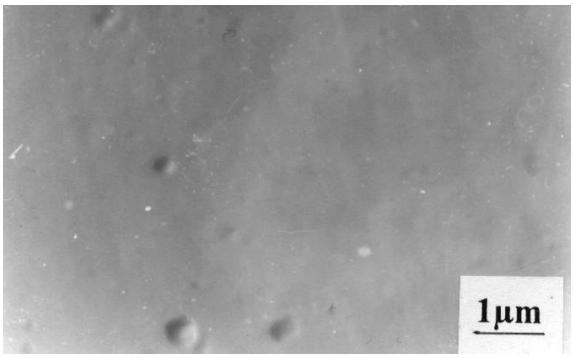


Fig. 3. Transmission electron micrograph (TEM) of the 4800 Å Al thin layer.

The thicker Al layers are amorphous with micro crystals (< 1 µm) distributed unevenly. In the case of Al growth this type of structure could be explained by hillocks generation at the Al-quartz interface, caused by the presence of free silicon atoms. Their availability is due to a chemical reaction between the silicon oxide and Al. It is well known [9] that Al possesses stronger affinity to oxygen than Si. So a replacement of Si from the SiO_2 is possible and as a result Al_2O_3 and free Si - atoms are formed. The latter could provoke the hillocks' nucleation.

SEM and TEM studies show that the surface morphology and structure of thin evaporated metal films depend on the layer's kind and thickness. These two parameters are responsible for the adhesion and long-term stability of resonators' structures.

3.2. Resonator acoustic parameters

The kind, thickness and morphology changes of electrodes may be partial responsible for resonator parameters. The measured Rq and Q parameters for different Au electrodes thickness do not show changes in the Rq average value (13, 49 and 13,63 Ω) and Q value (39 796 and 36 721) for Au thickness (200Å Cr - 580Å Au) and (200Å Cr - 1200Å Au) respectively. The same parameters are measured when Al electrodes are applied. The resistance measured for the samples with 4800Å Al thickness is 9,67Ω, which is twice lower than the thin (1200Å) Al film – 15,68Ω. The Q factor presents the same tendency - 53388 and 32846 respectively. These results support structure studies, which showed that when Al thickness increase a homogenization takes place leading to the decrease of Al layers resistivity.

Fig. 4. and Fig.5. illustrate the Rq and Q factor change as a function of the kind of the metal at the same mass loading. The resonators with Ag electrodes have the lowest resistance measured (9,09Ω); no more different is Rq of Al resonators (9,68Ω), while the Rq of Au

resonators significantly increases (13,49Ω). The Q factor values are the best ones concerning the Ag electrodes – 57583 and Al electrodes – 53388, while the Au electrodes are quite lower - 36721.

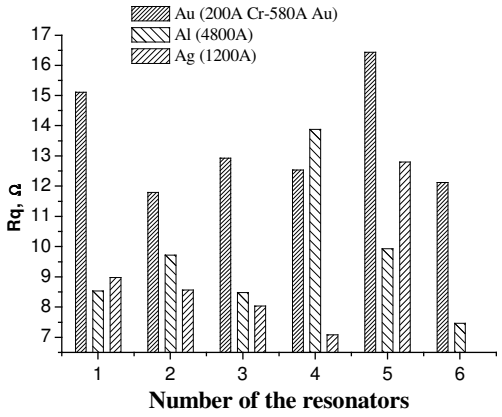


Fig. 4. Rq of the resonators with Au, Al and Ag electrodes
Au (200Å Cr - 580Å Au) Rq av. –13,49 Ω
Al 4800Å - Rq av. – 9,67 Ω
Ag 1200Å - Rq av. –9,09 Ω

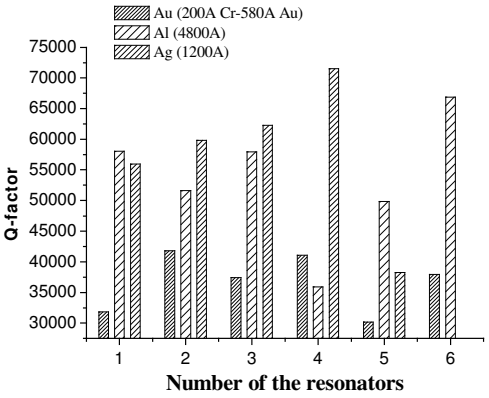


Fig. 5. Q-factor of the resonators with Au, Al and Ag electrodes
Au (200Å Cr - 580Å Au) Q av. –36721
Al 4800Å -Q av. –53388
Ag 1200Å -Q av. –57583

Conclusion is that the selection of the metal electrode type and its thickness are responsible for the resonators' Rq and Q- value to a great extent.

3.3. Amplitude frequency characteristics (AFC)

The AFC of all sample are measured and the typical ones for each group are shown. Fig.6 presents a typical amplitude frequency characteristic of quartz resonator with Au electrodes (200Å Cr - 1200Å Au).

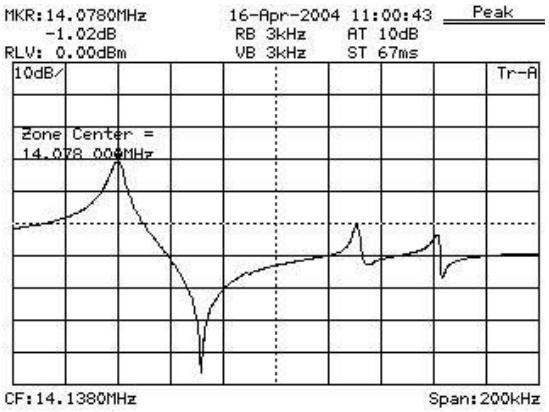


Fig.6. A typical amplitude frequency characteristic of quartz resonator - Au electrodes (200Å Cr - 1200Å Au)

Equal anharmonic couple strongly suppressed resonance is measured, 20 dB lower than the main resonance. Frequency shift of the anharmonic from the main TS mode resonance frequency resonance is 90 kHz and 120 kHz respectively. The difference between the anharmonic resonances is 30 kHz.

Figure 7 presents a typical amplitude frequency characteristic of quartz resonator with Au electrodes (200Å Cr - 580Å Au).

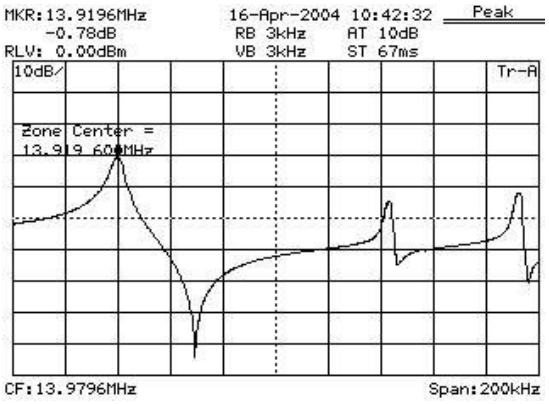


Fig.7. A typical amplitude frequency characteristic of quartz resonator – Au electrodes (200Å Cr - 580Å Au)

Compare to Fig. 6 there are more strongly expressed anharmonic resonances. The amplitude of the first anharmonic resonance is 13,9 dB lower than the main one, and the second anharmonic resonance is 11dB lower. In this case there is a less suppressed resonance (6dB) vs. the thick gold film resonators (Fig.6). The frequency sifts of the secondary resonances are 104 kHz and 154 kHz (difference 50 kHz), which is higher compare to those on Fig.6.

Figure 8 shows a typical amplitude frequency characteristic of quartz resonator with Al electrodes (4800Å). A couple of weak anharmonic resonance with the same amplitude suppressed with 18 dB vs. the main resonance is measured.

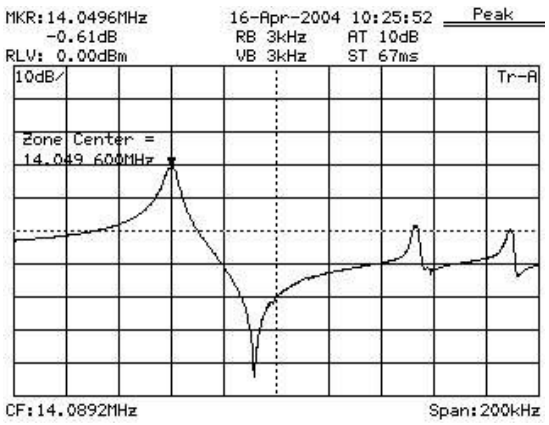


Fig.8. A typical amplitude frequency characteristic of quartz resonator – Al electrodes (4800Å)

The frequency shifts of the anharmonic resonance vs. the basic one are 91 kHz and 127 kHz (difference 36 kHz).

Figure 9 shows a typical amplitude frequency characteristic of quartz resonator with Ag-electrodes (1200Å)

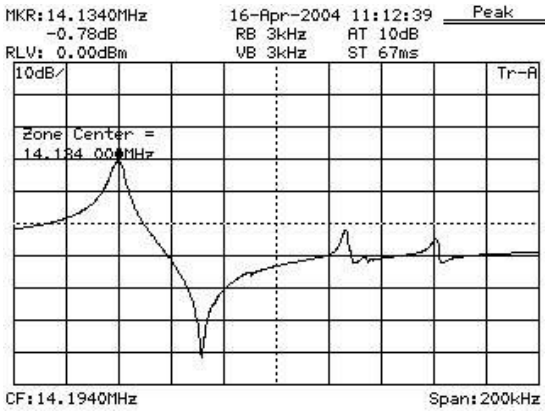


Fig.9. A typical amplitude frequency characteristic of quartz resonator – Ag electrodes (1200Å)

Comparing all AFCs it is obvious that, regardless the electrode metals the AFCs of all resonators are identical at the same mass loading (Fig.7, Fig.8 and Fig.9).
With changing the resonator mass loading, the secondary resonance can be suppressed.

4. CONCLUSION

It is shown, that surface morphology and structure of thin metal films, deposited at the same rate, is defined from the material type and thickness. The smoothest surface was obtained for the Au films. These parameters influence to a certain extent resonator characteristics.

At the same mass loading the best parameters are measured for resonators with Ag-electrodes, while the AFCs showed identical character.

Varying the types of the electrodes, their structure and thickness, the acoustical characteristics of resonators could be improved.

REFERENCES

[1] M.Hieda, R.Garcia, M.Dixon, T.Daniel, D. Allara, "Ultrasensitive Quartz Crystal Microbalance with Porous Gold Electrodes", Appl. Phys. Let., vol. 84, pp. 628-630, no. 4, 2004.

[2] J. Mizsei, V. Lantto, "In situ AFM, XRD and Resistivity Studies of the Agglomeration of Sputtered Silver Nanolayers", J. Nanopartical Res., vol. 3, pp. 271-278, no. 4, 2001.

[3] T. Nabi, H. Sambe, D. Ramaker, "AFM Study of Topographical Changes on Aluminum Surfaces in Sulfuric Acid under Low Current Anodic Dissolution", J. Electroanalytical. Chem., vol. 501, pp. 33-40, no. 1-2, 2001.

[4] H.W. Schumaher, B. Kracke, B. Damaschke, "Modification of Thin Gold Films with a Scanning Forse Microscope", Thin Solid Films, vol. 264, pp. 268-272, no. 2, 1995.

[5] S.Kundu, S. Hazra, S. Banerjee, M.K. Sanjal, S.K. Mandal, S. Chaudhuri, A.K. Pal, "Morphology of Thin Silver Film Grown by DC Sputtering on Si (001)", J. Appl. Phys., vol. 31, pp. L73-L77, no. 4, 1998.

[6] V. Georgieva, D.B. Dimitrov, G. Beshkov, L. Spassov, "Influence of Vacuum Rapid Thermal Annealing on the Sheet Resistance of thin Al Films Deposited on Quartz Crystal", Proc. of 14-th EFTF, Torino, Italy, pp. 225-228, March 14-16, 2000.

[7] V. Georgieva, L. Spassov, Ts. Angelov, M. Marinov, "The Effect of Rapid Thermal Annealing on Quartz Resonators", Proc. of 16-th EFTF, St. Peterburg, Russia, pp.C-061-C-064, March 12-14, 2002.

[8] V. Lazarova, Ts. Yordanov, L. Spassov, G. Beshkov, "Vacuum Thermal Annealing of Quartz Resonators", Vacuum, vol. 69, pp.379-383, 2003.

[9] E.P.Trifonova, V. Lazarova, L. Spassov, N. Efremova, "Composite Microhardness of Quartz-Al Structures", Cryst. Res. Technol., vol. 34, no. 3, pp.391-395, 1999.

EFFECT OF HYDROGEN FLUX ON LOOP CONTROL MODEL IN PASSIVE HYDROGEN MASER

Q. Wang [1], G. Busca [2], F. Droz [3], P. Rochat [3]
[1] Observatory of Neuchatel, l'Observatoire 58, 2000 Neuchâtel, Switzerland
[2] Kytime Sarl, Ecluse 11, 2022 Bevaix, Switzerland
[3] Temex Neuchâtel Time SA, Vauseyon 29, 2000 Neuchâtel, Switzerland

ABSTRACT

This paper introduces the flux-dependent Spin Exchange (SE) shift and broadening, together with the cavity pulling effect into the loop model of the Passive Hydrogen Maser (PHM), completing the previous study given in [1]. In contrast to the Active Hydrogen Maser (AHM), the PHM frequency at SE tuning is slightly dependent on the H flux, even neglecting the spin exchange parameter ε_H. In addition it depends upon typical parameters as the interrogation power and the phase selection for the Ultra Stable Oscillator (USO) and Auto Cavity Tuning (ACT) loops.

For some phase selection and at low H flux the model shows that the maser output frequency could be independent from the cavity frequency or even decrease by increasing the cavity frequency. The zero cavity pulling and the negative cavity pulling situation is confirmed by the experiment.

1. INTRODUCTION

Previous studies performed on the PHM [1,2] have not addressed the effect of the flux variation on USO and ACT loops. Even if such effects are relatively small, as shown in the present study, they are of practical importance and start to pay a sensible role at the level of the stability of the Galileo Space Passive Hydrogen Maser (S-PHM) which has demonstrated a flicker floor below 1×10^{-14} [3].

2. FLUX-DEPENDENT PARAMETERS

For PHM based on the 'Single frequency modulation, phase discrimination' approach, in the conventional operational phase selection $\phi_d=\arctan\left(\frac{\beta}{\omega_m}\right)-\frac{\pi}{2}$ (max H error signal in USO loop), the maser frequency offset from atomic resonance frequency is determined by the cavity pulling effect [1]:

$$\omega-\omega_0=\frac{\alpha\beta^2+\omega_m^2(1+S_0)}{\alpha(\beta^2+\omega_m^2)}\frac{\gamma_2}{\beta}(\omega_c-\omega) \qquad (1)$$

where, ω: Interrogation angular frequency, i.e. maser oscillation frequency

ω_0: Actual atomic H hyperfine transition angular frequency

ω_c: Cavity resonance angular frequency

ω_m: Modulation angular frequency

β: Half cavity line width in rad/sec

γ_2: Transverse or coherence relaxation rate for zero interrogation power

α: Oscillation parameter

S_0: Saturation factor at resonance

The hydrogen flux plays a significant role in PHM: the actual hyperfine transition frequency ω_0 is affected by the SE shift, the line width is broadened by SE effect, the oscillation parameter α and the saturation factor at resonance S_0 are both functions of the hydrogen density.

Hereafter the magnetic inhomogeneity shift and broadening are assumed to be negligible, which is close to our experimental situation.

2.1 Spin Exchange frequency shift

Taking into account of the SE shift $\Delta\omega_e$:

$$\omega-\omega_0'=\frac{\alpha\beta^2+\omega_m^2(1+S_0)}{\alpha(\beta^2+\omega_m^2)}\frac{\gamma_2}{\beta}(\omega_c-\omega)+\Delta\omega_e \qquad (2)$$

ω_0': Hyperfine transition angular frequency not shifted by SE.

For PHM with fast frequency modulation, the SE shift is determined by [4]:

$$\Delta\omega_e=-\frac{(\bar{\lambda}_{ex}+2\varepsilon_H\sigma_{ex})\hbar\bar{v}_r}{4\mu_0\mu_B^2\eta' Q_c}\frac{\alpha}{1+S_0}\gamma_2+2\varepsilon_H\gamma_{2e} \qquad (3)$$

where, σ_{ex}: H spin-exchange cross-section

$\bar{\lambda}_{ex}$: H Spin-exchange frequency shift cross section

ε_H: Frequency shift parameter related spin-exchange collisions

μ_0: Magnetic permeability of vacuum

μ_B: Bohr magnetron

$\hbar$: Planck's constant divided by 2π

η': Filling factor of the bulb (geometric constant)

$\bar{v}_r$: Mean relative velocity of hydrogen atoms

Q_c: Loaded quality factor of the cavity, $Q_c=\frac{\omega_c}{2\beta}$

γ_{2e}: Transverse SE relaxation rate

The $\bar{\lambda}_{ex}$ related term is dominant and more than 10 times bigger than the ε_H related term which is

associated to the interruption of the hyperfine interaction during SE collisions between hydrogen atoms.

2.2 Spin Exchange broadening contribution to the line width

The line width is contributed by several compositions:

$$\gamma_2=\gamma_b+\gamma_{2w}+\gamma_{2e} \quad (4)$$

$$\gamma_1=\gamma_b+\gamma_{1w}+\gamma_{1e} \quad (5)$$

Where,

γ_1 : Unsaturated longitudinal or hydrogen population difference relaxation rate

$\gamma_b+\gamma_{1w}$ and $\gamma_b+\gamma_{2w}$ are limiting longitudinal and transverse relaxation rates at zero flux. They do not depend on atomic flux and are constant for a specified storage bulb.

γ_b : Storage or geometric relaxation rate

γ_{1w}, γ_{2w}: Longitudinal and transverse wall relaxation rate respectively, due to atomic collisions with the wall of the storage bulb.

$$\gamma_{1w} \cong \frac{4}{3}\gamma_{2w} \quad (6)$$

γ_{1e}, γ_{2e}: Longitudinal and transverse spin exchange relaxation rate respectively, dependent on atomic flux Ψ

$$\gamma_{2e}=\frac{\sigma_{ex}\bar{v}_r}{V_b\gamma_b}\Psi \quad (7)$$

$$\gamma_{1e}=2\gamma_{2e} \quad (8)$$

Ψ : H atomic flux entering the cavity in the (1,0) hyperfine state

V_b : Volume of the storage bulb

2.3 Oscillation parameter and saturation factor at resonance

The oscillation parameter α is function of flux, and the saturation factor at resonance S_0 is the function of flux and interrogation power:

$$\alpha=\frac{\mu_0\mu_B^2\eta' Q_c\Psi}{\hbar V_b\gamma_1\gamma_2} \quad (9)$$

$$S_0=\frac{k_c}{\gamma_1\gamma_2}P_{in} \quad (10)$$

Where,

k_c : Constant independent of the atomic flux

P_{in} : Interrogation power

For the Engineering Model (EM) of the Galileo S-PHM, the following physics parameters are obtained from the experiment and analysis :

Q_c=6730, V_b=0.4 liter, η'=0.5, γ_b=2.9 rad/s, γ_{2w}=2.3 rad/s, ω_m=78.5 krad/s, $k_c=4.9\times10^{12}$.

Therefore the flux or interrogation power dependent parameters are deduced as following:

$$\gamma_1=5.86+1.48\times10^{-12}\Psi \quad (11)$$

$$\gamma_2=5.11+7.39\times10^{-13}\Psi \quad (12)$$

$$\alpha=\frac{8.73\times10^{-12}\Psi}{(5.11+7.39\times10^{-13}\Psi)(5.86+1.48\times10^{-12}\Psi)} \quad (13)$$

$$S_0=\frac{4.88\times10^{12}P_{in}}{(5.11+7.39\times10^{-13}\Psi)(5.86+1.48\times10^{-12}\Psi)} \quad (14)$$

$$\Delta\omega_e=-5.93\times10^{-16}\Psi-\frac{9.25\times10^{-14}\Psi}{(5.86+1.48\times10^{-12}\Psi)\left(1+\frac{4.88\times10^{12}P_{in}}{(5.11+7.39\times10^{-13}\Psi)(5.86+1.48\times10^{-12}\Psi)}\right)} \quad (15)$$

Under the nominal operation condition and the interrogation power of −82dBm, $\Psi=5.6\times10^{12}$ atoms/s, α=0.37, S_0=0.43, γ_2=9.2 rad/s, $\Delta\omega_e$=-0.033 rad/s corresponding to a relative frequency shift of -3.7×10^{-12}.

3. SPIN EXCHANGE CROSSING POINT

The following calculation is performed considering two different arbitrary values of the atomic flux, marked respectively by subscripts 1 and 2 and assuming that a single value of the cavity resonant frequency exists for which the maser oscillation frequency ω does not depend on the atomic beam flux.

From the general cavity pulling equation:

$$\omega-\omega_0'=F(\omega_c-\omega)+\Delta\omega_e \quad (16)$$

The two maser oscillation frequencies ω_1 and ω_2 have a common value of ω_t when the cavity is tuned to ω_{c_SE} :

$$\omega_{c_SE}=\omega-\frac{\Delta\omega_{e1}-\Delta\omega_{e2}}{F_1-F_2} \quad (17)$$

For the AHM, the cavity pulling factor is:

$$F=\frac{\gamma_2}{\beta} \quad (18)$$

The spin exchange shift coupled with the condition for self-sustained maser oscillations [5] is given by

$$\Delta\omega_e=\frac{-0.29\times2\bar{v}a_0^2\hbar}{\mu_0\mu_B^2Q_c\eta'}\gamma_2 \quad (19)$$

$\bar{v}$: Mean velocity of hydrogen atoms, $\bar{v}_r=\sqrt{2\bar{v}}$

a_0 : Radius of the first Bohr orbit

In this approximation corresponding to neglecting the ε_H related term, both F and $\Delta\omega_e$ are in direct proportion to the flux-dependent γ_2, the quantity $\gamma_{21}-\gamma_{22}$ factorizes on the right side of Eq. (17) and we obtain

$$\omega_{c_SE}=\omega+\frac{0.29\times 2\bar{v}a_0^2\hbar\beta}{\mu_0\mu_B^2 Q_c\eta'} \qquad (20)$$

and $\omega_t=\omega_0'$

The results show that, with the assumption made, the cavity SE tuning point ω_{c_SE} is independent of the atomic density. For an AHM the various cavity pulling curves at different fluxes cross at the single point representing the unshifted oscillation frequency ω_0'.

However, in PHM the cavity pulling factor is

$$F=\frac{\beta^2+\omega_m^2\left(\frac{1+S_0}{\alpha}\right)\gamma_2}{\beta^2+\omega_m^2}\frac{}{\beta} \qquad (21)$$

Even omitting the negligible ε_H term of Eq. (3), the spin exchange shift is

$$\Delta\omega_e=-\frac{\bar{\lambda}_{ex}\hbar\bar{v}_r}{4\mu_0\mu_B^2\eta' Q_c}\frac{\alpha}{1+S_0}\gamma_2 \qquad (22)$$

so Eq. (17) becomes

$$\omega_{c_SE}=\omega-K_{wf}\frac{\frac{\alpha_1}{1+S_{01}}\gamma_{21}-\frac{\alpha_2}{1+S_{02}}\gamma_{22}}{\frac{1+S_{01}}{\alpha_1}\gamma_{21}-\frac{1+S_{02}}{\alpha_2}\gamma_{22}} \qquad (23)$$

and $\omega_t\neq\omega_0'$

where K_{wf} is a constant for a certain physics package,

$$K_{wf}=-\frac{\bar{\lambda}_{ex}\hbar\bar{v}_r\,\beta\left(\beta^2+\omega_m^2\right)}{4\mu_0\mu_B^2\eta' Q_c\omega_m^2} \qquad (24)$$

Eq. (23) is a complex function of flux-dependent parameters γ_2, α and S_0, and the flux cannot be factorized. Therefore for PHM the crossing of various cavity pulling lines at different fluxes is not at a single point, although the spread is relatively small.

Taking the example of S-PHM EM values at the interrogation power of –82dBm for three fluxes, i.e. the purifier currents of 3.5A, 4.0A and 4.5A, which correspond the low H_2 pressure of 0.03,0.05 and 0.07 mBar, and the useful H atomic flux of 3.5×10^{12}, 5.6×10^{12} and 7.6×10^{12} atoms/s, respectively (see Fig. 1), from Eq. (2) we obtain

For Ipur=3.5A: $\omega-\omega_0'=-0.0236+0.0000121\cdot(\omega_c-\omega)$

For Ipur=4.0A: $\omega-\omega_0'=-0.0328+0.0000143\cdot(\omega_c-\omega)$

For Ipur=4.5A: $\omega-\omega_0'=-0.0397+0.0000167\cdot(\omega_c-\omega)$

As shown in Fig. 2, the SE cavity tuning point is at $\omega_{c_SE}-\omega=3487$ rad/s (555Hz) with the spread of 528 rad/s (84Hz), and the corresponding maser oscillation frequency ω_t is slightly offset by 0.018 +/- 0.008 rad/s from the value of the unshifted atomic transition frequency ω_0'.

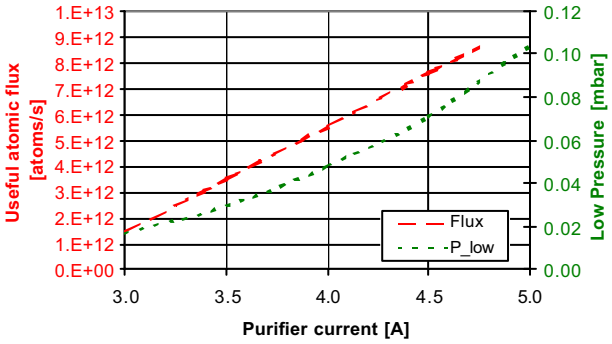
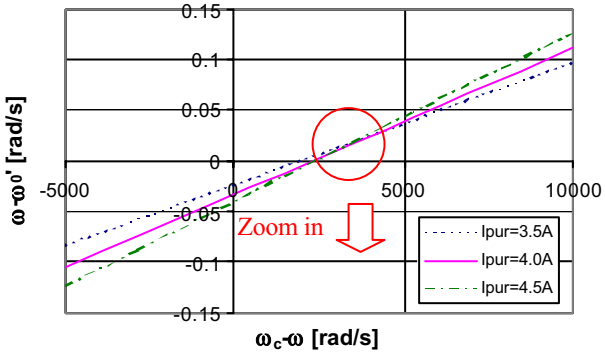


Fig. 1 Atomic flux and H_2 low pressure vs. purifier current for S-PHM EM



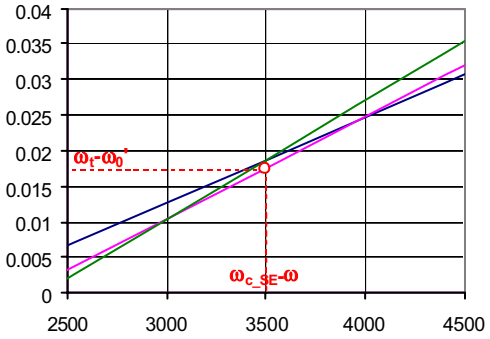


Fig. 2 Maser frequency offset vs. cavity frequency detuning for different purifier current

4. AUTO CAVITY TUNING POINT

In a PHM the microwave cavity is tuned by the ACT to the interrogation frequency, i.e. $\omega_{c_ACT}-\omega=0$. Therefore in our case the ACT tuning point is smaller than the 'SE tuning point' by approximately 555 Hz. (as shown in Fig. 2).

According to Eq. (16),

$$\omega_{c_ACT}=\omega=\omega_0'+\Delta\omega_e \qquad (25)$$

ACT tuning point ω_{c_ACT} varies with the SE shift $\Delta\omega_e$, which is dependent on the atomic flux and the interrogation power (see Fig. 3,4). The effect to ACT tuning point is negligible.

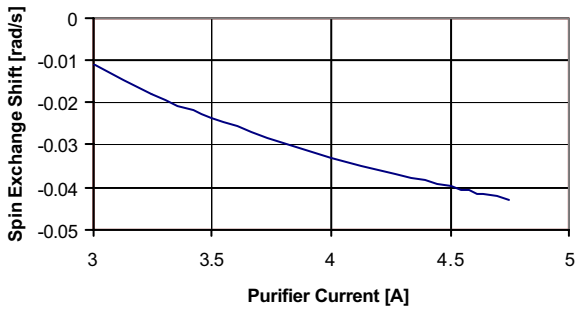


Fig. 3 SE frequency shift vs. purifier current for P_{in} =-82dBm

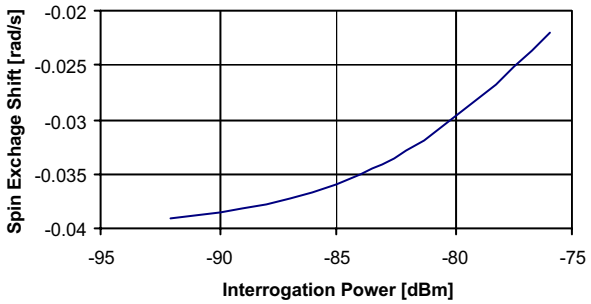


Fig. 4 SE frequency shift vs. interrogation power for Ipur=4A

5. CAVITY PULLING FACTOR

As shown in Eq. (21), for the conventional phase selection, the sign of the cavity pulling factor is always positive for any level of the flux.

Taking into account of the atomic flux impact on γ_2, α and S_0 (which is also dependent on the interrogation power) determined by Eq. (12-14), F is the function of atomic flux Ψ and the interrogation power P_{in}.

Fig. 5 gives the variation of the cavity pulling factor against the purifier current, for different values of the interrogation power.

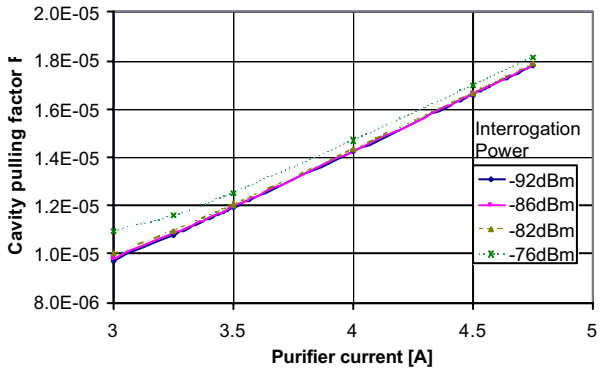


Fig. 5 Cavity pulling factor vs. purifier current for different interrogation power

In case the phase shift in H loop is $\phi_a+\Delta\phi$ ($\Delta\phi$: Phase offset) [1], the cavity pulling factor relative error is:

$$\frac{F'}{F}=\frac{\sec(\Delta\phi)}{\sqrt{\beta_2+\omega_m^2}\left[\alpha\beta_2+\omega_m^2(1+S_0)\right]}\left\{\omega_m\left[(2\alpha-1)\beta_2+\omega_m^2\right]\cos\left[\Delta\phi+\arctan\left(\frac{\beta}{\omega_m}\right)\right]\right.$$
$$+\beta\left[2\omega_m^2+\alpha\left(\beta^2-\omega_m^2\right)\right]\sin\left[\Delta\phi+\arctan\left(\frac{\beta}{\omega_m}\right)\right]$$
$$\left.+\omega_m S_0\left[\left(\omega_m^2-\beta^2\right)\cos\left[\Delta\phi+\arctan\left(\frac{\beta}{\omega_m}\right)\right]+2\beta\omega_m\sin\left[\Delta\phi+\arctan\left(\frac{\beta}{\omega_m}\right)\right]\right]\right\} \quad (26)$$

Replaced with typical physics parameter values, $\frac{F'}{F}=K$ is a function of the interrogation power, flux and phase offset.

At nominal Pin=-82dBm, K varies with the phase offset for different purifier currents (shown as Fig. 6):

$$K_{_3.25A}=1+0.387854\ \mathrm{Tan}[d\phi]$$
$$K_{_3.5A}=1+0.313778\ \mathrm{Tan}[d\phi]$$
$$K_{_4A}=1+0.261692\ \mathrm{Tan}[d\phi]$$
$$K_{_4.5A}=1+0.252738\ \mathrm{Tan}[d\phi] \quad (27)$$

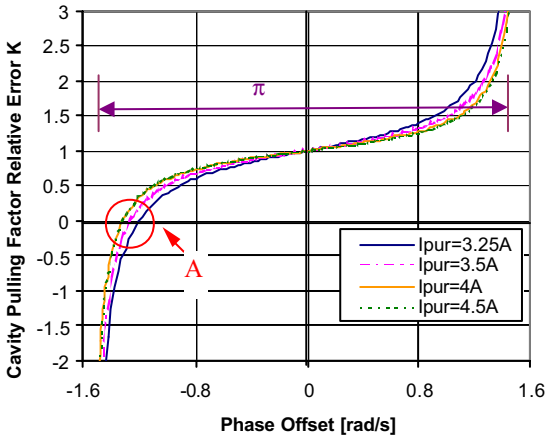


Fig. 6 Cavity pulling factor relative error vs. phase offset for different purifier current at P_{in} =-82dBm

Fig. 6 shows that

1. At $\Delta\phi$ =0, i.e. phase selection of ϕ_a, the cavity pulling is F.
2. When $\Delta\phi=\pm\frac{\pi}{2}$, the cavity pulling is infinite, which indicates that the H error signal is null for this particular phase and only a cavity error signal is present in USO loop. This is in agreement with the design of the loop phase selection: ϕ_a selected for obtaining a maximum H error signal in USO loop, and the ACT loop phase shifted by 90° from ϕ_a for obtaining a pure cavity error signal in the cavity loop.
3. Region A of Fig. 6 shows that for every hydrogen flux there exists a certain phase offset, for which the cavity pulling is zero, i.e. a pure H error signal is obtained in USO loop.
4. Also in region A near the zero cavity pulling, for a certain phase offset, the cavity pulling factor goes from positive to negative when the flux decreases.

Fig. 7 shows the variation of the cavity pulling factor relative error against the purifier current, for the nominal interrogation power and $\Delta\phi$ =-1.24 rad/s.

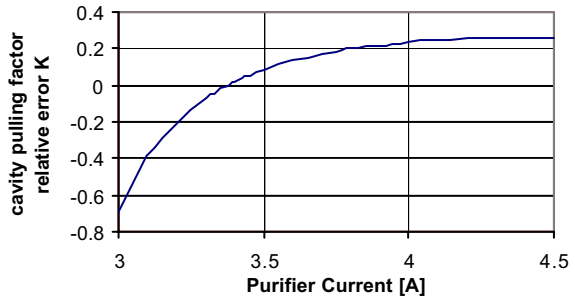


Fig. 7 Cavity pulling factor relative error vs. purifier current at P_{in} =-82dBm and $\Delta\phi$ =-1.24 rad/s

6. EXPERIMENTAL RESULTS

Fig. 8 gives measurement results performed on S-PHM EM in a configuration different from the one to which refer the previous calculation reported in Fig. 2. It shows that the cavity pulling factor, represented by the slope of the curve, decreases from positive values to zero and to negative values by decreasing of the H flux. For the purifier current of 3.25A, the cavity pulling is negative. The measurement also shows that in the present experiment the difference of SE tuning and ACT tuning is 1.3 +/- 0.3 kHz. These results confirm the validity of the loop model and in particular the possibility of obtaining zero cavity pulling in the USO loop.

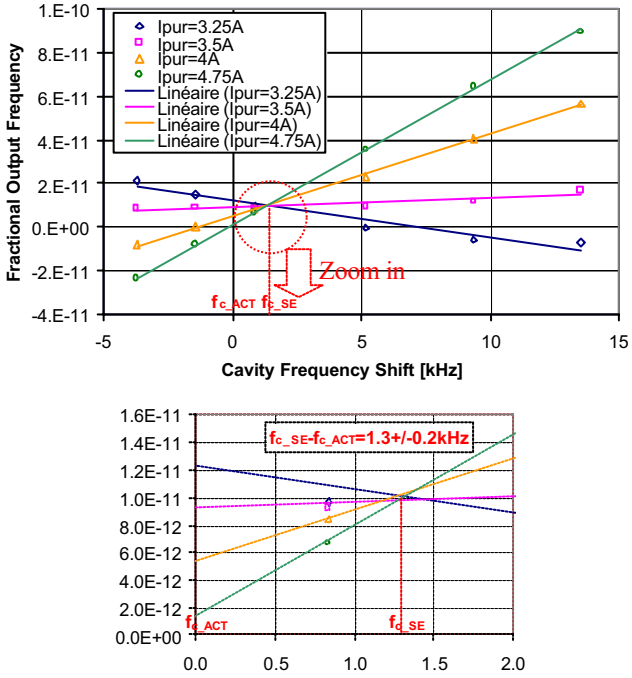


Fig. 8 Fractional output frequency vs. cavity frequency shift for different purifier current obtained from the experiment and their linear fits

7. CONCLUSION

The present study has demonstrated that the SE tuning in PHM is basically different than in the AHM due to the passive mode of the operation. In addition typical parameters such as interrogation power, error signal phase selection for the H and cavity loops, as well as the H flux, are playing a role which is no more negligible for the stability actually reached by the Galileo S-PHM. The loop model selected for this PHM is completed with the introduction of these small effects.

The model is validated by the experimental results. It shows that the SE tuning point is slightly dependent upon the H flux and in particular that it is possible to select the phases of the H loop error signal for reducing or even nullifying the cavity pulling effect in the H loop. The study of the practical advantages of this phase selection is out of the scope of the present work. Due to the implication on the cavity and H error signal levels of the phase selection, a careful assessment of the zero cavity pulling technique requires the determination of the resulting trade-off between the short and long term stability - which is possible only with an ad hoc measurement campaign.

REFERENCES

[1] G. Busca, Q. Wang, M. Belloni and L. Mattioni, Cavity Pulling in Galileo Passive Hydrogen Maser, Proceeding of the 2003 IEEE International Frequency Control Symposium Jointly with the 17th European Frequency and Time Forum, pp. 86-89.

[2] S. A. Kozlov, V. A. Logachev, Influence of the Various Factors on the Passive Hydrogen Maser Frequency Instability, 1997 IEEE Frequency Control Symposium, pp. 286-294.

[3] F. Emma, F. Droz, P. Pochat and L. Mattioni, Onboard Clocks for Galileo Development Status, International Workshop on Galileo Time, Torino, Italy, Sep. 27-28, 2004.

[4] J. Vanier and C. Audoin, The Quantum Physics of Atomic Frequency Standards, Adam Hilger, Vol. 2, 1988.

[5] J. Vanier, R. F. C. Vessot, Cavity Tuning and Pressure Dependence of Frequency in the Hydrogen Maser, Applied Physics Letters, April 1, 1964, pp. 122-123.

TOWARDS AN OPTICAL FREQUENCY STANDARD AT 729 NM

G. Hagel, C. Champenois, D. Guyomarc'h, M. Houssin,
C. Lisowski, M. Vedel, M. Knoop, and F. Vedel

Physique des Interactions Ioniques et Moléculaires (UMR 6633 CNRS-UAM1),
Université de Provence, Centre de St. Jérôme, Case C21, 13397 Marseille Cedex 20, France

ABSTRACT

The development of an optical frequency standard is a possible way to improve the actual stability (in terms of the Allan variance estimator) of today's atomic clocks. One solution is to lock the frequency of a local oscillator (a laser) to the narrow transition of a cooled single ion [1]. The calcium ion is a very promising candidate because of its ultra-narrow dipole-forbidden transition at 729 nm. In this paper, we will summarize the progress we have made regarding the spectral quality of the local oscillator at 729 nm.

INTRODUCTION

The study of trapped ions is of great interest for quantum information [2], spectroscopic issues [3] and development of frequency standard in the optical domain [4]. In our group, we study the opportunity to develop a frequency standard based on a single Ca^+ ion stored in a miniature trap [5]. The proposed clock transition is the electric quadrupole transition $4S_{1/2}$-$3D_{5/2}$ at 729 nm (see figure 1), with a natural linewidth below one hertz. A frequency standard on such a narrow transition requires an ion sufficiently cold to access the Lamb-Dicke regime (which implies the reduction of the discrete Doppler spectrum to few vibrational sidebands) [6].

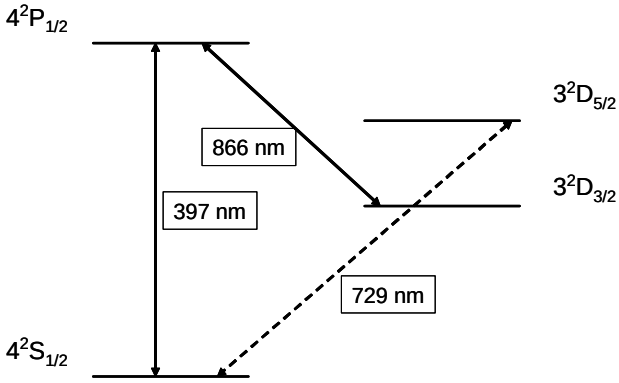


Figure 1:Energy levels of Ca^+

We have recently proposed a novel method to control the localization of a single ion in a trap [7]. We have also evaluated the systematic frequency shifts for such a frequency standard as well as the best interrogation protocol for a minimum clock linewidth [8]. One of the major points to control in view of a frequency standard application concerns the spectral quality of the local oscillator. The progress we have made on this topic will be resumed in this contribution.

ION COOLING

The calcium ion is trapped at the center of a miniature Paul-Straubel trap. The transitions involved in the cooling of the ion are the $4S_{1/2}$-$4P_{1/2}$ transition at 397 nm and the $4P_{1/2}$-$3D_{3/2}$ transition at 866 nm (see figure 1).
In order to reach the Lamb-Dicke regime on the very narrow transition, the localization of the ion in the trap should be inferior to a fraction of the clock wavelength. This is achieved by adding compensating DC voltages to the confining pseudo-potential in order to set the average position of the ion in the zero-field point of the trapping potential. The degree of localization of the ion can essentially be measured from observation of vibrational sidebands in the excitation spectrum of the narrow clock transition. This can be realized by visualizing the quantum jump statistics for different laser frequencies of the clock transition. Figure 2 shows a typical signal of these quantum jumps where the fluorescence signal at 397 nm is represented as a function of time.

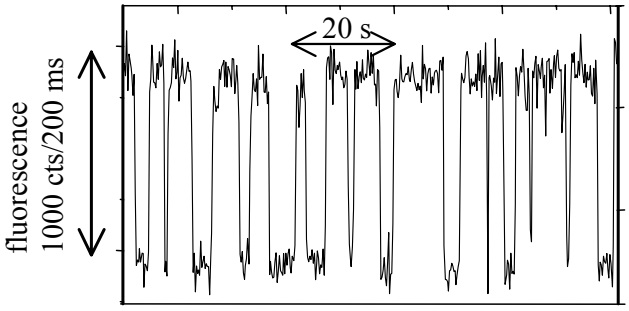


Figure 2 : Quantum jumps of a single Ca^+ ion

STABILIZATION OF THE 729 NM LASER

At present, the local oscillator at 729 nm is a broad area laser diode mounted in an external cavity in Littrow configuration [9]. This multimode diode has a nominal optical output power of 100 mW. Recently, we have improved the stabilization of this laser on a reference cavity. This cavity is a Fabry-Perot long of 10 cm length provided by an ULE (ultra low expansion) spacer. The width in frequency of one interference fringe of this Fabry-Perot is 10 kHz (corresponding to a finesse of 10000). The frequency of the laser is stabilized on this

reference cavity by the Pound-Drever-Hall technique [10].
The output of the external cavity is shaped by three cylindrical lenses (L1,L2,L3) and one spherical one (L4) to reduce the astigmatism of the beam and to produce an almost spherical waist. The beam passes

The reflection signal from the cavity is demodulated at 40 MHz, which produces a dispersion shape which allows a very high frequency lock signal. This signal is separated in three components. The low frequency part (below 300 Hz) drives, via a PZT actuator, the length of the extended cavity of the laser with an important gain

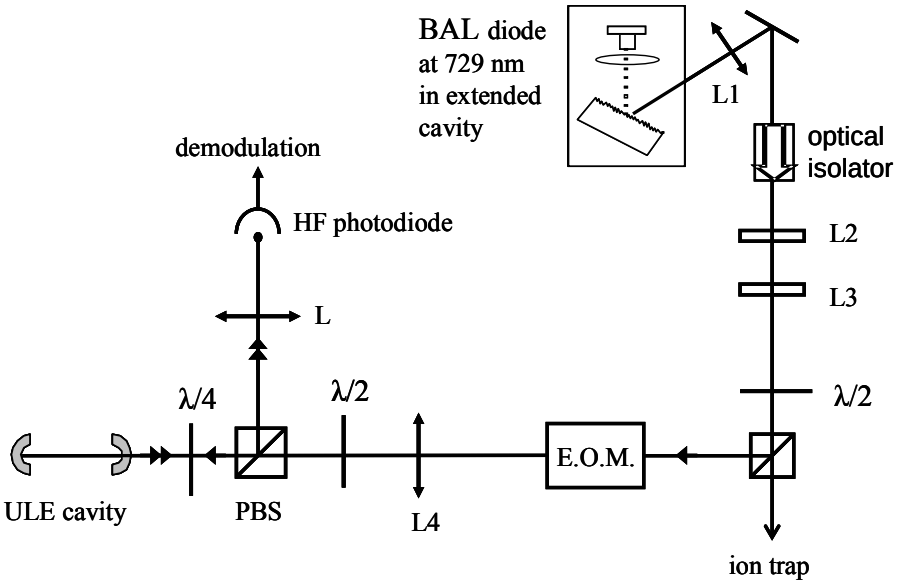


Figure 3 : set-up of the stabilization of the 729 nm laser diode

then through an electro-optic modulator which is resonant at 40 MHz and mounted as a phase modulator (see figure 3). With a beam waist of 200 µm at the entrance of the cavity, we succeed to couple almost 15-20 % of the input beam power into the high-finesse cavity. By the way of a polarization beam splitter and a quarter-wave plate, the reflection beam is detected by a high frequency pass band photodiode (DC to 125 MHz).

and an integrator device. The high frequency part drives with a small gain the input current of the laser (directly on the anode of the laser diode). Finally, the intermediate frequency corrections are applied to the modulation input of the diode's current supply with an adjustable gain. The optical power sent to the cavity is almost 100 µW.
The lock described has allowed us to reduce the jitter of

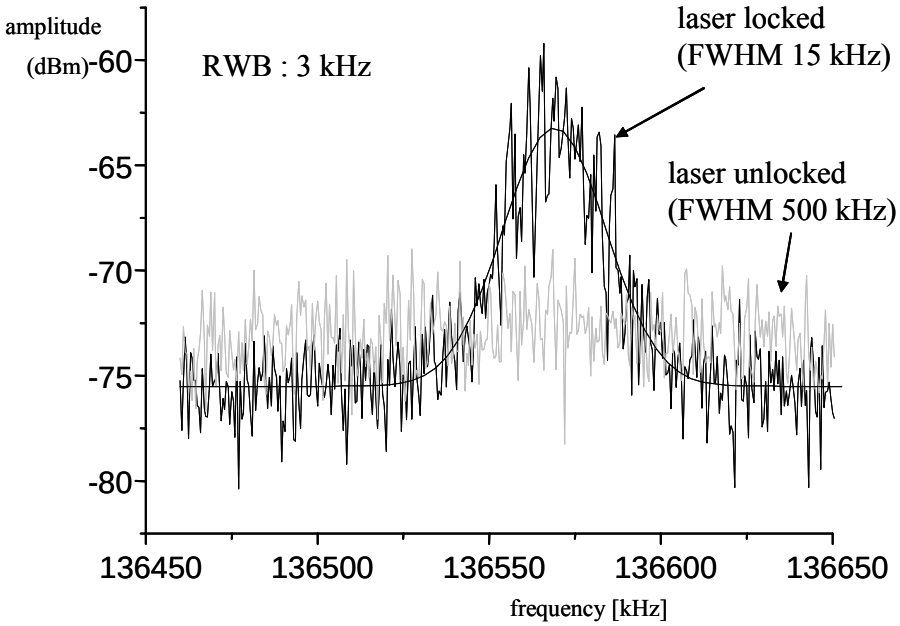


Figure 4 : reduction of the laser jitter

the laser from 500 to 15 kHz. These widths have been measured with an auto-correlation method [11]. We mix two signals coming from the laser on a high-frequency photodiode. The first is sent through an optical fiber of 10 km length and the second passes through an acousto-optic modulator (mounted in double pass for the first order). The resulting frequency beat signal between both signals is monitored on a spectrum analyzer (HP ESA-L1500A). Some precautions are necessary to employ this technique. First of all, the resolution band width (RBW) of the spectrum analyzer should be smaller than the width of the analyzed signal. Secondly, the recording time is the limit in time on which this method can provide information. With these precautions taken, an important reduction of the spectral linewidth, obtained with the described technique, has been observed as depicted in figure 4. At twice the AOM driving frequency, the jitter of the laser can be seen directly, the measured linewidth of the beat signal is 15 kHz. This result is promising, however it does not take into account the absolute stabilization of the cavity, which will be one of the important steps that remain to study.

We have used the described diode laser to excite the proposed clock transition directly, to observe quantum jumps in a single trapped calcium ion and to monitor the quantum jump statistics for the measurement of the natural lifetime of the metastable $3D_{5/2}$-level [3]. It has turned out that the amplified spontaneous emission (ASE) background emitted by the diode is too important to allow the use of this laser as a clock laser. Moreover, the possibility to reduce the linewidth of the diode laser down to the values we aim (a few hertz on one second) is not guaranteed. Indeed, the Schawlow-Townes limit corrected by Henry [12] gives for this diode laser a spectral linewidth of a few kHz. As a consequence, we have started to build a home-made titanium-sapphire laser which has much tighter intrinsic spectral specifications [13]. In this solid-state laser, the high-frequency corrections are made on an intra-cavity EOM, allowing to achieve a very high stability.

As a next step, the reference cavity will be mounted in a vacuum vessel, to isolate it from acoustic and environmental noise. The cavity will be mounted in a way to minimize mechanical vibrations and thermal fluctuations. At least two stages of active thermal stabilization will be added. We will eventually opt for a new reference cavity, longer than the actual one and with a higher finesse (length 20-25 cm, finesse > 100 000). With the TiSa laser and the stabilization of this new reference cavity, we hope to obtain in a near future a local oscillator with a jitter of a few hertz at one second and with stability in the same order of magnitude. In parallel, we will continue to explore the opportunity of an "all-diode" set-up, which is one of the main advantages of the calcium ion. Its realization depends mainly on the commercial existence of single-mode laser diodes at 729 nm.

[1] P. Gill, *Proceedings of the Sixth Symposium on Frequency Standards and Metrology* (World Scientific Ltd, Singapore, 2002)

[2] M. Riebe, H. Häffner, C. F. Roos, W. Hänsel, J. Benhelm, G. P. T. Lancaster, T. W. Körber, C. Becher, F. Schmidt-Kaler, D. F. V. James, and R. Blatt, *Deterministic quantum teleportation with atoms,* Nature **429**, 734 (2004)

[3] M. Knoop, C. Champenois, G. Hagel, M. Houssin, C. Lisowski, M. Vedel, and F. Vedel, *Metastable level lifetimes from electron-shelving measurements with ion clouds and single ions*, Eur. Phys. J. D **29**, 163 (2004)

[4] S.A. Diddams, Th. Udem, J.C. Bergquist, E.A. Curtis, R.E. Drullinger, L. Hollberg, W.M. Itano, W.D. Lee, C.W. Oates, K.R. Vogel, and D.J. Wineland, *An Optical Clock Based on a Single Trapped Hg^+ Ion* ,Science **293**, 825 (2001)

[5] C. Champenois, M. Knoop, M. Herbane, M. Houssin, T. Kaing, M. Vedel, and F. Vedel, *Characterization of a miniature Paul-Straubel trap*, Eur. Phys. J. D. **15**, 105 (2001)

[6] R. H. Dicke, *The effect of collisions upon the Doppler width of spectral lines*, Phys. Rev. **89**, 472 (1953)

[7] C. Lisowski, M. Knoop, C. Champenois, G. Hagel, M. Vedel, and F. Vedel, *Dark resonances as a probe for the motional state of a single ion*, arxiv:physics/0407135

[8] C. Champenois, M. Houssin, C. Lisowski, G. Hagel, M. Knoop, M. Vedel, and F. Vedel, *Evaluation of the ultimate performances of a Ca^+ single-ion frequency standard,* Phys.Lett. A. **331**, 298 (2004)

[9] M. Houssin, P. Courteille, C. Champenois, M. Herbane, M. Knoop, M. Vedel, and F. Vedel, *Linewidth reduction by 6 orders of magnitude of a broad-area 729-nm diode laser,* Applied Optics **42**, 4871 (2003)

[10] R. W. P. Drever, J. L. Hall, F. V. Kowalski, J. Hough, G. M. Ford, A. J. Munley, and H. Ward, *Laser phase and frequency stabilization using an optical resonator* , Appl. Phys. Lett. **31**, 97 (1983)

[11] T. Okoshi, K. Kikuchi, A. Nakayama, *Novel method for high resolution measurement of laser output spectrum*, Electron. Lett. **16**, 630 (1980)

[12] C. H. Henry, *Theory of the phase noise and power spectrum of a single mode injection laser*, IEEE J. Quant. Elec. **19**, 1391 (1983)

[13] B. de Beauvoir, C. Schwob, O. Acef, L. Jozefowski, L. Hilico, F. Nez, L. Julien, A. Clairon, and F. Biraben, *Metrology of the hydrogen atoms : Determination of the Rydberg constant and Lamb shifts*, Eur. Phys. J. D **12**, 61 (2000)

The vortices frequency measurement in piezoceramic vortex flow meter

Krzysztof Weiss, Sylwester Gawor, Aleksander Lisowiec, Mariusz Tupaj

Tele and Radio Research Institute

kweiss@itr.org.pl

Abstract

The strut positioned in the stream of flowing medium induces in this medium vortices with frequency depending on medium velocity and strut shape and diameter. This relation is described by Strouhal Number. The vortices frequency measurement gives information about the medium flow velocity. In most commercial vortex flow meters the vortices existence is detected by pressure measurement or by mechanical vibration of bluff body measurement.
It is also possible to measure them by ultrasound waves. The ultrasonic beam located downstream from the small strut, is transmitted through the vortex pattern.. The measuring system consists of transmitting and receiving transducers positioned on opposite sides of pipe in which the medium flow is measured. The vortices change the structure of received signal. This signal is detected and the vortices frequency is measured . In the paper the principle of vortex flow meter operation and detector circuits are described, and results of practical measurements are presented.

1. Principle of vortex flowmeter acting

The Strouhal Number for round strud is the quotient of vortex number *n in[Hz]* times strud diameter *D in[m]* and flow velocity *V in [m/s]*.

$$S = \frac{n \cdot D}{V}$$

This number for Reynold's Number in range 400< Re <5000 is a constant value 0,21. Below Re = 400 its value falls and is equal to 0,12 for Re=60.
Transforming this equation one can obtain relation for flow speed calculation:

$$V = \frac{S \cdot n}{D}$$

As *D* and *S* are known, measurement of the vortices frequency gives the flow velocity. This frequency can be measured by ultrasonic beam passing through the flowing medium perpendicular to the flow direction and direction of the strud. The vortices induced by the strud, modulate the frequency of passing signal because of Doppler effect. Frequency change depends on the ratio of disturbing bar diameter to distance between transducers and on sound velocity in a medium between transducers.

$$\Delta f = \frac{f_0 \cdot V_m \cdot d}{2h \cdot V_d}$$

where: f_0 - *transducer's work*
V_m- *medium's speed*
V_d - *velocity of sound in a medium*
d – disturbing bar diameter
h – distance between transducers

The principle of vortex flowmeter is presented in fig 1.

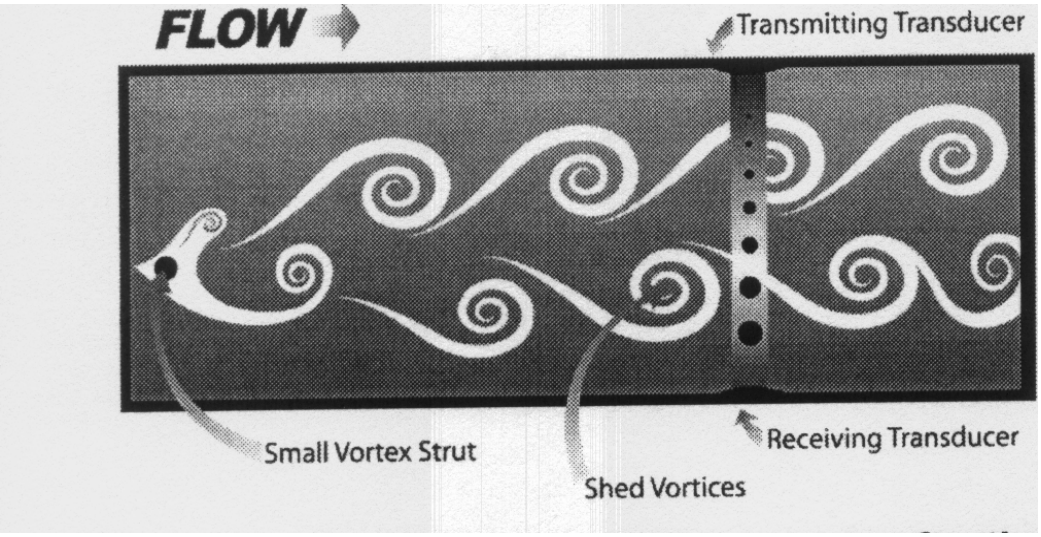


Fig 1. Vortex flowmeter principle of acting.

2. The instrument construction

The flowmeter probe consists of metal cylinder with square window perpendicular to the cylinder axis. In this window, on opposite sides, perpendicular to the cylinder axis are mounted two piezoceramic transducers. One of them is working as ultrasound transmitter and the second as ultrasound receiver. The disturbing strud is mounted parallel to transducers surfaces on the edge of window. The frequency modulation is detected in electronic circuit, where the frequency modulation is converted to electric signal corresponding to the frequency of vortices. The flowing medium speed is calculated from equation:

$$V = k \cdot f$$

Where: k – coefficient typical for measuring head
f - vortices frequency

The flowmeter electronics is physically placed on three printed circuit boards. One of the boards contains the oscillator that produces the input signal to transmitting transducer, the other board contains the signal processing circuitry that amplifies, demodulates and further transforms the output signal of receiving transducer The third board contains clamps for connecting power and input and output signals. The third board will in future contain the pressure and temperature measurement electronics.

The block diagram of the transducer output signal processing circuit is presented in fig. 2. The main blocks are input amplifier/limiter with quadrature detector (built around SA

604A integrated circuit), post detector amplifier (1/2 LM 358), Schmidt comparator (1/2 LM 358) and frequency/current conversion circuit.

The input amplifier is broadband matched to 50 Ω transmission line. The transducer output signal level is about 0.1 mV RMS (depending on the actual transducer used). SA 604A has enough amplification to convert it to square wave at its demodulator output. In the quadrature demodulator a 1 MHz quartz resonator has been used as a parallel resonant circuit. It offers very good sensitivity and works very well up to 5 kHz modulation frequency and 1 kHz deviation as detector linearity is not a specially important issue. At 100 Hz modulation frequency and 100 Hz deviation, the signal at the detector output is 300 mV RMS.

The signal at SA 604A demodulator output is further amplified by a factor of three by operational amplifier (first half of LM 358). The 3 dB cutoff frequency of the amplifier circuit is about 3 Hz. Another operational amplifier (second half of LM 358) forms a comparator with hysteresis (the hysteresis is set at about 400 mV to prevent the transitions at comparator output without any flow) that transforms the amplified signal to pulse train suitable for external frequency meter and frequency/current converter.

There are 7 V and 12 V voltage regulators on the signal processing board. The 12 V fixed regulator supplies power to most of the circuits. The 7 V regulator (MIC5205) supplies power to SA 604A and LM 358.

One of the problems is direct transmission (through electromagnetic radiation) of oscillator output signal to the input of SA 604A IF amplifier. If this direct signal is comparable or greater than transducer output signal (100 μV RMS) it diminishes the effective deviation of signal at demodulator input causing lower level of demodulated signal.

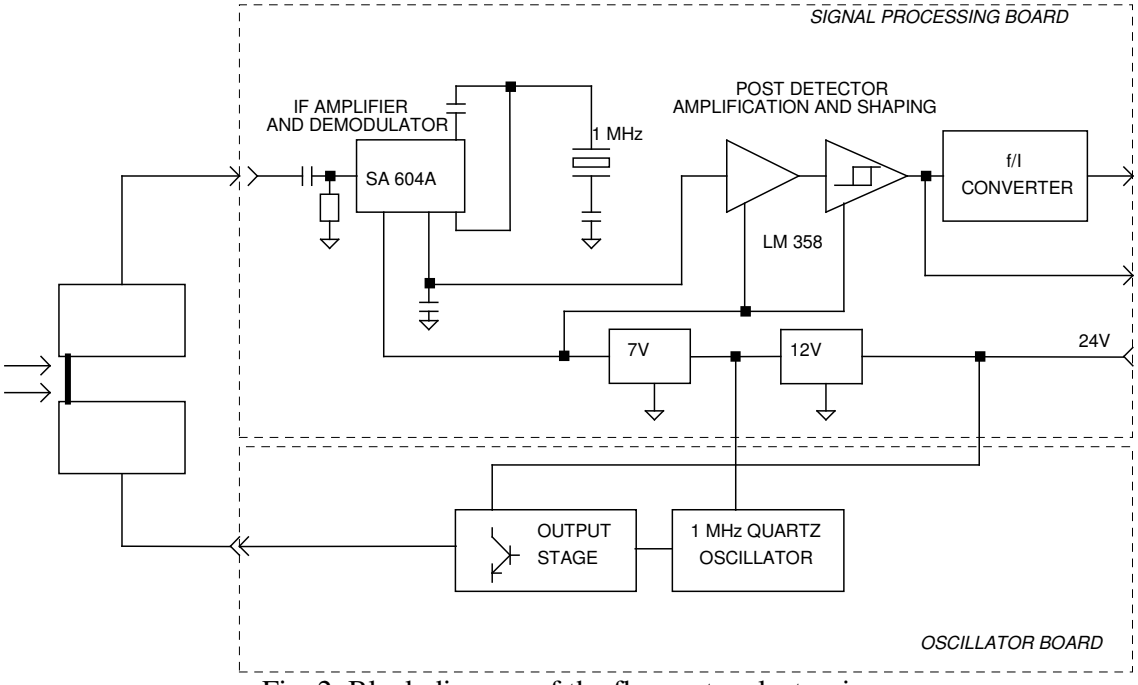


Fig. 2. Block diagram of the flowmeter electronics

3.Experimental results

The aerodynamic tunnel with diameter 75 mm was equipped with vane anemometer with flow velocity measurement range 0,3÷30 m/s and resolution 0,1 m/s. In this tunnel

different kinds of measuring heads and electronic circuits were investigated in the flow range 0,5÷15 m/s. The typical characteristics are presented in fig 3.

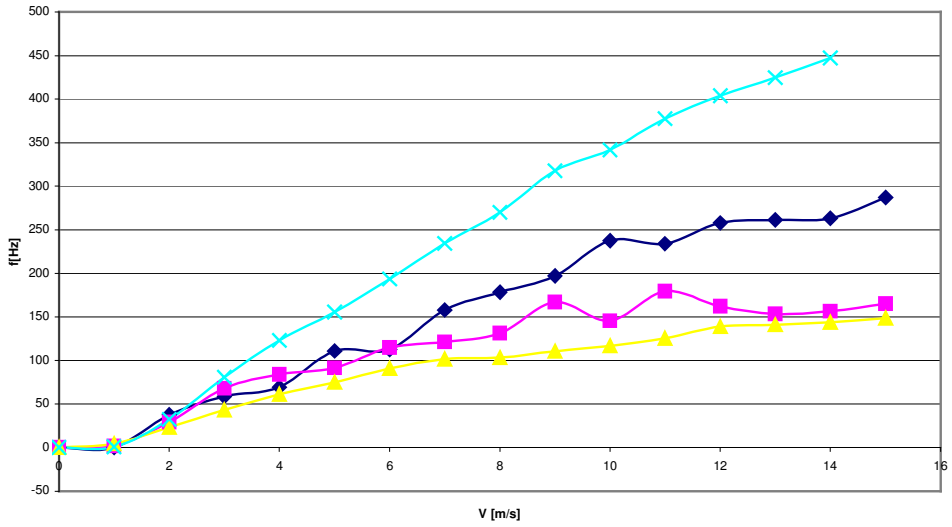


Fig.3. Typical characteristics of vortex flowmeter.

We observed the nonlinearity of frequency versus flow velocity characteristics. These characteristics were dependent of head shape and construction and on parameters of detector circuit.

The quadrature detector output signal has been sampled by digital oscilloscope and converted by Fourier transform into frequency domain. The spectra for 1.5 m/s , 3 m/s, 9 m/s and 15 m/s flows are presented below.

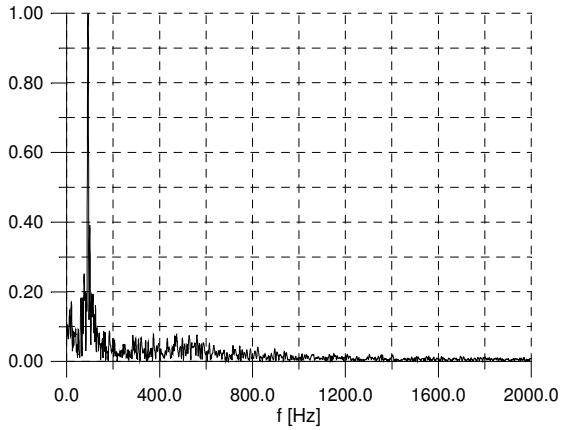


Fig.4 Fourier transformation for 1,5 m/s

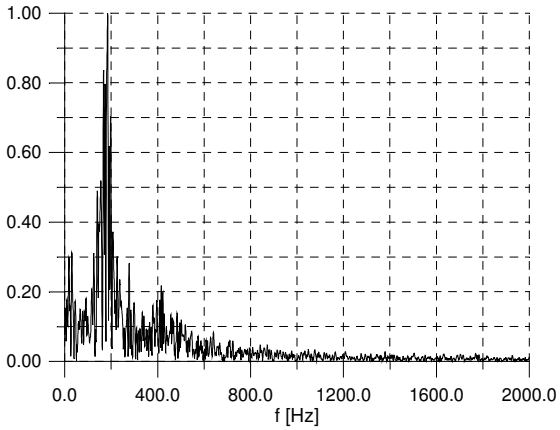


Fig.5 Fourier transformation for 3 m/s

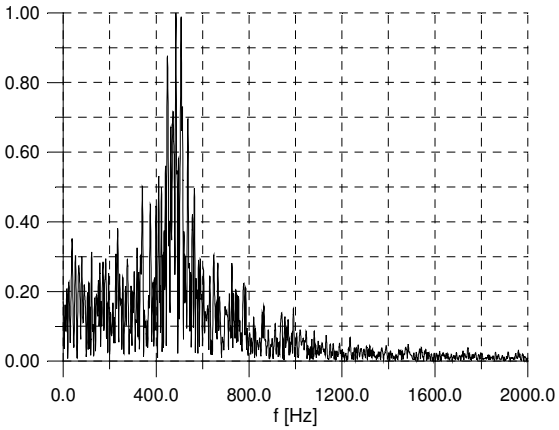


Fig.6 Fourier transformation for 9 m/s

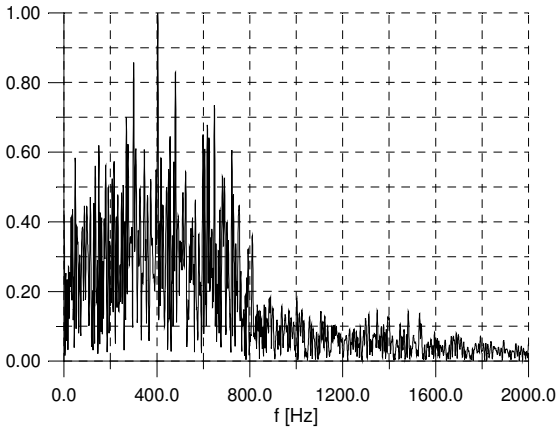


Fig.7 Fourier transformation for 15 m/s

It can be observed that for small flow velocity the spectral line corresponding to vortices frequency is quite strong. As the flow speed increases, the spectrum becomes more complex and many other spectral lines appear. Some of them move with flow speed increase and some of them remain stationary. Around the speed at which the flow speed – frequency characteristic bends, the intensity of parasitic lines becomes higher than the intensity of main line.

Further measurements have been carried in aerodynamic tunnel with flow speeds up to 60 m/s. Optimized shape of inlet enables linear flow speed measurement up to 30 m/s and monotonic characteristic up to 40 m/s.

4.Conclusions

The vortex flowmeter is predisposed to small and medium flow speeds measurement. Maximum measurable flow speed with head shape and construction used in our instrument is 30 m/s. The limitation on maximum measurable flow speed comes from vortices caused by head body parts. To increase the measurement range it is necessary to modify the channel shape so as to eliminate unwanted vortices. It is also possible to read the flow speed from Fourier transform of the signal at the detector output, however this necessitates the use of complex microprocessor circuit and still the level of unwanted vortices should be as low as possible.

References

H. Schichling. Grenzschicht – Teorie. Izdatielstwo „Nauka” Moskwa 1969. p.p. 19-56
http://www.efunda.com/designstandards/sensors/flowmeters/flowmeter_vtx.cfm

Shanghai Astronomical Observatory's Hydrogen Masers and Their Applications

Z.C.Zhai,C.F.Lin,W.Q.Zhang,J.X.Peng,Y.P.Zhang,R.H.Zhang,Y.Cai and P.H.Yang

Shanghai Astronomical Observatory, Chinese Academy of Sciences

80 Nandan Road, Shanghai-200030

Abstract

During more then 30 years, Shanghai Astronomical Observatory has developed three generations of hydrogen masers, 6 laboratory models and more than 50 engineering units of active hydrogen masers as well as 5 compact passive hydrogen maser. They have been used in several fields, including one is in Australia. This paper describes their design features, applied performance specifications and their applications.

The Development of Applied Hydrogen Maser

A new generation of hydrogen maser has been developed[1] since 1985, and so far 46 this kind of hydrogen masers have been built and put to use in China. The new masers are compact, relatively light-weight, rugged, and easily transportable standards, Entire standard was to be self-contained in a single standard rack cabinet. All electronics including the phase-lock receiver were to be mechanically and electrically integrated into the overall design. It is a relatively small instrument weighing 150Kg. The left of Fig1 shows the photograph of the complete maser system with 108cm high, 52cm wide, 68cm deep. All parameters of the maser can be monitor-controled either by a microcomputer built in inside front panel or by microcomputer in office through an IEEE 232 interface.

A comprehensive series of environmental and stability tests on the new masers has been evaluated[2]. Table 1 shows their performance specifications.

A cavity auto-tuning system (CAT) is installed in the masers. The masers with auto-tuning system can be worked alone either in CAT ON manner or in CAT OFF manner. It shows that there is only minimal degration of the maser short-term stability due to the introduction of the injected signal of the cavity auto-tuner. However, the long-term stability of the maser has greatly been improved.

The Development of Compact Hydrogen Maser[3]

At present, Shanghai Observatory meanwhile is developing the third generation of hydrogen maser, a passive hydrogen maser which employs a compact cavity design, resulting in significant size and weight reductions compared to a conventional maser. The right of Fig1 shows the photo of the miniature passive hydrogen maser which is 70cm high, 40cm wide, 52cm deep and weighs 40kg.

1. Quantum hydrogen discriminator

A "magnetronic" cavity was used with more mechanical strong to be built. The dimension of the cavity, which is made of aluminum ,is inner diameter about 15.2cm , height 14cm. The Q-factor of the cavity is about 5500.

4-pole state selector is used in the compact passive hydrogen maser. The bore diameter of the magnet is 1mm and its length is 59mm.

Other designs in the physical part of the passive hydrogen maser are similar to transportable active hydrogen masers(SOHM-3) produced by Shanghai Observatory.

2. Electronic system

Electronics of the passive hydrogen maser comprises two frequency lock-loops. 10MHz crystal oscillator frequency is locked to the hydrogen emission line and resonant cavity frequency is locked to the crystal oscillator frequency.

3. Data measurements

With the SOHM-3 active hydrogen maser as a reference, the initial test of the passive maser indicates that the frequency stabilities exhibit approximately $\sigma(2, \tau)=3\times10^{-12}\tau^{-1/2}$($1s<\tau<1\times10^{4}s$). At present, there is still some work to be done to analyze the various factors influencing the passive hydrogen maser so as to improve its performance.

The Application Aspects

To equip Chinese VLBI network and satisfy some scientific/engineering subjects in China, Shanghai Observatory has commercially produced active hydrogen masers, SOHM-3, since 1989, and so far more than 50 SOHM-3 hydrogen masers have been built. The following is a brief introduction to their applications.

1. Applications for VLBI

Chinese VLBI network consists of two stations, Sheshan, Shanghai and Nanshan, wurumoqi, with a 25-meter antenna each station, and one mobile station with 6-meter antenna, located FengHuang shan, kunming.

Shanghai Observatory's 6 hydrogen masers have been used in the VLBI network, with two H-masers each station.

Also, for VLBI work, one of Shanghai Observatory's H-maser went to Australia to join ATNF VLBI Observations in 1993.

2. Applicalions for positioning system

Atomic clocks are the essentials for navigation-positioning system. Some units of Shanghai Observatory's H-masers have been used in Chinese radio and satellite navigation-positioning system.

3. Applications for communication and tracking system

Radio communication and tracking station need high precision time frequency standards. Several units of Shanghai Observatory's H-masers have been used in these systems, including one of which being used in mobile tracking station.

4. Applications for T/F metrology

Atomic hydrogen maser can be used as a metrological reference standard to evaluate secondary standards such as Rb, crystal oscillator, and even Cs standards etc. 5 of Shanghai Observatory's H-masers have been used for these purposes.

5. Applications for time-keeping

Hydrogen masers, of course, can be used for time-keeping, especially hydrogen masers with cavity auto-tuning. 6 of Shanghai Observatory's hydrogen masers have been in continuous operation to be as reference standards for time-keeping.

References

[1] Z.C.Zhai, H.X.Huang, G.X.Jiang et al,"The new generation of hydrogen maser at Shanghai Observatory", proc.41st Annu.symp.Frequency control (1987) 82

[2] Z.C.Zhai,C.F.Lin,H.X.Huang et al, " Performance evaluation of the Shanghai Observatory's new H-maser", Electromagnetic metrology,(International Academia publishers,Beijing 1989)

[3] C.F.Lin,T.X.Liu,Z.C.Zhai et al, "Miniature passive hydrogen maser at Shanghai Observatory", proc. 2001 IEEE international Frequency control symposium and PDA Exhibition (2001) 89.



Figure 1. A photograph of the transportable active hydrogen maser and the passive hydrogen maser constructed at Shanghai Observatory

Table 1. Performance data for Shanghai Observatory's active hydrogen maser

Items	Performance							
Frequency Stability	τ CAT ON	10ms 3E-11	100ms 4E-12	1s 4.0E-13	10s 5E-14	100s 8E-15	1000s 6E-15	10^5s 5E-15
Accuracy	5×10^{-13}							
Reproducibility	$\pm3\times10^{-14}$ (after 24 hours switching off)				$\pm2\times10^{-13}$ (after transportation)			
Temperature sensitivity(1)	1.4×10^{-14} □ (20 □-35□)				1.8×10^{-14} □ (35□– 20□)			
Temperature sensitivity(2)	$<5.3\times10^{-14}$ (From 20□ to 35□ and to 20□ in the 24 hours)							
Magnetic Sensitivity	1.4×10^{-14}/G							
VoltageVariance (220V± 20%)	< 4×10^{-14}							
Phaser noise (5MHz)	1Hz	10Hz	100Hz	1000Hz	10000Hz			
	-100dBc	-120dBc	-140dBc	-150dBc	-150dBc			
Frequency resolution	± 7×10^{-14}							

IMPROVED THEORY FOR HIGHER-ORDER BLACKBODY RADIATION SHIFT IN Cs CLOCKS

V. G. PAL'CHIKOV, YU. S. DOMNIN

Institute of Metrology for Time and Space at National Research Institute for Physical-Technical and Radiotechnical Measurements, Mendeleevo, Moscow Region 141570, Russia ; E-mail: vitpal@mail.ru , Fax: +007-095-534-0609

ABSTRACT

A theoretical method is presented for calculating the higher-order terms of series in powers of the black-body radiation (BBR) field for the Stark-state wavefunctions, dipole transition matrix elements and corresponding frequency shifts of hyperfine splitting in the ground states for Cs atoms. A theoretical approach for calculating the BBR shifts in Cs atoms is based on the Green function method for summation over all intermediate states and exact Dirac-Fock wavefunctions for the resonant transitions to the first excited s-, p- and d-states. By comparing the calculated BBR shift with results of measurements employing atomic frequency standards, the BBR effects on the ground state of Cs atoms are analyzed.

Keywords: Black-body radiation, hyperfine structure, ac Stark effect.

1. INTRODUCTION

Although the BBR effects are most dramatic in Rydberg atoms it is important to note that the ac Stark shifts are quite substantial even in ground state systems, in particular, on the reference ^{133}Cs hyperfine transition $6^2S_{1/2}$ [(f=3,M_f=0)- (f=4,M_f=0)] in atomic cesium fountain [1]. Therefore, the BBR shift is large enough to be observable in the modern measurements, which could limit the accuracy of this kind of frequency standard.

In our paper, a general method is presented for calculating the higher-order terms of series in powers of the BBR field for the Stark-state wavefunctions and corresponding frequency shifts of hyperfine splitting in the ground states for Cs atoms. The ac polarizabilities and hyperpolarizabilities are calculated, starting from a formal expansion of the second- and fourth-order energies in terms of the complete set of field-free wavefunctions. We used a multiconfiguration Dirac-Fock program for determination of the radial matrix elements. Alternatively, we have been developing a semi-empirical model potential method which, by simplifying the description of the valence electron – atomic core interaction, makes it possible to study complex atoms by conventional few-electron configuration-interaction methods. Atomic units are used throughout

2. GENERAL FORMULATIONS

The spectral properties of an atom in a field F of a light wave are determined by the quasienergies, corresponding to the nonperturbed atomic spectrum. Except for some special cases the perturbation of spectra in the first nonvanishing order turns out to be quadratic with respect to the electric field amplitude F of the wave (dynamic or ac Stark effect). In the general case, the energy shift $\Delta E = E - E_0$ of an atomic level under the action of monochromatic radiation with electric-field strength

$$\vec{F}(t) = F\,\mathrm{Re}\left\{\vec{e}\,\exp\left[(\vec{k}\cdot\vec{r} - \omega t)\right]\right\} \quad (1)$$

(where $\vec{e}$ is the unit polarization vector and $\vec{k}$ is the wavevector) is the root of the secular equation

$$\det\left|\left|\Delta E\,\delta_{MM'} - U_{MM'}\right|\right| = 0, \quad (2)$$

where $U_{MM'}$ is the matrix element of the atom-field interaction operator.

The ac polarizabilities and ac hyperpolarizabilities are calculated starting from formal expressions for the second- and fourth-order quasienergy in terms of the field-free wave functions $|0\rangle$ and the reduced quasi-energy Green functions G [2]:

$$\Delta E^{(2)} = U^{(2)}_{MM} = -\frac{F^2}{4}\,\alpha(\vec{e},\omega) = -\langle\langle 0|\,HGH|0\rangle\rangle, \quad (3)$$

$$\Delta E^{(4)} = U^{(4)}_{MM} = -\frac{F^4}{64}\,\gamma(\vec{e},\omega) = -\langle\langle 0|\,HGHGHGH|0\rangle\rangle + \langle\langle 0|\,HGH|0\rangle\rangle\langle\langle 0|\,HG^2H|0\rangle\rangle \quad (4)$$

$$G_E = \sum_m \frac{|m\rangle\langle\langle m|}{\varepsilon_m^{(0)} - E}, \quad (5)$$

where the double brackets indicate the integration over the spatial variables and averaging over time, α and

γ are polarizability and hyperpolarizability, $\varepsilon_m^{(0)} = E_m^{(0)} + \omega$ is the quasi-energy, H is the interaction Hamiltonian.

$$H(\vec{r}, t) = V(\vec{r})e^{-i\omega t} + V^{+}(\vec{r})e^{i\omega t}, \quad (6)$$

In the dipole approximation

$$V(\vec{r}) = -\frac{1}{2}F(\vec{e}\cdot\vec{d}), \quad (7)$$

where $\vec{d}$ is the atomic dipole moment.

After averaging over time, the frequency dependent electric field of blackbody radiation has the form:

$$\frac{1}{8\pi}F^2(\omega) = \frac{1}{3}u(\omega, T), \quad (8)$$

where $u(\omega, T)$ is the spectral energy density

$$u(\omega, T) = \frac{\alpha^3}{\pi^2}\frac{\omega^3}{e^{\omega/kT}-1}. \quad (9)$$

Therefore, in the lowest order the total shift of the state |a> due to BBR effect can be calculated by using second order perturbation theory [3]:

$$\Delta E_a = \frac{1}{4}\sum_b |< a|r|b >|^2 \times \int_0^\infty F^2(\omega)\left(\frac{1}{\omega_a - \omega_b - \omega} + \frac{1}{\omega_a - \omega_b + \omega}\right)d\omega$$
$$= \frac{2\alpha^3}{3\pi}\sum_b |< a|r|b >|^2 \int_0^\infty \omega^3 \left(e^{\omega/kT} - 1\right)^{-1} \times \left(\frac{1}{\omega_a - \omega_b - \omega} + \frac{1}{\omega_a - \omega_b + \omega}\right)d\omega \quad (10)$$

By consideration of (6) as a quasistatic perturbation [3], it is necessary to take into account the average (rms) value of the BBR electric field

$$< F^2(\omega) >= \frac{1}{2}\int_0^\infty F^2(\omega)d\omega = \frac{4\pi^3}{15}\alpha^3(kT)^4$$
$$= (8.319V/cm)^2\left[T(K)/300\right]^4 \quad (11)$$

The ac Stark effect in the hyperfine-structure (hfs) occurs as a result of disturbance of atomic levels under the influence of the interaction between electron and nuclear magnetic dipole moment (in the first order) as well as the interaction with BBR electric field (in the second order). If the field are weak enough the center of multiplet is shifted and there occurs the splitting of sublevels of hfs-multiplet to second order in the field (6).

The general expression for the second order corrections in(6) for the energy level f, M_f is [4]:

$$\Delta E(f, M_f) =< f, M_f| HGH_{hfs}GH| f, M_f > + 2 < f, M_f| H_{hfs}GHGH| f, M_f > - < f, M_f| H_{hfs}| f, M_f > \times < f, M_f| HGGH| f, M_f > \quad (12)$$

Here f is the total atomic moment, M_f is its projection, H_{hfs} is the non-relativistic Hamiltonian of hfs interactions including the contact, magnetic dipolar and tensor terms, respectively. In particular case of the ground S-state , the energy shift may be represented in the form [4]

$$\Delta E(f = I + 1/2, M_f) = -\frac{\alpha^2\mu_I}{3}F^2\{\alpha_0 + \alpha_1(1 - Q/\alpha^2\mu_I) \times \frac{(I + 3/2)(I + 1/2)}{I(2I + 1)}\} \quad (13)$$

$$\Delta E(f = I + 1/2, M_f) = -\frac{\alpha^2\mu_I}{3}F^2\{\alpha_0 + \alpha_1(1 + Q(2I + 3)/\alpha^2\mu_I(2I - 1)) \times \frac{(I^2 - 1/4) - 3M_f^2}{I(2I + 1)}\} \quad (14)$$

In the formulas (13), (14), I is the nuclear spin, μ_I is the nuclear magnetic moment, Q is the quadruple moment, α_0 and α_1 are scalar and tensor parts of polarizability (radial parts) :

$$\alpha_0 = \frac{A_{hfs}}{3\alpha^2\mu_I} < a| rg_1g_1r'|a > - \frac{2}{9} R_a(0)\sum_k \frac{R_k(0)}{E_k - E_a} < k| rg_1r'|a >,$$

$$\alpha_1 = \frac{2}{15}[< a| rg_1r'^{-3} g_1r''|a > + 2 < a| r^{-3}g_1r' g_1r''|a >], \quad (15)$$

where $R(0)$ is the radial wavefunction at $r = 0$, A_{hfs} is the hyperfine structure constant (without electric field), g_1 is the radial Green function (see the next section).

The general expressions for fourth-order energy shift are much more complicated. Taking into account the hfs-interaction as the first-order perturbation potential, we can express the fourth-order Stark energy shift by means of the one-particle Green's function for valence electron [5]:

$$\Delta E^{(4)} = -\frac{F^4}{4!}\{2[< a| h(GD)^4|a > - < a| DGD|a >< a| hGDG^2D|a > - < a| DGD|a >< a| hG^2DGD|a >] + 2[< a| DGh(GD)^3|a > - < a| DGD|a > \times < a| DGhG^2D|a >] + < a| (DG)^2h(GD)^2|a > - 2 < a| DG^2D|a >< a| DGDGh|a > - < a| DG^2D|a >< a| DGhGD|a > + < a| h|a > [< a| DG^2D|a >^2 - 2 < a| DG(DG)^3|a > + 2 < a| DGD|a > \times < a| DG^3D|a > - < a| (DG)^2(DG)^2|a >]\}. \quad (16)$$

Here h represents the full set of hfs interactions, D is the z-component of the electric dipole moment of the atom. The final expression for $\Delta E^{(4)}$ may be written in the form in analogy to the second-order contributions (13)-(14) and has been omitted here for brevity.

3. ANALYTICAL EXPRESSION FOR GREEN FUNCTION

The problem of exact calculations of the quantities (13)-(16) is very difficult and involves calculating spectral sums over the complete set of unperturbed states, including direct summation over the discrete spectrum of intermediate states and integration over the continuous spectrum of intermediate states, which represent the Green function (5).

In the present work the second- and the fourth-order BBR contributions are calculated using Sturmian Green's functiom method [2]. It provides a reliable control for the accuracy of the calculation with the finite sums of discrete states. The angular part of the Green's function $G_E(\hat{r}_1, \hat{r}_2)$ is simply the product of spherical harmonics, while for the radial part $g_l(E; r_1, r_2)$ we have taken an expansion in Sturmian functions :

$$G_E(\hat{r}_1, \hat{r}_2) = \sum_{lm} g_l(E; r_1, r_2) Y_{lm}(\hat{r}_1) \times Y^*_{lm}(\hat{r}_1),$$

$$g_l(E; r_1, r_2) = \frac{4Z}{\nu} \times \sum_{k=0}^{\infty} \frac{U_{kl}\left(\frac{2Z\, r_1}{\nu}\right) U_{kl}\left(\frac{2Z\, r_2}{\nu}\right)}{k + \lambda_l + 1 - \nu}, \quad (17)$$

where $\nu = \frac{Z}{\sqrt{-2E}}$ and

$$U_{kl}(x) = \sqrt{\frac{k!}{\Gamma(k + 2 + 2\lambda_l)}}\, x^{\lambda_l} \times \exp\left(-\frac{x}{2}\right) L_k^{2\lambda+1}(x). \quad (18)$$

Here $L_m^l(x)$ is the Laguerre polynomial and $\Gamma(x)$ is the gamma function. The radial functions $R_{nl}(r)$ are obtained from the residues of the Green's function at the poles of $g_l(E; r_1, r_2)$:

$$\nu = \nu_{nl} = n_r + \lambda_l + 1, \quad (19)$$

with $n_r = 0,1,2, \ldots$ the radial quantum number, λ_l the effective orbital momentum and ν_{nl} the effective principal quantum number. These effective quantum numbers for the ground and metastable levels, are discussed in [6].

4. NUMERICAL RESULTS AND DISCUSSIONS

The quantitative relation between shift and splitting of atomic levels with respect to the BBR is thus determined by relations between the set of the second-order contributions (polarizabilities) and the fourth-order contributions (hyperpolarizabilities), presented above.

In the present paper, we estimate the temperature-dependent shift induced by the BBR electric field on the $\nu=6S_{1/2}(f=3,M_f=0)$- $(f=4,M_f=0)$ clock transition in Cs atoms, including calculations for the higher-order field contributions (hyperpolarizabilities):

$$\frac{\delta\upsilon}{\upsilon_0} = -16.8 \times 10^{-15}\left(\frac{T}{300K}\right)^4 \times \left[1 + 0.014\left(\frac{T}{300K}\right)^2 - 7.56 \times 10^{-11}\left(\frac{T}{300K}\right)^4\right] \quad (20)$$

The first term in Eq. (20) corresponds to the lowest order of BBR shift (polarizability).The second term is a correction due to separation in the frequency for D1 and D2 lines [7]. The last one connected with higher-order Stark effect (hyperpolarizability). Note, that re-calculated value for the hyperpolarizability in (20) is much smaller, compared with previous calculated value, presented in [8]. These results (20) have been obtained by using a multiconfiguration Dirac-Fock (GRASP) program code for matrix elements of polarizability and the model potential calculations for the higher-order terms. Alternatively, by applying model potential approach, the main numerical term in (20) is

$$\frac{\delta\upsilon}{\upsilon_0} = -17.4 \times 10^{-15}\left(\frac{T}{300K}\right)^4$$

These re-calculated values for the BBR shift in the lowest order agrees both with a measurement [1] and previous theory [7,8] and differs from the recent data, reported in [9].

5. ACKNOWLEDGEMENTS

The authors thank Andre Clairon (Paris Observatoty, France), Andreas Bauch and Stephen Weyers (Physikalisch-Technische Bundesanstalt, Germany) for discussions, consultations, and interest in this study. This work was supported in part by the Ministry of Science and Technogies of the Russian Federation [Contract No. GNTD/GK-021(00)] and the extra-budget Foundation of the State Standard of the Russian Federation.

REFERENCES

[1] E. Simon, P. Laurent, and A. Clairon, "Measurement of the Stark shift of the Cs hyperfine splitting in an atomic fountain", *Phys. Rev.,* Vol. A57, pp.436-439, 1998.

[2] N. L. Manakov, V.D. Ovsiannikov, and L.P. Rapoport, "Atoms in a laser field.", *Phys. Reports, Vol* 141, pp. 319-433, 1986.

[3] J. W. Farley and W. H. Wing , "Accurate calculation of dynamic Stark shifts and depopulation rates of Rydberg energy levels induced by blackbody radiation. Hydrogen, helium, and alkali-metal atoms.", *Phys. Rev.,* Vol. A23, pp.2397-2424, 1981.

[4] N. L. Manakov, V.D. Ovsiannikov, and L.P. Rapoport, "Stark effect on hyperfinr structure of atoms.", *Opt. Spektrosk., Vol* 38, pp. 424-427, 1975 [Opt.Spectrocs. (USSR), *Vol* 38, pp. 242-245, 1975].

[5] V. G. Pal'chikov, " Higher-order Stark effect and transition probabilities on hyperfine structure components of hydrogen-like atoms", *Hyperfine Interaction*, Vol.127, pp. 287-292, 2000.

[6] A. Derevianko, W. R. Johnson, V. D. Ovsiannikov, V. G. Pal'chikov, D. R. Plante, and G. von Oppen. "Higher-order Stark effect on an excited helium atom", *Phys. Rev.,* Vol.60A, pp. 986 -995, 1999.

[7] W. M. Itano, L. L. Lewis, and D. J. Wineland, "Shift of $^2S_{1/2}$ hyperfine splitting due to blackbody radiation", *Phys. Rev.,* Vol.25A, pp. 1233-1235, 1982.

[8] V. G. Pal'chikov, Yu. S. Domnin, and A.N. Novoselov. "Black-body radiation effects and light shifts in atomic frequency standards", *J.Opt.*, Vol. 5B, pp. S131-S135, (2003).

[9] S. Micalizio, A. Godone, D. Colonico, F. Levi, and L. Lorini, "Black body radiation shift of the ^{133}Cs hyperfine transition frequency", *Phys. Rev.,* Vol.69A, pp. 053401(1-8), 2004.

SHORT AND LONG TERM STABILITY OF KVARZ HYDROGEN MASERS CH1-76 AND CH1-75A

K. Kalliomäki, T. Mansten

[1] MIKES, Electricity Group, Otakaari 7B, FI-02150, Espoo, Finland

Keywords: Hydrogen maser, frequency stability, aging

Abstract

MIKES (Centre for Metrology and Accreditation) has two Kvarz hydrogen masers, passive CH1-76 and active CH1-75A. The phase difference between them is monitored using a high-resolution phase comparator, giving short-term stability as a result. The long-term behaviour is studied with a GPS-monitored Cs-clock as a reference.

The aging of the passive CH1-76 maser seems to diminish exponentially; the time constant is approximately 180 days. At the beginning aging was $5 \cdot 10^{-15}$ per day but now it is less than $3 \cdot 10^{-16}$ per day. There is something peculiar in the active maser CH1-75A, because the aging has been relatively high, $2 \cdot 10^{-15}$ per day with CAT (Cavity Automatic Tuning) on.

The short-term stability of the phase difference of CH1-76 and CH1-75A is excellent, Allan FDEV is $1 \cdot 10^{-12} / \tau$ between 1 and 10 s. At longer integration times it approaches $4 \cdot 10^{-15}$.

1 Introduction

The Electricity Group of MIKES takes care of official time and frequency in Finland and calibrates customers' standards. An excellent short-term stability is important in calibration of atomic reference oscillators of telecommunication industry and naturally for the study of the performance of our own Cs-atomic clocks. As a solution we have acquired two hydrogen masers, first a passive Hydrogen Maser CH1-76, later an active Hydrogen Maser CH1-75A and a corresponding 5/10 MHz Phase Comparator to match the excellent short-term stability of the masers. It is capable of better than $1 \cdot 10^{-13}$ resolution in one second. Masers have been continuously cross-checked using the above-mentioned phase comparator.
Nowadays we can measure the short-term stability of practically any oscillator. In principle the short-term stability of our measurement system is 10 fs (SD) but the stability of the current reference, the passive hydrogen maser, limits it to 1 ps. The active hydrogen maser should improve the stability of our reference by one decade.

As a long term frequency and phase reference we have used HP high stability Cs-clock (HP5071A) which in turn is continuously compared to our four GPS-receivers.

Now both masers have been running continuously, the passive one since February 2003 and the active one since December 2003.

2 Measurement system

Our primary time and frequency standard is HP5071A high stability version supported by another HP5071 standard version and four GPS-receivers. This system guarantees correct long term (> 1 day) frequency, short-term accuracy is limited by inherent stability of Cs-clock to $1 \cdot 10^{-11}$ level (τ = 1s). The phases of second ticks of both hydrogen masers are continuously compared to that of HP5071A at 10-minute intervals using time interval counter HP 53132A.

Because the short term stability of a hydrogen maser is up to two decades better compared to that of caesium clock, we can't use caesium as a reference in short term stability studies. The only way is to compare the masers to each other. The resolution of a phase difference measurement system must be better than 0.1 ps. A special phase comparator (Kvarz) is capable of resolving 0.01 ps, when 5 or 10 MHz outputs of masers are compared.

Masers are sensitive to ambient fluctuations. They need accurately (< ± 0.1°C) controlled ambient temperature. In our case masers are located in a standard laboratory where temperature may vary ±1°C until our new laboratory building is ready. Therefore, we can't obtain the best possible accuracy at the moment.

3 Results

Fig. 1 shows the long-term stability of CH1-76. At the beginning the aging rate was $5 \cdot 10^{-15}$/day but it has decreased exponentially down to less than $0.2 \cdot 10^{-15}$/day during 2 years. It has suffered some electronic component failures, too, see Fig. 1.

Fig. 2 shows the long-term stability of CH1-75A. In this case the aging rate has been stable, around $-2 \cdot 10^{-15}$/day (CAT ON). This rate is out of specifications ($0.3 \cdot 10^{-15}$/day when cavity autotuning (CAT) operates). Therefore, Dr Demidov from Kvarz visited us in Nov. 2004 (MJD 53330) trying to solve the problem, but at the moment the aging rate seems to be the same as earlier.

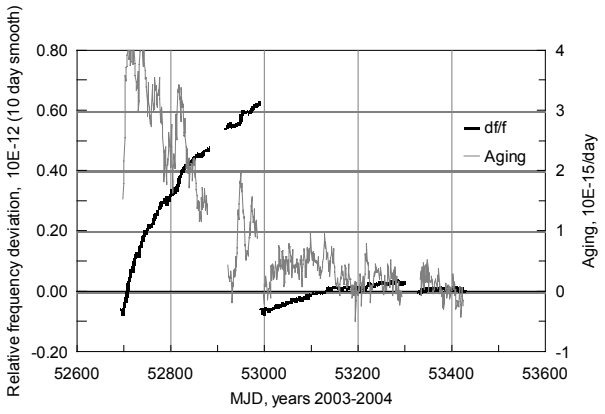


Fig. 1: The long term stability of the passive hydrogen maser CH1-76. Repairs MJD 52910 and 53320, frequency adjustment MJD 52990

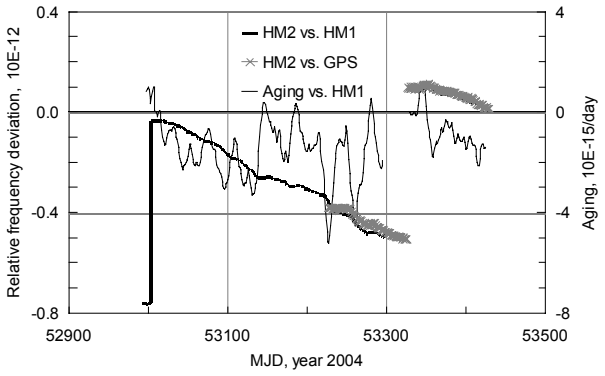


Fig. 2: The long term stability of the active hydrogen maser CH1-75A

Fig. 3 shows the DAC readings of the CAT system and maser frequency (frequency adjustment is removed). Assuming that the unit of the DAC reading is about $1 \cdot 10^{-15}$ (according to the manual) the aging without CAT is around $20 \cdot 10^{-15}$ /day and the CAT "gain" only 10.

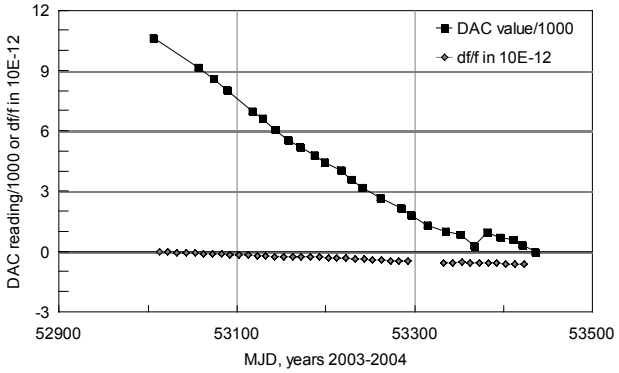


Fig. 3: DAC reading and frequency of active maser CH1-75A

Fig. 4 shows the short-term stability (Allan deviation) of the difference of 5 MHz outputs. Linear drift is eliminated.

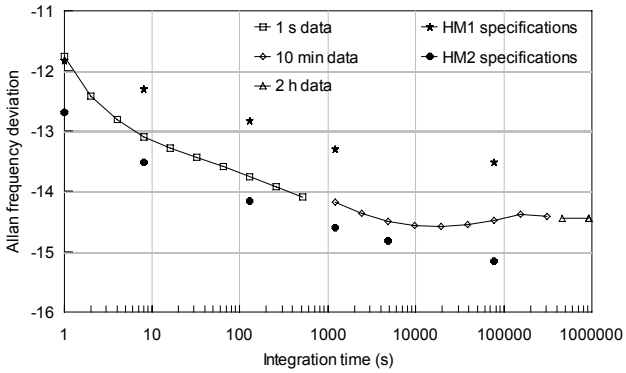


Fig. 4: Stability on of the maser CH1-76, CH1-75A as a reference

The short-term stability of CH1-76 (passive maser) seems to be better than its specifications (stars). Without another active maser it is difficult to judge whether CH1-75A is within its specifications (circles) but it can't be far from them.

4 Discussion

The passive hydrogen maser CH1-76 is clearly within its specifications and its aging rate is very low. The only problem is the reliability of the associated electronics; two simple but fatal component failures have occurred within one year. Our CH1-76 is an old production model; probably new versions have more reliable electronics. The maintenance provided by Kvarz works well but the loss of continuous phase/time is a problem.

The active hydrogen maser CH1-76A works well but its aging rate is still out of its specifications. According to the opinion of Dr Demidov, the problem is in physics package, not in CAT electronics. Its short-term stability is probably better than that of the passive maser as it should be. Slight temperature fluctuations, (up to ±1°C) in our laboratory make it difficult to judge its stability more precisely. The flicker floors seem to be reached in approximately one day.

References

[1] Progress in the Development of IEM Kvarz Passive Hydrogen Masers, N. A. Demidov, A. V. Pastukhov, and A. A. Uljanov, 31st Annual Precise Time and Time Interval (PTTI) Meeting, Dana Point, California, December 7-9,1999, pp. 579-588

THE QUARTZ RESONATOR WITH INDEPENDENT (ORTHOGONAL) DUAL-MODE EXCITATION

Alexander N. Lepetaev, Anatoly V. Kosykh
Omsk State Technical University, Omsk, RUSSIA
644050, Mira Avenue 11. intecs@mail.ru

It is known, that thermocompensated oscillators with best parameters of stability may be performed using dual-mode resonators in which the first oscillation mode is used as reference signal, and another mode – as temperature sensor [1, 2, 3]. However simultaneous excitation of two oscillation modes in one resonator is the difficult problem which has been not solved up until now. Oscillator circuits with dual-mode excitation can provide stable excitation of both modes only at the certain parities of resistance of these modes [3, 4]. At changing of resonator temperature a failures of activity both the basic and a temperature oscillations (dips are independent each of other) may take place; thus the parity of motional resistance of modes can strongly change and fall outside the limits area of stable dual-mode excitation.

It is possible to change a situation if use a resonator excitation by a method of parallel field. The matter is that at use of this method there is the additional degree of freedom connected with a choice of an azimuth of an electrode gap. For each oscillation mode there is such azimuth of a gap at which the given mode can be completely suppressed. Then other oscillation mode can be raised without spurious, but it is necessary that the excitation circuit had sufficient activity and provided excitation of this oscillation mode even at of activity dips. The problem of single-mode excitation is well studied. There is a set of oscillator circuit which can provide excitation of one oscillation mode even at greater changes of its resistance. In this case the problem of dual-mode excitation turns to a problem of creation of such electrodes system at which two independent oscillators excites two oscillation modes, and there is no interaction of these modes.

It is possible to offer some variants of this problem decision.

VARIANT 1. «THE MATCHED ELECTRODE GAP».

The essence of this variant is that for creation of a parallel field bilateral excitation electrodes are used, and gaps in the top and bottom electrodes are shifted each from other. The schematic figure of design cross-section is shown at Fig. 1.

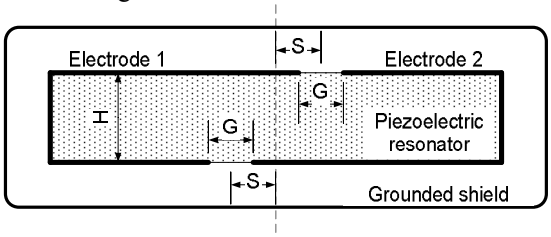


Fig. 1. Resonator with excitation by bilateral electrodes with gape shift (cross-section).

At this figure H is the thickness of resonator plate, G is the width of an excitation gap, and S is the value of shift of a gap concerning the plate center. In such design the effective electric field is directed under a corner to a plane of a plate. Because the polarization vector of concrete oscillations mode also directed under a corner to a plate plane, it is possible to provide an orthogonality of a polarization vector to effective vector of electric field by choosing of shift S. Thus the given mode of oscillations no longer interacts with excitation electrodes. For this statement confirmation the numerical modeling of electric fields in this construction has been carried out (program FlexPDE 4.0.0d was used). For calculation following parameters were used: thickness of a plate of resonator H = 0.54 mm (corresponds to SC-cut resonator excited on third harmonic with resonance frequency of 10 MHz), width of gap width G = 0.1 mm. Anisotropy of dielectric properties was not considered, as it is insignificant (ε changing vs. field direction does not exceed 3 %). The average value of quartz absolute dielectric constant, equal to 40 pf/m, was used in calculation. The potential of the shield was considered equal to zero. Fig. 2 show the picture of lines of electric potential near gap area at S = 0.2 mm.

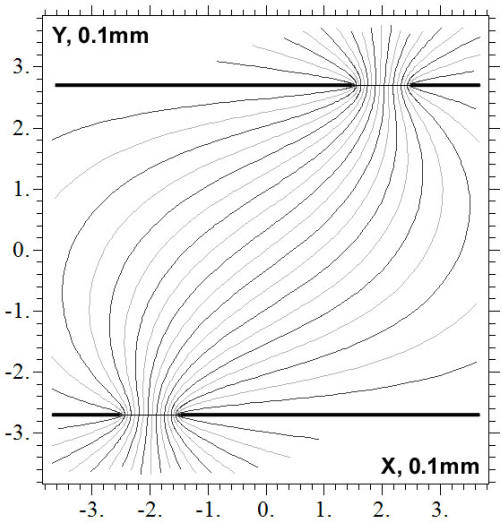


Fig. 2. Lines of electric potential level.

Fig. 3 show the picture of an excitation vector field weighed by relative amplitude of polarization vector on the third harmonic. Also here schematically represented the direction of an electric displacement vector (created by oscillations of the given mode).

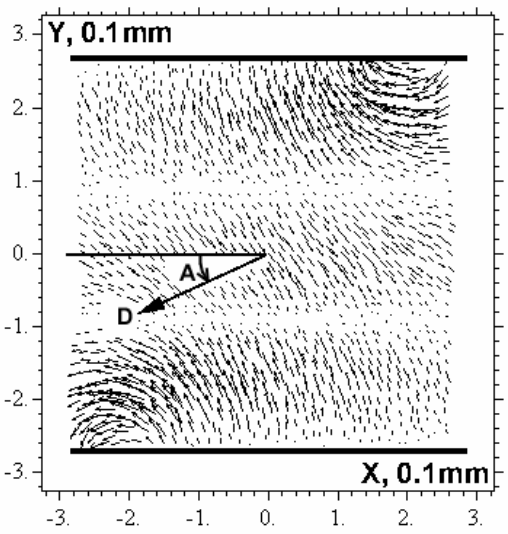


Fig.3. A picture of electric field vectors weighed by amplitude of polarization at excitation on the third harmonic.

It is visible from these figures, that the electric field has some primary direction. Energy interaction of electric field and the electric displacement created by the given mode of oscillations, is defined by expression:

$$effect = \int_V \left(\bar{E} \cdot \bar{D}\right) dV \qquad (1)$$

As gradients of fields in a direction along an electrode gap much less than corresponding gradients in a perpendicular to gap direction, the estimation of interaction energy is possible by integration in this plane of section. At Fig. 4 the values of expression (1) are resulted at integration in a plane of section depending on shift S (at different values of a angle between a direction of projection of polarization vector on a plane of section and a plane of a plate (corner A on Fig. 3)). Numbers on lines of these plots correspond to value in degrees of corner A.

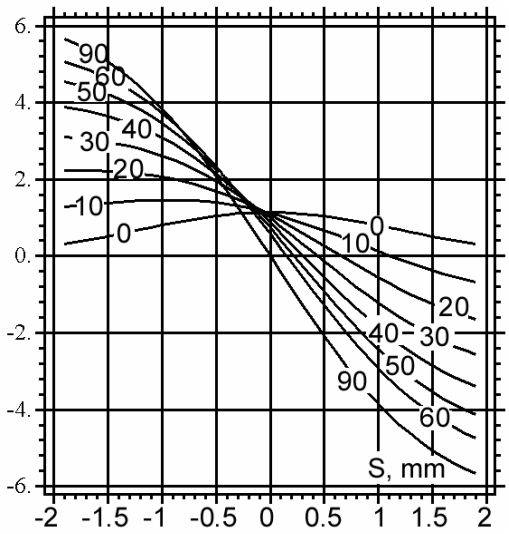
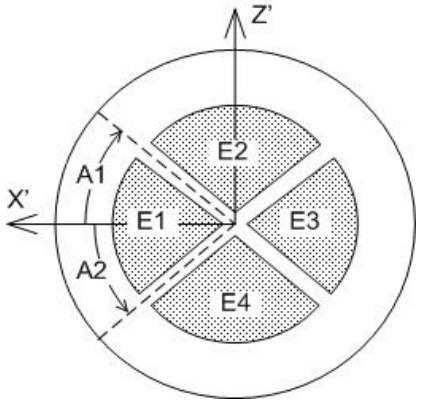


Fig. 4. Normalized couple energy of oscillation.

It is visible from this Figure that for some range of polarization vector corners there are such electrode shift S values at which the energy of electrodes and oscillations interaction becomes equal to zero. Then it is possible to offer a design of the resonator with orthogonal excitation which is schematically represented on Fig. 5. At this figure the Y'-axis is perpendicular to drawing planes and directed aside the observer. Two electrode gaps divide electrodes into 4 parts. Azimuths of gaps are designated as A1 and A2. The azimuth of a gap is lay off from X'-axis aside Z'-axis Four electrodes are labeled as E1, E2, E3, E4 accordingly. For definiteness, we shall realize numbering from X'-axis clockwise. An azimuth of gaps selects to provide full (or partial) suppression of one mode in one gap and other mode in other gap.



(1)

Fig. 5. The schematic image of the resonator with orthogonal modes excitation.

In such resonator two opposite electrodes must connected to the common wire, and two remained electrodes must connected to two separate oscillation circuits (we shall call it as G1 and G2) which will excite different modes of oscillations (are named as «1» and «2»). Possible variants of connections are listed below in the table.

Variant	1	2	3	4
E1	Comm.	Comm.	G1	G1
E2	G1	G1	Comm.	Comm.
E3	Comm.	Comm.	G2	G2
E4	G2	G2	Comm.	Comm.
Gap E1-E2	Mode 1	Ortog. 2	Mode 1	Ortog. 2
Gap E2-E3	Ortog. 2	Mede 1	Mode 2	Ortog. 1
Gap E3-E4	Ortog. 1	Mode 2	Ortog. 1	Mode 2
Gap E4-E1	Mode 2	Ortog. 1	Ortog. 2	Mode 1

Here «Ortog. 1» and «Ortog. 2» corresponds to such values of shift S at which orthogonality of a field of excitation in relation to corresponding mode of oscillations is provided. «Mode 1» («Mode 2») are designate, that the S-value is chosen to provide the maximal coupling with the mode 1 (mode 2).

Now the resonator turns from a bipolar component to a two-port component that complicates its model. At Fig. 6 one of possible variants of representation of the equivalent scheme of such resonator, presented in the form of macromodel of programs MicroCAP is

introduced. Here components E1…E4 are the voltage sources controlled by a voltage (VofV) (they modeled a direct piezoeffect), G1…G4 – the current sources controlled by voltage (IofV) (they model reverse piezoeffect). Four resistors of 1t (1 teraOhm) value, connected in parallel to inputs, do not influence on properties of model and are necessary only for elimination of "hanged" nodes. Parameters of model are values of motional resistance, frequencies and Q-factors of modes C and B at zero value of electrode shift S and at parallel connection of electrodes segments at which it turns to an ordinal bipolar component. Parameters of mode C are measured at connection c1b1 to c1b2, and c2b1 to c2b2, and parameters of mode B are measured at connection c1b1 to c2b1 and c1b2 to c2b2. Parameters kc1, kc2, kb1, kb2 corresponds to correction factors depending on size of gape shift S, for four gaps. Index b and c in designations of these factors corresponds to a designation of a mode excited by a gap, figures in the end designate conditional number of a pole of an electrode for other mode (an equivalent of "plus" and "minus"). An individual (by default) value of these factors corresponds to zero value of electrode shift S. Values of these factors can be received from Fig. 4. by normalization on value at S=0.

VARIANT 2. «THE COMPENSATED EDGE».

As in the first variant only a half of full electrode gap is used for excitation of any mode, a quite logical step is eliminate one of grounded electrodes.

In some cases using of first variant is impossible. As apparently from Fig. 4, at small values of angle A (see Fig. 3) for maintenance of separation from one of oscillations modes the value of shift S can appear excessively greater. At A = 0 the full separation is impossible at any value of S.

It is possible to change an electrode configuration to reduce value of S and to make its comprehensible even at A=0. For this purpose the model of the resonator with grounded electrode removed has been considered. Calculations have shown that removal of such electrode almost does not influence on energy of interaction, and plots on Fig. 4 almost do not change. It is related by that the grounded screen (which is present at model of the resonator on Fig. 1) incurs a role of the removed electrode.

However it is possible to change a configuration of an electric field near to electrode edge if to add there an additional grounded electrode performed in the form of a narrow strip, parallel to electrode edge. We shall name this additional electrode as "compensator". The schematic diagram of cross-section of such design is shown at Fig. 7.

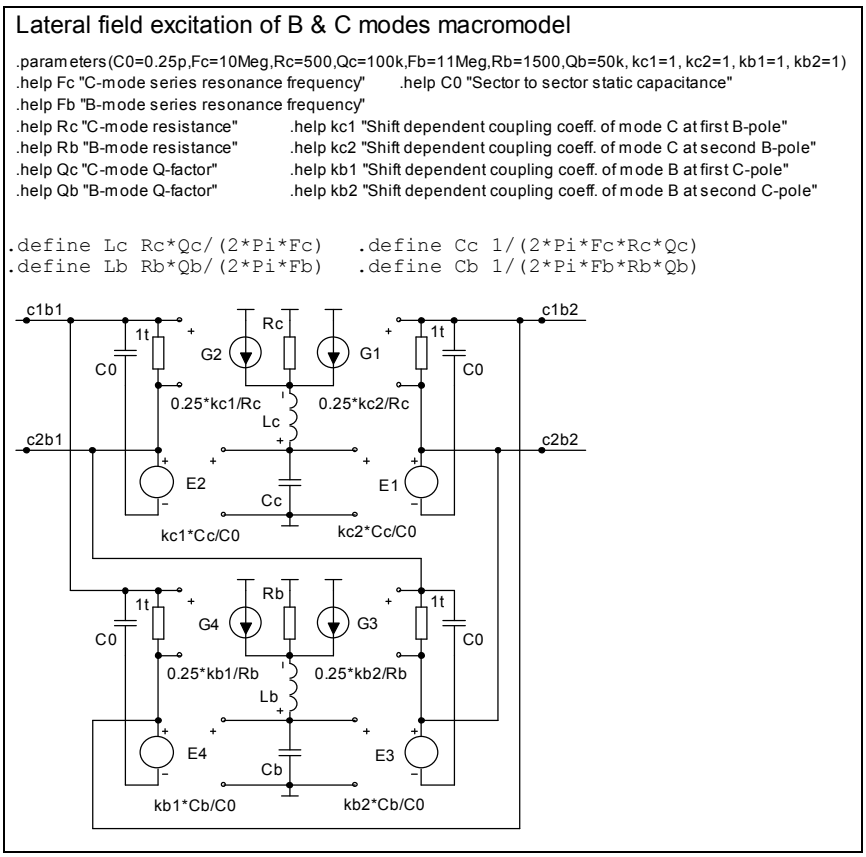


Fig.6. The macromodel of the two-port resonator

The results of coupled energy calculation (similar to shown at Fig. 4) at W=0.4 mm and G=0.1 mm are given at Fig. 8. It is visible from this figure, that now for any value of a angle A there is value S at which the separation with the given mode is carried out.

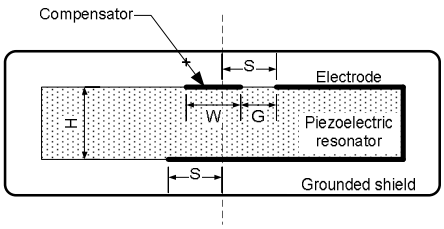


Fig. 7. The schematic image resonator design of with a free electrode and the compensator

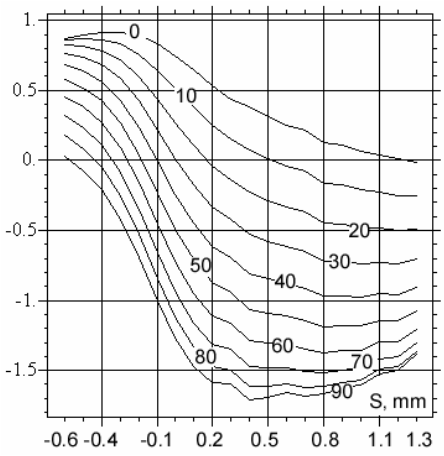


Fig. 8. Normalized free edge coupled energy with compensator.

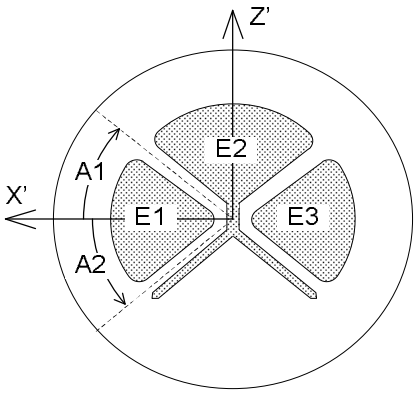


Fig. 9. The schematic image of the resonator with orthogonal excitation of modes and use of the compensator

Fig. 9 presents the schematic drawing of a resonator electrodes design for such method. In such design there is only one variant of orthogonal excitation when electrode E2 is the common wire, electrode E1 is connected to G1, electrode E3 is connected to G2, gap E1-E2 is used for excitation of a mode 1, gap E2-E3 is used for excitation of a mode 2, free edge E1 is separated from a mode 2, free edge E3 is separated from a mode 1 (the separation is made with use of the compensator).

PRACTICAL CALCULATIONS FOR SOME KNOWN CUTS.

For definition of concrete values of parameters of a design of the resonator (azimuths of gaps A1 and A2, and values of size of shift S) calculations of orientation of vectors of polarization have been made for some cuts. The angle φ between X'-axis and a projection of a vector of polarization to a plane of the resonator (the positive direction corresponds to turn from X'-axis in a direction to Z'-axis) and an angle θ between Y'-axis and a vector of polarization was defined. Results of calculations are listed below in the table.

Cut	Rotate angles		Mode C		Mode B	
	α (°)	β (°)	φ (°)	θ (°)	φ (°)	θ (°)
BT	0	-49.2	0	90	90	49.2
AT	0	35.25	-90	35.3	180	90
V1	5	35	-68	37.1	-171.6	78.2
DP	13.9	34.67	-38.3	48.1	-155.3	58.8
FC	15	34.67	-35.7	49.9	-153.3	57.0
IT	19.1	34.25	-26.4	56.8	-144.4	49.5
SC	22	34.5	-22.3	61.1	-140.9	47.5
RT	15	-34.5	170.2	76.1	44.8	44.3
TC	13.9	-32	172.6	78.3	45.1	41.4

Angle A (Fig. 3) may be obtained as: A = 90 – θ, and value of S may be obtained from curves in Fig. 8.

The feature of resonators with parallel excitation field is the small range of the electrical tuning, therefore it is expedient to use the resulted decision for ovenized oscillators [6] or for thermocompensated oscillators on base of frequency synthesizers [5]

REFERENCES

1. US Pat. No. 4079280, J. Kusters, et al., Quartz resonator cut to compensate for static and dynamic thermal transients, March 1978.
2. A. V. Kosykh, I. V. Abramzon, V. P. Bagaev, "Dual mode oscillators with resonators exited on B and C modes", Proc. 1994 IEEE Frequency Control Symp., 1994, pp. 578-586.
3. I. V. Abramzon, "Two-mode quartz resonators for digital temperature compensated quartz oscillators". Proc. 1992 IEEE Frequency Control Symp., 1992, pp. 442-447.
4. S. S. Shodowski, "Resonator self-temperature-sensing using a dual-harmonic-mode crystal oscillator", Proc. 43rd Ann. Frequency Control Symp., 1989, pp. 2-7.
5. A. Kosykh, V. Bagaev, B. Ionov, A. Lepetaev, S. Zavjalov, A. Vasiliev. The New Method of Statistic Piecewise-Linear Interpolation and Its Application To DTCXO Creation. Proc 1993 IEEE AFCS, pp. 687 – 697.
6. Y. Watanabe, K. Ozaki, S. Goka, T. Sato, H. Sekimoto. Frequency Stability and Phase Noise Characteristics of Digital OCXO Using Dual-Mode Excitation. Proc 2001 IEEE FCS and PDA Exhibition, pp. 790 – 79.

TWO-DIMENSIONAL COLLIMATION OF A CONTINOUS CESIUM FOUNTAIN IN AN OPTICAL LATTICE

N. Castagna[1], G. Di Domenico[1], G. Dudle[2], F. Füzesi[1], M.D. Plimmer[1], A.V. Taichenachev[3] P. Thomann[1] and V.I. Yudin[3]

[1]Observatoire Cantonal, Rue de l'Observatoire 58, Neuchâtel, CH 2000 Switzerland Natascia.Castagna@ne.ch
[2]Swiss Federal Office of Metrology and Accreditation (METAS), CH 3003 Bern-Wabern Switzerland
[3]Novosibirsk State University Pirogova 2, Novosibirsk 630090, Russia and Institute of Laser Physics SB RAS, Lavrent'eva 13/3 Novosibirsk 630090 Russia

Keywords: Zeeman and Stark effects, mechanical effects of light on atoms, molecules, and ions, laser cooling, atomic and molecular beam sources and techniques.

Abstract

We present studies of two-dimensional transverse laser collimation of a continuous cold cesium fountain. We compare both red-detuned and blue-detuned optical molasses resulting from polarisation gradient cooling on all hyperfine components of the D2 line. We also give results for magnetically induced laser cooling on the F=3-F'=2 transition and show how it enables efficient diagnosis of transverse magnetic fields.

1 Motivation

The S.I. second is defined in terms of the frequency of the transition between the ground state hyperfine levels of the Cs atom. The best realisations of the second use atomic fountain clocks operating in a pulsed mode. Our group has developed a continuous Cs fountain clock, (FOCS1) now installed at the Federal Office of Metrology and Accreditation (METAS) in Bern. Continuous fountains have the advantage of a lower atomic density for a given signal-to-noise ratio than their pulsed counterparts, which helps minimise collisional shifts. Furthermore, the use of such a fountain enables one to eliminate the Dick effect [4,10] associated with local oscillator noise. In order to take full advantage of the continuous fountain approach, however, we need to increase the useful flux by collimation of the atomic beam.

In previous work, our collaboration reported collimation in a two-dimensional lattice using Zeeman induced degenerate Raman sideband cooling in a far-off-resonance lattice [2,5]. This provided temperatures as low as 1.6 μK and a fourfold gain in flux compared with the simple moving molasses (transverse temperature 60 μK). However, with this technique the atoms are prepared in the magnetically sensitive F=3, m_F=3 state, unsuitable for a primary time standard. A more promising approach proposed by two of us [18] would be to use the a.c. Stark effect of a linearly polarised laser tuned near resonance with the Cs D1 or D2 lines, since this prepares atoms in the m_F=0 sublevel, ideal for clock applications. This technique is harder to put into practice due to hyperfine depumping by the Stark shifting laser. Furthermore, in both approaches, the efficiency of the process is improved if atoms are first pre-cooled to a temperature below 10 μK so that they are all trapped in the potential wells of the 2D lattice. This pre-cooling could be achieved using a separate lattice located just below the sideband cooling one, the pre-cooling laser being tuned to a suitable transition. In addition, however, one might arrange for Sisyphus cooling to take place in the same lattice as that producing sideband cooling. Thus, we have undertaken a broad study of sub-Doppler cooling in our 2D linear-θ-linear lattice, the geometry of which is described in [5](Fig. 1).

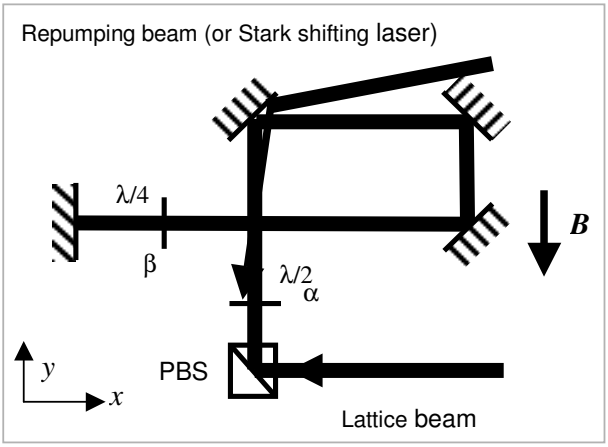


Fig. 1: Optical lattice geometry. Two wave plates allow us to control the linear polarization (α and β are the tilt angles with respect to the vertical). PBS: Polarizing Beam Splitter; **B** is the static magnetic field.

Here a laser tuned to a hyperfine component of the Cs D2 line is used in a phase-stable, power-recycling geometry with coplanar, counterpropagating beams intersecting at right angles. This is simpler to implement than the minimal lattice with three independent coplanar beams interesecting at 120° and has the intrinsic advantage of higher power and thus deeper lattice potentials [6]. We have compared laser cooling on all hyperfine components of the D2 using a repumping laser on either the D1 or D2 lines. Note that the use of a D1 repumper precludes the existence of second-order processes such as non-degenerate Raman transitions. While studies on the cycling 4-5' transition are ubiquitous [13], and those on the 3-2' transition well documented [19,21], work on the other lines is somewhat rare.

In this article we describe two cases of interest. The first is where $\theta=\pi/2$ ($\alpha=22.5°,\beta=0$ in Fig. 1), the lin-perp-lin configuration with maximum polarisation gradient. The second is laser collimation induced by a transverse magnetic field when $\theta=0$ (*i.e.* $\alpha=\beta=0$).

2 Experimental set-up

The apparatus, sketched in Fig. 2, is essentially the same as that described in [5]. Cs atoms from a moving optical molasses are launched continuously in a parabolic flight with an initial longitudinal velocity v_z of 3.6 m/s [1]. In a region located 22 cm above the source, the slow beam is collimated using the 2D linear-θ-linear optical lattice, $1/e^2$ intensity radius = 5.7 mm, truncated at a radius of 9 mm.The lattice laser is tuned near resonance with a given hyperfine component of the Cs D2 line. The lattice beams are reflected from gold-coated mirrors to maximise reflectivity (R>97 %) and minimise birefringence; improvement is expected from silver ones [8]. At the end of the flight, the atomic flux is probed via laser-induced fluorescence of the F=4 to F'=5 transition of the D2 line. By adding a D2 re-pumping beam in the probe region tuned to the F=3 to F'=4 transition, we can detect the atoms arriving in both the F=3 and F=4 hyperfine levels. We measure transverse temperatures by translating laterally the whole detection system (probe and repumping lasers and detection optics) in the horizontal plane and measuring flux as a function of position. The data are fitted to a gaussian curve (collimated atoms) superimposed on a slope representing the wings of a much broader distribution of uncollimated atoms. Depending on the atomic flux, a resolution of of 0.2 µK can be achieved.

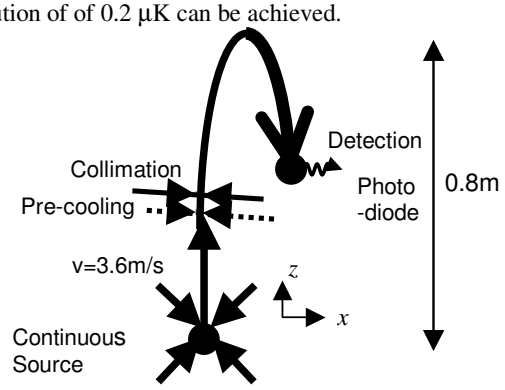


Fig. 2: Scheme of the experiment. The source of the atomic beam is a six beam moving molasses, four of which are shown (horizontal beams in and out of the page).

3. Polarisation gradient cooling and the role of the repumping transition.

The Cs D2 line has six hyperfine components resulting from the hyperfine splitting of the $6s\ ^2S_{1/2}$ and $6p\ ^2P_{3/2}$ levels (Fig. 3). The vast majority of laser cooling experiments on the Cs atom use the cycling F=4-F'=5 transition, with a rempumping laser tuned to a hyperfine component of the D2 line. The lowest temperatures reached are about ten times the recoil limit. In our work geared towards sub-recoil cooling in an optical lattice, however, we need to consider the possiblity of cooling on other hyperfine

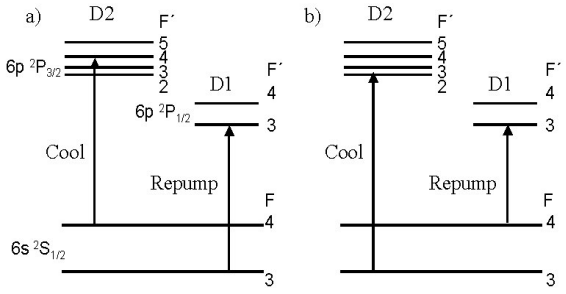


Fig. 3: Energy levels of Cs used for laser cooling and repumping transitions. The wavelengths are D1: 895 nm, D2:852 nm.

Line		F'=2	3	4	5
D1	F=3		1/4	7/12	
	F=4		3/4	5/12	
D2	F=3	1	3/4	5/12	0
	F=4	0	1/4	7/12	1

Table 1: Branching ratios for spontaneous emission F-F' of the hyperfine components of the Cs D1 and D2 lines.

components of the D2 line as well as the use of a D1 repumping laser. The reasons for this are several fold, each transition having its own particular advantage or special feature. In this section we consider in turn each of the six hyperfine components. We have studied the effects of lattice laser power, repumper power and frequency as well as the influence of the angle θ between the polarisation vectors of the lattice beams. Depending on the particular repumping hyperfine transition (a component of the D1 or D2 line), whether or not the repumping laser is exactly resonant or just close to resonance, uniformly polarised or not, laser cooling is more or less efficient, or even impossible. Given the huge number of possible configurations, we do not yet have a complete picture. Figs 5 and 7 show how the atomic flux density at the end of our fountain varies as a function of D2 laser frequency for the repumping laser tuned to different hyperfine components of the D1 line. (The D1 laser is a travelling wave directed nearly along the *y*-axis of the lattice (*xy*) plane, the pumping laser of Fig. 1. As a result, this laser alone cannot produce polarisation gradient cooling). For the results presented in this section, the polarisation of the lattice laser beams is lin-perp-lin with polarisation vectors inclined at 45° to the vertical. Magnetic fields are cancelled to within a few milligauss using the technique described in section 4. The role of branching ratios and dark states is found to be essential to explain our observations. Let us consider each transition in turn.

F=4-F'=5 and F=3-F'=2

The 4-5' transition is currently the one used to collimate the continous beam in the FOCS1 fountain clock. At present, two independent standing waves tuned to the 44'-45' crossover (44' +175 MHz) produce a 2D optical molasses. The resulting temperature is 7 µK. In our work on optical lattices we have

obtained a lower temperature (3.6 μK) for the same detuning (Fig. 4). This improvement we ascribe to the greater phase stability of the folded lattice compared to the two-standing-wave molasses. Fig. 4 shows how the transvserse temperature varies with lattice power in this case for detunings around the crossover.

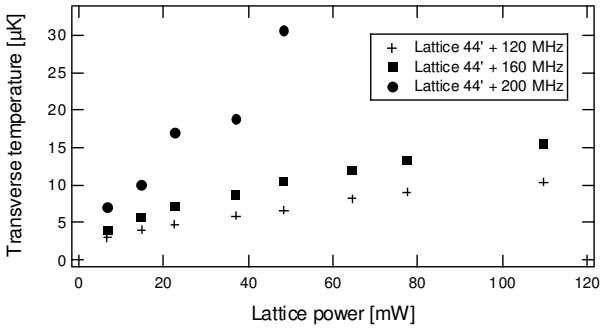


Fig 4.: 2D Sisyphus cooling using an optical lattice tuned between the 44' and 45' components of the D2 line. Variation of transverse temperature with lattice power for different detunings. The repumping laser is tuned to D2 34'.

It would be natural to implement this geometry for the next fountain clock FOCS2. Comparable results can be obtained, however, if one cools instead with a lattice on the F=3-F'=2 transition. In this case, the laser must be tuned to the blue to produce grey molasses. With blue detuning, atoms accumulate around regions of low light intensity where there is less scattering and hence less heating due to spontaneous emission. In our case, we obtain similar temperatures to those with the 4-5' transition. Under the most favourable conditions ($\delta = +3\Gamma$, $I \approx 3$ mW/cm^2) the required laser power is lower and the flux density about 10% higher. For this reason and because we are interested in a pre-cooling mechanism for Stark sideband cooling (where the atoms are trapped in F=3), we have studied in detail the 3-2' transition. Like the 4-5' hyperfine component, it is a closed transition so the repumping laser plays only a minor role. At high laser intensity we have found a departure from usual grey molasses similar to the case reported by Lucas *et al.* [12] for Rb. In a search for the lowest temperature and maximum flux density, we have studied the effect of laser power for a fixed detuning (Fig. 6). There appear to be two zones leading to different slopes of temperatures versus power, one at low power where grey molasses operates,. the other at larger power where higher temperatures are obtained. In neither case, however, does the temperature vary as the ratio of Rabi frequency squared to detuning, as it would for Sisyphus cooling [3].

A search for an optimum angle between the laser polarisation vectors also led us to consider the case of $\theta=0$ where the influence of transverse magnetic fields is decisive. This is discussed in section 3 below.

F=4-F'=4 and F=3-F'=3

For transitions of the type $\Delta F=0$, laser cooling occurs for blue detuning. This means that, for example, a laser red-detuned from the 4-5' line produces some cooling on the 4-4' line too.

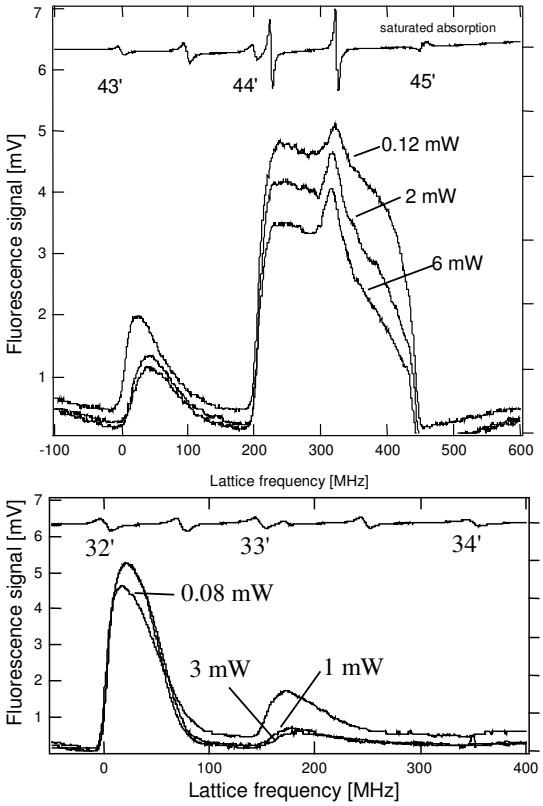


Fig. 5: Two-dimensional collimation of the continuous Cs fountain as a function of D2 laser frequency for different D1 repumping powers. Upper figure: lasers as in Fig. 3a):Lower figure: lasers as in Fig. 3b). The dispersion curves are the saturated absorption error signals used to identify the hyperfine components.

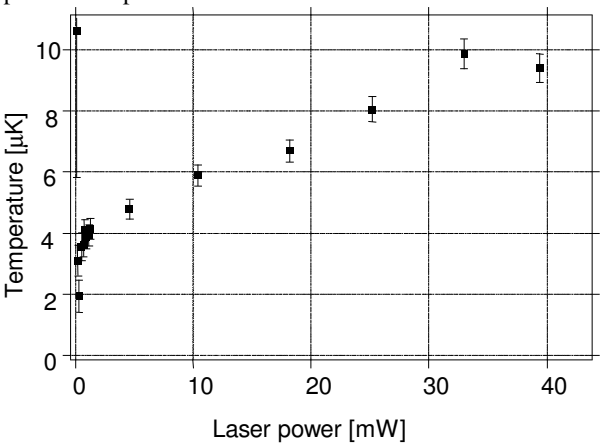


Fig. 6: Laser cooling on the F=3-F'=2 transition. Study of transverse temperature as a function of laser power for a blue detuning of 3Γ. Repumping laser 0.1 mW tuned to D1 43'

The 4-4' line is of interest to us because in the Stark sideband cooling scheme, the lattice laser might be locked to the 4-4' transition. For a lattice laser exactly resonant with 44', we should get no Sisyphus cooling due to the 44' line itself, just the far red wing of the 45' transition. Indeed some studies

elsewhere considered only the role of the 4-5' transition, whatever the laser detuning [7]. However, by detuning the lattice to the blue of 44', we might produce some Sisyphus cooling [11] which would help pre-cool untrapped atoms. The cooling efficiency on open transition, such as 44', is determined by the effectiveness of the repumping laser. Different results are obtained for a D1 repumper tuned to either the 3-3' or 3-4' hyperfine components (Figs. 5 and 7). In particular the 3-3' transition repumps atoms more efficiently back to F=4 from where they begin a new cooling cycle. For the D2 line, the branching ratios of these transitions are inverted as shown in Table 1. This means that, for example, wheras the D1 3-3' repumper works best for a given cooling transition, for a D2 repumper, 3-4' works better. In the simplest Stark cooling scheme, the D1 F=3 laser is blue detuned by a few linewidths from 3-3' so its repumping efficiency is less. This degrades the pre-cooling efficiency and suggests the need for an additional laser tuned to 3-4' to optimise the pre-cooling process. The most striking aspect in Figs. 5 and 7 is perhaps the existence of a peak midway betwen the 44' and 45' transitions. This "*crossover cooling*" feature, where both red-detuned and blue-detuned molasses seem to co-exist, merits further investigation.

For comparison we have studied laser collimation due the 3-3' transition using a D1 repumper laser on 4-3' or 4-4'. Indeed one might image a Stark sideband cooling scheme in which the far-off-resonance lattice is locked to 33' so as to trap atoms in F=4. The Stark shift would be produced by a laser tuned to the blue of 4-4' which would serve as a weak repumping laser. Using in our case a D1 repumper resonant with 4-3', we have measured temperatures as low as 3.0(5) µK for P=1 mW and a detuning of 3Γ. The maximum flux is about three times lower than for the 3-2' transition, making this component the less attractive of the two for simple collimation. The curve of flux versus detuning is of comparable width for both 3-2' and 3-3.'

F=4-F'=3 and F=3-F'=4

For completeness we have also studied laser cooling on the remaining two hyperfine components of the D2 line. For the cooling laser tuned to the blue of the 4-3' transition, laser collimation can be obtained provided the D1 repumping laser is locked to the hyperfine component 33' (Fig. 5) but not 34' (Fig. 7).

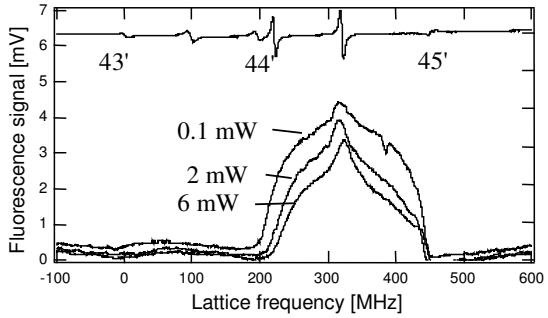


Fig. 7: Two-dimensional laser collimation using the F=4-F'=3,4,5 components of the D2 line. Here the D1 repumper is tuned to 3-4' and no collimation is observed on the D2 4-3' component.

In order for laser cooling on an open transition to be efficient, there must be effective repopulation of the ground state hyperfine level by the repumping laser. Two factors govern this effectiveness: hyperfine branching ratios and the existence of dark states. Hyperfine depumping implies heating by the repumping process. If a mean number of spontaneous photons $\langle N_{sp} \rangle$ close to unity is emitted per cooling cycle, laser cooling can occur without too much heating. On the other hand, if $\langle N_{sp} \rangle$ is too high, say close to 2 or more, there will be greater heating due to spontaneous emission. In the case of cooling on the 4-3' transition, $\langle N_{sp} \rangle$ =4/3 with a D1 repumper on 3-3' but 12/5 with a D1 repumper on 3-4'. Dark states come into play when the repumping laser is of constant polarisation (no gradient). An atom falls into a state from which no re-excitation is possible and is lost forever from the cooling cycle. Yet hyperfine redistribution seems to predominate- otherwise one would expect D1 3-4' (no dark state) to be a better repumper than 3-3' (one dark state, here m_F=0 for vertical linear polarisation).

When we lock the cooling laser to 4-3' and scan the D1 repumper frequency, we observe some interesting curves whereby cooling occurs when the repumper laser is detuned from resonance but not when it is on resonance. This implies that at resonance hyperfine depumping by the repumper is too strong.

In the context of Stark sideband cooling, we could produce a far-off-resonance lattice to trap atoms in F=3 with a laser tuned to the 4-3' hyperfine component. By detuning the lattice laser to the blue, some Sisyphus pre-cooling could be obtained provided 3-3' were an efficient repumper (which it is for the D1 line).

Finally, let us consider the 3-4' transition. Quite simply, we never observe laser cooling on this hyperfine component, whatever the experimental parameters. We believe this is due to two factors. First, the branching ratio from $^2P_{3/2}$ F=4' means the F=4 ground hyperfine level is populated more than F=3 so the repumping laser is a critical element (Table 1). Second, both repumping transitions 4-3' and 4-4' have dark states for our linearly polarised D1 repumper so neither can be effective enough. Further investigation, using for example a repumper with a polarisation gradient , is needed to elucidate this matter and might also shed light on the D2 4-3' cooling transition as well.

4. Magnetically induced laser cooling

For the Stark-induced degenerate Raman sideband cooling mentioned in the introduction, it is desirable to cancel magnetic field components *in situ* in the region of the optical lattice. The best way to do this is to use the atoms themselves as a probe. To detect fields at the few milligauss level, a measurement of the Zeeman shift (a few kilohertz) on the optical transition is too insensitive due to the large natural width of the D2 line (5.3 MHz). Non-degenerate Raman transitions between F=3 and F=4 ground hyperfine levels would be a suitable spectroscopic tool but require phase-locked lasers. By far the simplest way is to use magnetically induced laser cooling (MILC). Here, in absence of a polarisation gradient, laser collimation can still take place

provided a transverse magnetic field is present. The measurement of the atomic flux density at the end of the fountain thus provides a measure of such a field. One can then adjust the field components to minimise the flux. We have studied laser collimation essentially on the F=3-F'=2 component of the D2 line. A detailed communication on this subject has been submitted for publication elsewhere [16], so here we provide only a brief summary including some data not supplied therein.

Although MILC was described more than a decade ago [17,20,23,24], most work has been on J to J+1 transitions. Indeed, we observe the effect for red-detuned molasses using the F=4-F'=5 component of the Cs D2 line. By contrast, published data on J to J-1 transitions are rarer. The group of Metcalf mentions results for the textbook case of 1D collimation of a thermal ^{87}Rb beam on the F=1 to F'=0 transition using a weak transverse magnetic field and a $\sigma^+\sigma^+$ standing wave with no polarisation gradient [9] and obtained unpublished data for the lin-parallel-lin case [15]. Nienhuis *et al.* [14] calculated a graph of reduced force versus reduced velocity for 1D cooling of a J=2 to J' =1 transition in a similar configuration but with a strong magnetic field. The only work on Cs of which we are aware is that of Valentin *et al.* [21,22,25] who studied the F=3 to F'=2 component of the D2 line. In all these experiments, a thermal atomic beam was collimated in a 1D standing wave to sub-Doppler transverse temperatures in either strong magnetic fields (Larmor frequency $\Omega_Z \geq$ ground-state light shift Δ_{LS}) or weak ones ($\Omega_Z < \Delta_{LS}$). In all of the above, results describe cooling to zero transverse velocity in low fields and non-zero velocity in high fields. All calculations were numerical simulations and, to the best of our knowledge, no expressions for the transverse temperature $T_\perp$ as a function of $B_\perp$ were given. Our own experiment is different in that it concerns 2D collimation of a continuous Cs fountain using an optical lattice with parallel linear polarisation vectors; the magnetic fields are always weak ($B_\perp$<1.2 G) and the atoms, already cold to begin with (longitudinal temperature 75 µK, $T_\perp$ =60 µK), are transversely cooled to about 5 µK.

In short, we have measured flux density and transverse temperature as a function of transverse magnetic field for different laser powers (1 to 32 mW) and detunings (3Γ to 15Γ). As shown in Fig. 8, the characteristic features of this process are minimal cooling in zero transverse field and a quadratic increase in *T versus* B away from the maximum. The T^2 coefficient varies as (const.+1/*P*) where *P* is the lattice power (Fig. 9). The flux density versus magnetic field curves exhibit a dip, the width of which ΔB varies as the cube root of the lattice intensity. This behavoiur can be understood from simple theoreticla arguements. Whereas grey molasses, produced on the 3-2' transition with a polarisation gradient is destroyed by an applied magnetic field, be it transverse or longitudinal, the mechanism studied here only works when a transverse magnetic field is applied. By measuring the flux density, we are able to cancel out transverse fields at the few milligauss level in this way.

5. Conclusion

We have studied two-dimensional cooling of a continous cold Cs fountain using a phase-stable, power recycling optical lattice. As shown by Table 2, the use of a lattice in place of two independent beams leads to higher flux and lower temperatures. We have investigated polarisation gradient laser cooling on all hyperfine components 3-F' and 4-F' of the D2 line using either a repumping laser tuned to a hyperfine component of the D1 or D2 lines. In addition to the well-known 45' cycling transition, the 32' line looks useful both for beam collimation and preparation of atoms for more efficient degenerate Raman sideband cooling. The 33' and 44' components could be used for Sisyphus cooling in the same lattice as that used for Raman cooling. Collimation on the 43' transition can be obtained with a suitable repumping transition. By contrast, with a repumping laser of uniform polarisation, no cooling looks possible on the 34' line, no matter what the repumping laser frequency. Using the 32' transition with no polarisation gradient, we have identified magnetically induced laser cooling as a sensitive way to detect and hence cancel transverse magnetic field components at the few millgauss level.

Cooling mechanism	Flux (10^6 at.s^{-1})	$T_\perp$ (µK)	Comment
2-beam Sisyphus 44'-45' crossover	0.2	7	Used on FOCS 1
Sisyphus 44-45 crossover (lattice)	2.6	3.6	Characteristic peak
Grey molasses 32'+3Γ in lattice	1.9	3.5	
Zeeman-induced degenerate Raman sideband	0.5	1.6	Requires magnetic fields

Table 2: Flux of atoms in m_F=0 and lowest transverse temperatures measured at the end of the fountain obtained from different collimations schemes thus far investigated. 45' : $^2S_{1/2}$ F=4-$^2P_{3/2}$ F'=5 *etc.*

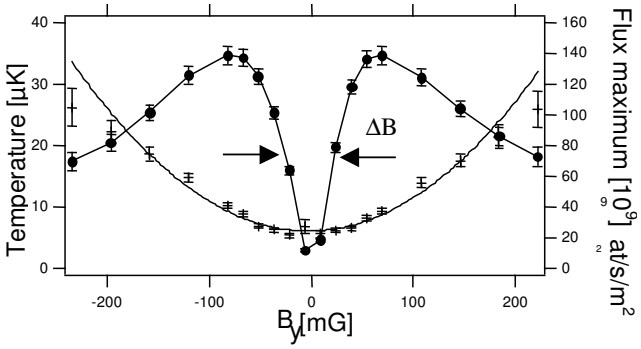


Fig. 8: Variation of transverse temperature (+) and flux (•) as a function of transverse magnetic field for P=8 mW. The lower solid line is a quadratic fit to the temperature data.

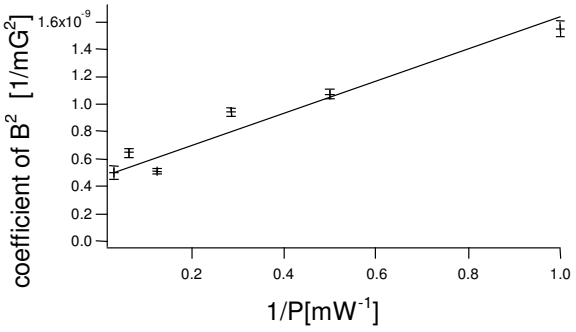


Fig. 9: Variation of B^2 coefficient of the temperature curve of Fig. 8 as a function of inverse lattice power (1/*P*) for laser cooling on the F=3-F'=2 component of the D2 line.

Acknowledgements

AVT and VIY are grateful to the Swiss National Science Foundation (FNRS) for funding a stay in Neuchâtel during which much of this work was performed. The authors thank D. Boiron and S. Pádua for kindly supplying copies of Ph.D. theses. The experimental work was supported by FNRS and METAS.

References

[1] P. Berthoud, E. Fretel and P. Thomann, "Bright, slow, and continuous beam of laser-cooled cesium atoms", *Phys. Rev. A* **60** ,R4241-4244 (1999).
[2] N.Castagna, G. Di Domenico, G. Dudle, G. Mileti, M. Plimmer, A.V. Taichenachev, P. Thomann, V.I. Yudin "Collimation of a continuous cold cesium beam using Raman sideband cooling". Proc. EFTF 04, Guildford, UK.
[3.] J. Dalibard and C. Cohen-Tannoudji, "Laser cooling below the Doppler limit by polarization gradients: simple theoretical models", *J. Opt. Soc. Am. B* **6**, pp. 2023-2045 (1989).
[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" Proceedings of the 22nd annual PTTI Meeting pp. 487-508 (1990).
[5] G. Di Domenico, N. Castagna, G. Mileti, P. Thomann, A.V. Taichenachev and V.I. Yudin, "Laser collimation of a continouous beam of cold atoms using Zeeman-shift degenerate-Raman-sideband cooling" *Phys. Rev. A* **69**, pp. 063403-1 – 063403-9 (2004).
[6] G. Di Domenico, N. Castagna, M.D. Plimmer and P. Thomann, "On the stability of optical lattices" *arXiv physics /0412072.*
[7] H. Ellmann, J. Jersbland and A. Kastberg "Temperature in 3D optical lattices influence by neighbouring transitions" *Eur. Phys. J. D* **13** pp. 379-384 (2001).
[8] C.W. Goodwin, private communication (2004).
[9] R. Gupta, S. Padua, C. Xie, H. Batelaan and H. Metcalf., "Simplest atomic system for sub-Doppler laser cooling", *J. Opt. Soc. Am. B* **11**, pp. 537-541 (1994).
[10] A. Joyet, G. Mileti, G. Dudle, and P. Thomann, "Theoretical Study of the Dick Effect in a Continuously Operated Ramsey Resonator" *IEEE Trans. Instr. Meas.* **50**, pp. 150-156 (2001).
[11] M. Kajita, N. Kotake, A. Hasgawa, T. Morikawa, "Influence of neighboring transition on polarization gradient cooling" *Appl. Phys. B* **65**, 563-568 (1997).
[12] D. Lucas, P. Horak and G. Grynberg, "Sisyphus cooling of rubidium atoms on the D2 (F=1 → F'=1) line: The role of the neighbouring transitions", *Eur. Phys. J D* **7**, 261-267 (1999).
[13] H.J. Metcalf and P. Van der Straten "Laser cooling and trapping" (Springer 1999).
[14] G. Nienhuis, P. van der Straten and S.-Q. Shang, "Operator description of laser cooling below the Doppler limit", *Phys. Rev.* A **44**, pp. 462-474 (1991).
[15] S. Pádua,"Laser cooling and slowing of rubidium" PhD thesis, SUNY, Stony Brook 1993. Unpublished.
[16], M.D. Plimmer, N. Castagna, G. Di Domenico, P. Thomann. A.V. Taichenachev and V.I. Yudin "2D laser collimation of a cold cesium beam induced by a transverse magnetic field", submitted to JETP Lett.
[17] B. Sheehy, S.-Q. Shang, P. van der Straten, S. Hatamian and H. Metcalf, "Magnetic-Field-Induced Laser cooling below the Doppler limit", *Phys. Rev. Lett.* **64**, pp. 858-861 (1990).
[18] A.V. Taichenachev , A.M. Tumaikin, V.I. Yudin and L. Hollberg, "Two dimensional sideband Raman cooling and state preparation in an optical lattice", *Phys. Rev. A* **63** 033402 (2001).
[19] C. Triché, P. Verkerk and G. Grynberg,"Blue-Sisyphus cooling in cesium gray molasses and antidot lattices" *Eur. Phys. J. D* **5**, pp. 225-228 (1999).
[20] P.J. Ungar, D.S. Weiss. E. Riis and S. Chu, "Optical molasses and multilevel atoms: theory" *J. Opt. Soc. Am. B* **6** pp. 2058-2071 (1989).
[21] C. Valentin, M.-C. Gagné, J. Yu and P. Pilllet "One-Dimension Sub-Doppler Molasses in the Presence of Static Magneic Field", *Europhys. Lett*. **19**, pp. 133-138 (1992).
[22] C. Valentin, "Manipulations non-cohérente et cohérente d'un jet atomique de césium par laser: Mélasses optiques dans un champ magnétique et transfert adiabatique de population", PhD thesis, Université de Paris-Sud 1994.Unpublished.
[23] P. van der Straten, S.-Q. Shang, B. Sheehy, H. Metcalf and G. Nienhuis "Laser cooling at low intensity in a strong magnetic field", *Phys. Rev. A* **47** pp. 4160-4175 (1993).
[24] D.S. Weiss, E. Riis, Y. Shevy, P.J. Ungar and S. Chu, "Optical molasses and multilevel atoms: experiment", *J. Opt. Soc. Am. B* **6** pp. 2072-2083 (1989).
[25] J. Yu, M.-C. Gagné, C. Valentin, S. Guellati, J. Djeema, R. Yuan and P. Pillet ,"Effects of a Static Magnetic Field on Sub-Doppler Molasses: Manipulation of Cesium Atoms", *IEEE Trans. Instrum. Meas*. **42**, pp. 109-114 (1993).

An Optical Clock with Ultracold Calcium Atoms

T. Nazarova, Ch. Lisdat*, C. Degenhardt, H. Stoehr, Ch. Grain, P.-E. Pottie, H. Schnatz, U. Sterr, F. Riehle

Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany,
Tel: ++49 531 592 4312, fax ++49 531 592 4305, uwe.sterr@ptb.de
*Institut für Quantenoptik, Universität Hannover, Welfengarten 1, 30176 Hannover, Germany

ABSTRACT

Optical frequency standards with laser-cooled neutral alkaline earth atoms can reach low relative uncertainties and high frequency stabilities due to the high numbers of atoms (10^7) and narrow clock transitions involved. In our Ca-frequency standard with $3\cdot10^7$ atoms Doppler-cooled to a temperature of 12 µK we have realized a relative uncertainty of $1.2\cdot10^{-14}$. It is limited mainly by the black body shift due to the Ca oven an the temporal phase chirp during the excitation pulses. Here we present a detailed investigation of systematic uncertainties in the calcium frequency standard. The fractional frequency stability of $2\cdot10^{-14}$ in one second mainly limited by Dick effect and detection laser noise was obtained. Different noise sources that contribute to the stability of the frequency standard and future improvements are discussed.

1. INTRODUCTION

Optical frequency standards based on narrow transitions with exceptionally high Q-factors ($Q = \nu/\Delta\nu$) have a potential to compete with the primary standards for frequency and time, the Cs-clock, since for a given linewidth $\Delta\nu$, a higher quality factor Q is obtained for optical frequencies.

Two different concepts of optical frequency standards are pursued using single ions and neutral atoms. Single-ion standards are expected to reach levels of fractional inaccuracy of 10^{-18} [1]. However, the fractional instability is limited to the lower 10^{-15} range due to a limited signal-to-noise ratio.

Neutral atom standards based on alkaline earth atoms use a high number of atoms (10^6 to 10^8) with the prospect to achieve extremely low instabilities. Relative inaccuracies in the 10^{-16} range are expected [2] for ballistically expanding clouds and below 10^{-17} for atoms trapped in an optical lattice [3]. Due to the high stability this accuracy level can be reached in less than one second.

In this paper we present the status and future prospects of the Ca-frequency standard. During our last frequency measurement in October 2003 we achieved a fractional uncertainty of $1.2\cdot10^{-14}$ and a stability of $2\cdot10^{-14}$ in one second. We present our new experimental setup that will allow us to reduce the uncertainty by one order of magnitude due to the elimination of the black body shift of the Ca-oven. By optimizing the loading rate, an improvement of the stability to the theoretical limit given by the quantum projection noise (QPN) of $6\cdot10^{-16}$ in one second with 10^8 atoms per second is expected.

2. PREPARATION OF THE ULTRACOLD ENSEMBLE

The thermal beam of Ca atoms from an oven at 900 K is first slowed by a Zeeman slower from a velocity of 600 m/s to a velocity of about 40 m/s (Fig. 1). Then the slow atoms are deflected by 10° and collimated with a 2D molasses into a magneto-optical trap (MOT). Compared with the previous setup the new one provides an increased atom flux into the MOT and avoids the black body radiation from the hot oven that was limiting the recent measurements [2].

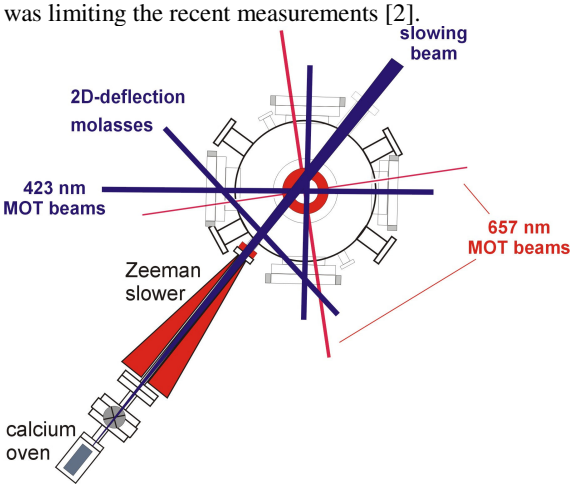


Fig. 1: Experimental setup for laser cooling of Calcium atoms

At the first cooling step Ca atoms are cooled on the allowed 423 nm transition $^1S_0$ – $^1P_1$ (Fig. 2) and captured in a MOT consisting of a magnetic quadrupole field with a gradient of 0.9 T/m and six independent laser beams with an intensity of 1.7 mW/cm^2 each and $1/e^2$ radius of 7.3 mm. In 10 ms about 10^7 atoms are collected in a cloud with a size of 0.3 mm x 0.3 mm x 0.5 mm at a temperature of 2 mK. That is close to the Doppler limit of 0.8 mK for this cooling transition. The lifetime of the first stage MOT is 1.2 s.

A second stage cooling is applied using the forbidden 657 nm transition $^1S_0$ – $^3P_1$. For efficient cooling the scattering rate is increased by quenching the upper metastable $^3P_1$ state by excitation of the 453 nm transition level to 4s4d $^1D_2$ state. With this method the atoms are further cooled down to 12 µK, corresponding to an rms velocity of 10 cm/s.

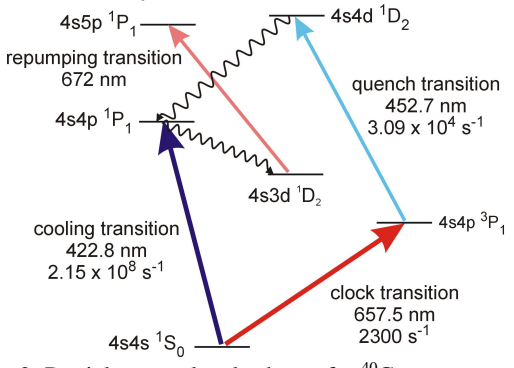


Fig. 2: Partial energy level scheme for ^{40}Ca

3. PROBING THE CLOCK TRANSITION

For precision spectroscopy of the calcium clock transition with a resolution close to the natural line width a highly stable spectroscopy laser and the utilization of atom interferometric methods are necessary. After the cooling sequence all fields are switched off to allow to probe the undisturbed clock transition $^1S_0$-$^3P_1$ at 657 nm. Since the atoms are nearly at rest the separation of the excitation region can be realized with laser pulses applied in a time sequence like in a 4-pulse Ramsey-Bordé atom interferometer [4] where two counterpropagating pairs of pulses act as beamsplitters. This method provides high spectral resolution and at the same time has the whole velocity distribution contributing to the signal. Both outputs of the atom interferometer are measured with high efficiency using an electron-shelving scheme.

3.1 Interrogation laser

The laser to interrogate the clock transition is an extended-cavity diode laser that is stabilized to a highly stable reference resonator described elsewhere [5]. To provide sufficient power for exciting the atoms two slave lasers are injection-locked to the ultrastabile master laser (Fig. 3).

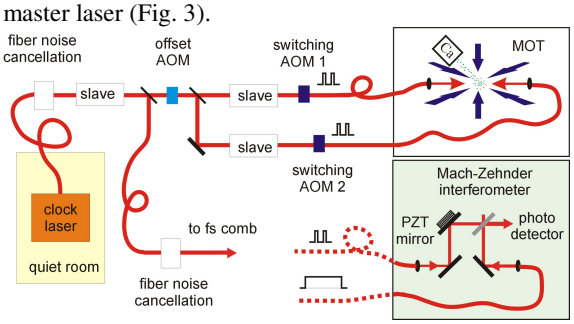


Fig. 3: Interrogation laser setup

The line width of the laser system of around one Hertz was determined by recording the beat between two identical but independent systems.

3.2 Influence of the temporal laser phase

The temporal phase during the spectroscopy pulses can be perturbed by thermal and transient effects in the laser system. For example a heating or electronic phase shifts in the acoustic-optical modulators (AOM) that cut the excitation pulses from the cw laser beam can lead to parasitic phase shifts. To quantify this effect the optical phase of the excitation pulses was measured with an optical homodyne Mach-Zehnder interferometer at different AOM center frequencies [5]. To rule out parasitic electronic phase shifts, e.g. of the photodiode, and to remove the background from a single beam, the measurements were performed on the positive and negative slope of the fringes and taking the difference between both signals.

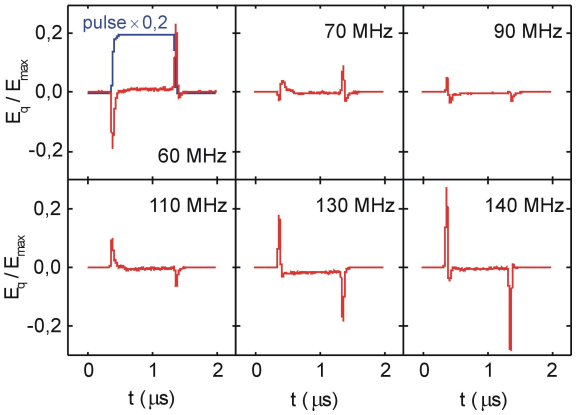


Fig. 4: The quadrature components of the electric field E_q of the excitation pulses normalized to the maximum field E_{max} for different center frequencies of the AOM.

We observed large phase excursions at the edges of the pulses (Fig. 4) which lead to frequency shifts up to 100 Hz depending on the AOM centre frequency. The measured frequency shifts as a function of the AOM center frequency is shown in Fig. 5. These measurements were compared with the frequency shifts calculated from the measured phase excursion by solving the optical Bloch equations for a 4-pulse atom interferometer.

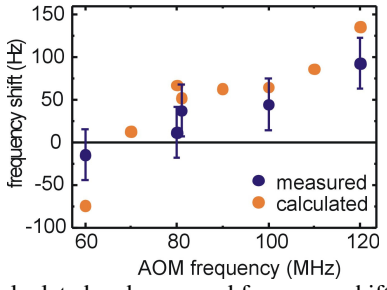


Fig. 5: Calculated and measured frequency shifts in the 4-pulse interferometer

The influence of the phase excursions in a 4-pulse atom interferometer was measured for different pulse separation times T (Fig. 6).

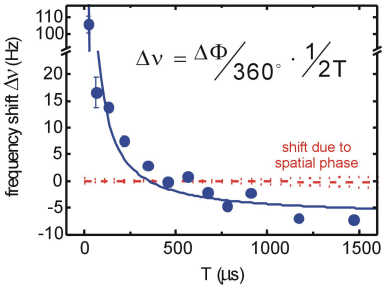


Fig. 6: Resolution dependent shift in a 4-pulse atom interferometer as a result of the influence of the phase excursions

The constant phase shift due to the phase excursions leads to a frequency shift proportional to $1/T$. Therefore this results can be used to correct the temporal phase during a frequency measurement.

4. FREQUENCY MEASUREMENT

The transition frequency of the calcium intercombination line at 657 nm was determined with respect to the primary standard on time and frequency, the caesium clock, using a self-referenced optical frequency comb generator. The measurement took place in October 2003.

Table 1 summarizes all contributions to the uncertainty budget of the calcium frequency standard.

The measured frequency of the $^1S_0$ (m = 0) to $^3P_1$ (m = 0) transition in ^{40}Ca is ν_{Ca} = (455 986 240 494 144 ± 5.3) Hz.

contribution	uncertainty u(Δν) / Hz	near future u(Δν) / Hz
black body shift due to Ca-oven (900 K)	3.9	0
laser phase		
- temporal	1.6	0.5
- spatial	1.0	0.6
asymmetry of line shape	0.2	0.05
quadr. Zeeman effect	0.1	0.1
drift of resonator	0.1	0.1
black body shift due to ambient temperature	0.07	0.07
cold collisions	0.06	0.06
quadr. Stark effect	0.02	0.02
sum in quadrature	4.3	0.8
uncertainty of Ca-standard	$1 \cdot 10^{-14}$	$2 \cdot 10^{-15}$
frequency measurement	3.1	
total uncertainty	5.3	

Table 1: Uncertainty budget of the frequency measurement in October 2003 (second column) and uncertainty expected with the new setup (third column)

The uncertainty is mainly limited by the black body shift due to the Ca-oven in the old setup an the temporal phase chirp during the excitation pulses. The measurement of the chirp-induced phase shifts will allow further reduction of the influence of the temporal phase, e.g. by operating the AOM at center frequency that minimized the shift. Since the black body shift from the oven will be avoided with the new setup using a Zeeman slower and a 2D molasses to deflect the atoms into the MOT a fractional uncertainty of the calcium frequency standard of $2 \cdot 10^{-15}$ can be reached in the next frequency measurement.

5. FREQUENCY STABILITY

To realize an optical clock the interrogation laser is locked to the atomic transition. Therefore any noise in the interrogation of the atomic line will lead to frequency fluctuations and degrades the stability of a clock.

The stability of the Ca-frequency standard during the last frequency measurement is shown in Fig. 7.

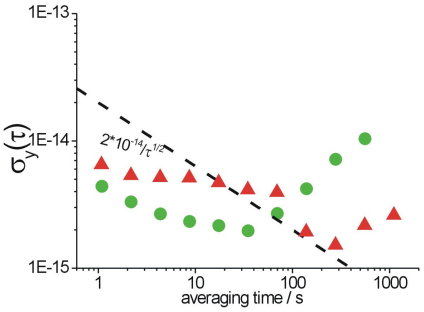


Fig. 7: Stability of the Ca-frequency standard. The stability of the Ca stabilized lasers with respect to a cavity-stabilized laser is shown by full triangles and the stability of two cavity-stabilized lasers is given by full circles. The calculated fractional instability of the Ca standard of $\sigma_y(\tau) = 2 \cdot 10^{-14} \tau^{1/2} / s^{1/2}$ is indicated by dashed line.

The stability for short times can be determined from a comparison between the Ca-stabilized laser and a laser stabilized to a second ultra-stable resonator. For times above the attack time of the lock to calcium (ca. 20 s) and below time where the drift of the resonator prevails (ca. 300 s) the stability of the Ca-frequency standard is close to $\sigma_y(\tau) = 2 \cdot 10^{-14} \tau^{1/2} / s^{1/2}$ (dashed line) in agreement with the calculated noise.

The relative frequency stability of an atomic frequency standard is described by the two-sample Allan deviation $\sigma_y(\tau)$ as a function of the averaging time τ [6]. For white noise in the detection, e.g. uncorrelated between single interrogation, it is given by [6-7]

$$\sigma_y(\tau) = \frac{1}{\kappa} \frac{1}{Q} \sqrt{\frac{T_c}{\tau}} \left(\frac{1}{S/N} \right) \qquad (1)$$

where $Q = 4 \cdot T \cdot \nu_{Ca}$ is the quality factor of the atomic line and κ is a constant depending on the shape of the resonance. For a Ramsey-Bordé interferometer this constant is equal κ = π [8]. T_c is the cycle time and *S/N* is the signal to noise ratio. Noise contributions to the Ca-frequency standard are listed in Table 2. The present stability of $2 \cdot 10^{-14}$ in one second is mainly limited by Dick effect and detection laser noise.

Noise source	σ_y(1s) (October 2003) T_c = 300 ms	σ_y(1s) (expected) T_c = 30 ms
QPN	$6 \cdot 10^{-16}$	$6 \cdot 10^{-16}$
Photon shot noise	$2 \cdot 10^{-16}$	$6 \cdot 10^{-17}$
Electronic noise	$7 \cdot 10^{-16}$	$2 \cdot 10^{-16}$
Detection Laser	$1 \cdot 10^{-14}$	$4 \cdot 10^{-17}$
Dick effect	$2 \cdot 10^{-14}$	$2 \cdot 10^{-17}$

Table 2: Noise and equivalent frequency stability contributions for the most recent measurement and expected improvements

5.1 Quantum projection noise

The quantum projection noise contribution to the frequency stability of an atomic frequency standard with N atoms is given by [9]

$$\sigma_{QPN,y}(\tau) = \frac{1}{\kappa}\frac{1}{Q}\frac{1}{C}\sqrt{\frac{T_c}{\tau}}\sqrt{\frac{1-p_e}{p_e N}} \qquad (2)$$

where C is the contrast of the Ramsey-Bordé fringes and p_e the excitation probability.
The quantum projection noise sets a theoretical limit to the relative frequency stability of $\sigma_{QPN,y}(1s) = 6 \cdot 10^{-16}$ in one second for $N = 3 \cdot 10^6$ atoms and cycle time of $T_c = 30$ ms.

5.3 Detection laser

The output of the interferometer is read out using a shelving detection method. Directly after the interferometry we shine in a 60 µs long laser pulse which is resonant with the $^1S_0$-$^1P_1$ transition at 423 nm with saturation parameter of $s \approx 1$. From the fluorescence induced by this pulse the number of atoms in the ground state can be determined. During the delay time of 1 ms the accelerated atoms leave the detection region and the atoms exited on the clock transition decay into the ground state. Therefore their number can be detected with the second detection pulse.
Since in the present detection method we use a running wave, the atoms are accelerated during the pulse. This leads to a change in the laser detuning due to the Doppler shift. This detuning results in a higher sensitivity to the frequency fluctuations of the detection laser. The contribution of the detection laser to the relative frequency stability is $\sigma_{\mathrm{det},y}(1s) = 1 \cdot 10^{-14}$ in one second. For improvement in this point we want to use a standing wave at higher saturation. Since in this method the atoms are at rest the detected signal is less sensitive to the frequency fluctuations. This will allow us in principle to reduce the contribution of the detection laser to $\sigma_{\mathrm{det},y}(1s) = 2 \cdot 10^{-17}$ in one second.

5.4 Dick effect

The Dick-Effect is an aliasing effect of the spectral density S_y of fractional frequency fluctuations of the interrogation laser at harmonics of the cycle frequency f_c due to the non-continuous interrogation of the atoms [10]. This aliasing effect leads to a frequency stability of the locked laser given by

$$\sigma^2{}_y(\tau) = \frac{1}{\tau}\sum_{k=1}^{\infty} S_y(kf_c)\left|\frac{g_k}{g_0}\right|^2 \qquad (3)$$

where g_y und g_k are the 0th and kth Fourier components of the sensitivity function *g*. The sensitivity function *g(t)* describes the change in the signal, e.g. the change in the excitation probability at the operation point of the frequency standard due to a small δ-like frequency fluctuation of the interrogation laser at time t. Fourier coefficients of the sensitivity function for a 4-pulse Ramsey-Bordé interferometer [11] for two different cycle times are shown in Fig. 8.

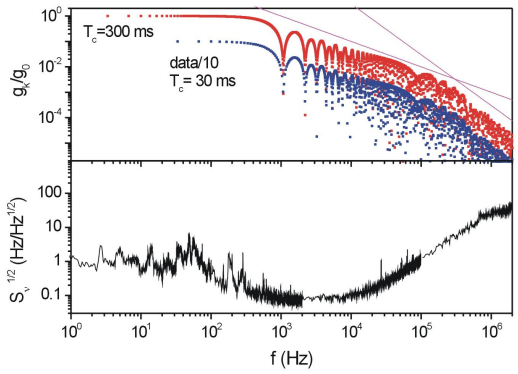


Fig. 8: Fourier coefficients of the sensitivity function (top) and spectral density of the frequency noise of the interrogation noise. For clarity, the data for a cycle time $T_c = 30$ ms are offset by a factor of 10.

To estimate the frequency stability due to the Dick effect the contributions of the measured noise spectrum at harmonics of the cycle frequency were weighted by the Fourier coefficients of the sensitivity function and summed up to $\sigma_y(\tau = 1s) = 2 \cdot 10^{-14}$.

Formula (3) shows the way to reduce the influence of the Dick effect. By increasing the cycle frequency from 3 Hz up to 33 Hz the contribution of the Dick effect can be reduced to $7 \cdot 10^{-15}$ in one second. This is due to less points in the low frequency range that have to be summed up.

To further minimize the Dick effect the frequency noise of the interrogation laser must be reduced especially in the low frequency range. At present the increase of the frequency noise of the interrogation laser towards lower frequencies is due to coupling of acoustic and seismic noise to the length of the resonator. With an improved setup we expect a reduction in the range above 10 Hz to $S_\nu^{1/2} \approx 0.1\mathrm{Hz}/\sqrt{\mathrm{Hz}}$ that would lead to an improved frequency stability of $2 \cdot 10^{-17}$ in one second.

6. CONCLUSION AND OUTLOOK

In conclusion we have shown the present status of the PTB Ca-frequency standard with ultracold calcium atoms. We have reached a relative total uncertainty of $1 \cdot 10^{-14}$ which now can be reduced to $2 \cdot 10^{-15}$ in the next measurement. Due to the new setup also an improvement stability approaching the quantum projection noise limit is expected.

7. ACKNOWLEDGEMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) and by the European Union through the program Cold Atoms and Ultraprecise Atomic Clocks (CAUAC).

REFERENCES

[1] A. A. Madej and J. E. Bernard, Single-ion optical frequency standards and measurements of their absolute optical frequency. In A. N. Luiten, editor, Frequency Measurement and Control, of Topics in Applied Physics, Springer, Berlin, Heidelberg, New York, vol. 79, pp. 153-194, 2001.

[2] U. Sterr, C. Degenhardt, H. Stoehr, C. Lisdat, H. Schnatz, J. Helmcke, F. Riehle, G. Wilpers, C. Oates, and L. Hollberg, The optical calcium frequency standards of PTB and NIST, @ C. R. Physique, vol. 5, pp. 845-855, 2004.

[3] H. Katori, M. Takamoto, V. G. Pal'chikov, V. D. Ovsiannikov, Ultrastable optical clock with neutral atoms in an engineered light shift trap, @ Phys. Rev. Lett., vol. 91, no 17, pp. 173005-1 – 173005-4, 2003.

[4] C. J. Bordé, Atomic interferometry with internal state labelling, @ Phys. Lett. A, vol. 140, pp. 10-12, 1989.

[5] C. Degenhardt, T. Nazarova, C. Lisdat, H. Stoehr, U. Sterr, and F. Riehle, Influence of chirped excitation pulses in an optical clock with ultra cold calcium atoms, @ IEEE Trans. Instrum. Meas., in print.

[6] D. W. Allan, Statistics of atomic frequency standards, @ Proc. IEEE, vol. 54, pp. 221-230, 1966.

[7] C. Audoin and B. Guinot, The Measurement of Time: Time, Frequency and the Atomic Clock, Cambridge University Press, Cambridge, New York, 2001.

[8] J. Vanier and C. Audoin, The Quantum Physics of Atomic Frequency Standards, Adam Hilger, Bristol and Philadelphia, 1989.

[9] W. M. Itano, J. C. Bergquist, J. J. Bollinger, J. M. Gilligan, D. J. Heinzen, F. L. Moore, M. G. Raizen, and D. J. Wineland, Quantum projection noise: Population fluctuations in two-level systems, @ Phys. Rev. A, vol. 47, pp. 3554-3570, 1993.

[10] G. Dick, Local oscillator induced instabilities in trapped ion frequency standards, in Proceedings of 19th Annu. Precise Time and Time Interval Conf., Redendo Beach, pp. 133-147, 1987.

[11] A. Quessada, R. P. Kovacich, I. Courtillot, A. Clairon, G. Santarelli, and P. Lemond, The Dick effect for an optical frequency standard, @ J. Opt. B, vol. 5, pp. S150-S154, 2003.

NEW CESIUM ATOMIC FOUNTAIN FREQUENCY STANDARD AT NICT

M. Kumagai, H. Ito, M. Kajita, M. Hosokawa
Atomic Frequency Standards Group
National Institute of Information and Communications Technology, JAPAN
mkumagai@nict.go.jp

ABSTRACT

National Institute of Information Communications Technology (NICT) has been developing Cesium atomic fountain primary frequency standards. For the construction of a new fountain, we have made the most use of the experiments and the results in the previous fountain. Also, we improved an optical system for the stable operation and installed a new microwave synthesizer. So far, we have achieved the frequency stability of $6\times10^{-13}/\tau^{1/2}$.

1. INTRODUCTION

NICT has been developing Cs atomic fountain frequency standards to aim at an operational primary frequency standard, whose frequency uncertainty is 1×10^{-15}, so that the standards can contribute to the accuracy of International Atomic Time (TAI) and Japan Standard Time (JST), which NICT have been maintained.

In 2002, we first succeeded to observe a Ramsey signal, which is narrower than 1Hz. Since then, we made several efforts to improve the SN ratio of the Ramsey signal. However, it did not become high enough to achieve the frequency stability at the order of $10^{-13}/\tau^{1/2}$[1,2,3]. In our first fountain, we found some points to be mended in its design (e.g. very long distance from the loading point to the microwave cavity, no selection cavity, low collection efficiency of the fluorescence and so on).

So, we have started to construct a new cesium atomic fountain with making the most use of the experience with the previous fountain. And, reviewing the whole system carefully, we improved the optical system and prepared stable microwave synthesizers.

This report presents the current status of the development of the new atomic fountain at NICT.

2. NEW ATOMIC FOUNTAIN

2.1 System Description

The design of the NICT new atomic fountain is shown in Fig.1. Our fountain consists of three parts; laser cooling area, microwave interaction region and detection zone. The detection zone is located between the laser cooling area and the microwave interaction region. A selection cavity, which pre-selects m_F=0 atoms, is installed just above the laser cooling area (about 8cm above the loading point). We use the rectangular cavity, whose resonance mode is TE_{102} for the state selection of the atoms. The loaded quality factor is about 100.

The laser cooling area consists of a trap chamber with several view ports and a pair of anti-helmholtz coils, which produce a magnetic gradient of 10mT/m. Cs atoms are trapped by the (0,0,1) cooling geometry. At present, magneto-optical trap (MOT) technique is used to load the atoms. In the future, we will try to capture the atoms without the magnetic gradient to reduce the collisional shift. An additional three

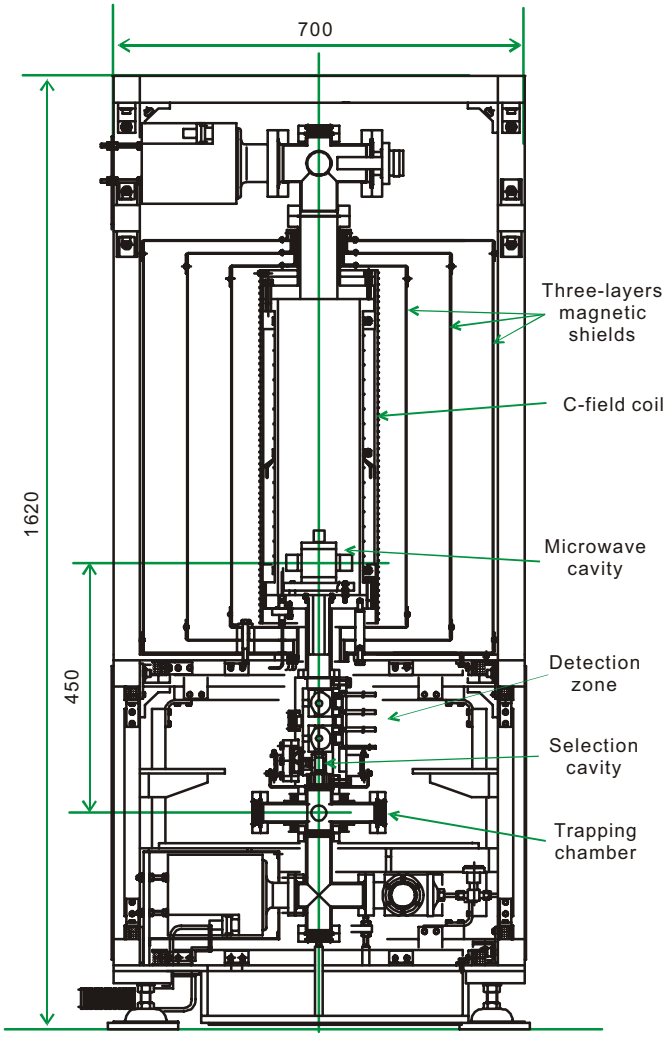


Fig. 1 NICT new atomic fountain

helmholtz coils are set on the flame, which supports the trap chamber to reduce the effect of the geomagnetism.
The microwave interaction region has a C-field coil and a cylindrical microwave cavity. The cavity has a TE_{011} resonance mode, whose loaded quality factor is about 15000. The whole of this region is surrounded by three-layers magnetic shield, which produce the shielding of order 10^4.

The detection chamber has three laser interaction ports where the detection beams are irradiated to the falling atoms. The highest and the lowest ports have the spherical mirrors each to increase the collection efficiency of the fluorescence. The fluorescence light collected by each mirrors is focused on a photo detector via a light pipe. Two photodiodes, which have a large active area ($1cm^2$), are used for the observation of the fluorescence from the falling atoms. The middle port is used for an optical pumping.

The trap chamber and the detection chamber are made of 316L stainless steel and the microwave interaction chamber is made of aluminium. Two microwave cavities and the spherical mirrors are made of oxygen-free copper.

Two graphite tubes are mounted between the laser cooling area and the detection zone, and between the detection zone and the microwave interaction region respectively as absorbers that prevent inflow of excessive Cs vapour into the microwave interaction region. The graphite tubes work as absorbers of the microwave leakage, too. By using two ion pumps and two NEG pumps, very high vacuum ($<2\times10^{-8}$ Pa) of the whole system has been achieved.

The height of total system is about 1.6m. The distance between the loading point and the microwave cavity is shortened to the length of 45cm, which is about 30cm closer than that of the previous fountain.

2.2 Improvements of the laser system

We use two extended-cavity diode lasers (ECDLs, New Focus TLB-6017) as master lasers. Both lasers are stabilized to Cesium D_2 absorption F=4-F'=5 line and F=3-F'=2 line independently by modulation transfer spectroscopy technique.

In each system, a portion of the laser beam is picked up from the main beams and that portion is divided into two beams; a pump beam and a probe beam. The pump is phase modulated by electro-optic modulator (EOM, New Focus 4001), which resonance frequency is 16MHz. The voltage controlled crystal oscillator (VCXO) is used as the source of 16MHz local reference. The modulated pump beam and the probe beams are counter-propagated and overlapped at the inside of the temperature controlled Cs vapour cell. The probe beam is detected by a Si PIN Photodiode (PD), and the AC component of the output is mixed with the local reference within a double balanced mixer. The error signal obtained by phase sensitive detection is amplified, integrated and fed back to the PZT and the current of the laser. In order to evaluate the frequency stability of the laser, we employ a femtosecond laser frequency comb system (Optical frequency synthesizer FC8003, Menlo system GmbH). The square root of the allan variance of the measured beat frequency is shown in Fig.2.

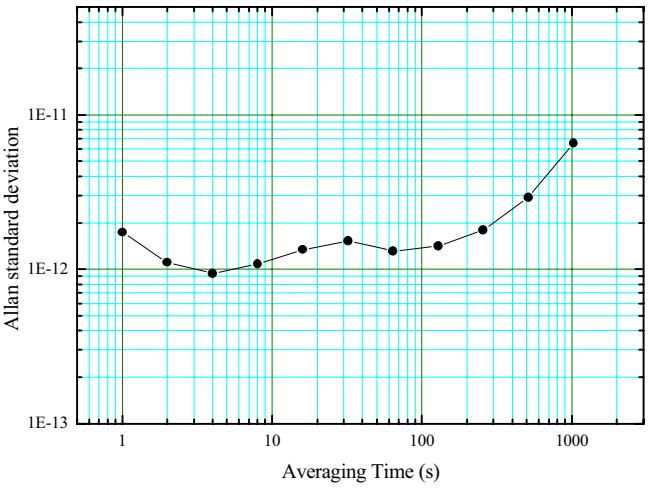


Fig. 2 Frequency stability of the master laser

The short-term stability of the laser is order of 10^{-12}. It means that the laser is stabilized within less than 1kHz, which is narrow enough for the primary frequency standards based on the Cs transition. By this method, both lasers are stabilized strongly to the Cs transition lines. We seldom have a laser dropout during the operation.

Two 150mW laser diodes (SDL5422) are injection-locked to the highly stabilized ECDL to amplify the laser power for the laser cooling. The frequency and the power of each laser beam are controlled with acousto-optics modulator (AOM). A direct digital synthesizer (DDS) makes synchronization of all AOMs. Each laser beam is delivered to the vacuum chamber through polarization maintaining fiber (PMF). Output port of each fiber has a collimation lens and a quarter-wave plate, which generate the circular-polarized beams. The diameters of the laser beams are 12mm for the vertical beams and 25mm for the horizontal beams with the power density of at least $10mW/cm^2$. A portion of the high power output is used for the detection.

The other ECDL stabilized to the F=3-F'=2 transition is used for repump. By using AOMs, the laser frequency is tuned to the F=3-F'=4 transition to pump the F=3 atoms back to the F=4 state. The repump beam is used for the laser cooling and the detection.

2.3 New room for atomic fountains

Last year, our T&F groups moved to a new building, where Japan Standard Time has been maintained under the high security. The building has several experimental rooms for the atomic standards. One of them is used for the development of our atomic fountains. The room itself is surrounded by one-layer magnetic shield, which reduces the effect of the geomagnetism to about 10%. The room is well temperature controlled within a range of ±0.1°C and the floor is isolated from the other ground by an air-suspension.

Thanks to an excellent environment and the achieved high laser stability mentioned above, we had no lasers dropout for one month. Although Tokyo had some

earthquakes of intensity 3 recently, our optical setup for the atomic fountain withstood those earthquakes. Without an intentional shock, the system will be operational continuously for more than one month. It becomes possible to measure the long-term frequency stability.

2.4 Microwave source

As the source of 9.2GHz microwave interrogation, we prepare a synthesizer of Spectra Dynamics Incorporation (SDI CS-1). All VCXOs and DDS are phase-locked to the 5MHz of the hydrogen maser, which is linked to UTC(NICT). Controlling DDS by PC, we tune the microwave frequency with a resolution finer than 1μHz. The output power of the synthesizer is controlled by a power servo with an uncertainty of 0.05dB.

In order to evaluate the frequency stability of the SDI synthesizer, we prepared two other 9.2GHz synthesizers; the synthesizer for space clock[4], which NIST have developed, and the synthesizer for our optically pumped primary frequency standard (NICT-O1)[5]. By measuring the stability of beat note between each pair of three synthesizers by dual-mixer time difference (DMTD) method [6], we evaluated the frequency stability. As a result, the frequency stability of the SDI synthesizer is $3\times10^{-13}/\tau$, which is almost the same as that of our hydrogen masers. So far, our microwave synthesizer does not set limit to the frequency stability of the atomic fountain.

3. FOUNTAIN OPERATION

3.1 Data acquisition

The operation cycle of the atomic fountain is as follows. Cs atoms are captured in a MOT (400ms), and then cooled in the optical molasses (20ms). The atoms are launched upward by (0,0,1) moving molasses. In order to avoid the inefficiency by a saturation effect of the horizontal beams at the launching process, the horizontal lasers are turned off at the initial stage of the moving molasses [3]. After the atoms reach to the equilibrium velocity, the horizontal beams are turned on again to prevent the horizontal dispersion. After giving the launch velocity to the atoms, the launched atoms are post-cooled to about 2μK by the polarization gradient cooling (PGC). All the laser beams are turned off by AOM and extinguished by mechanical shutters. In our system, in order to observe the Ramsey fringe narrower than 1Hz, the atoms must be launched with the initial velocity over 3.84m/s. At present, the atoms are launched with the initial velocity of 4m/s to the apogee of 82cm above the loading point.

Only m_F=0 components of the launched atoms (F=4) are pumped to the (F=3,m_F=0) state by the π–pulse microwave excitation through the selection cavity located at 8cm above the loading point. Atoms remaining in the F=4 state are blasted away by the radiation pressure of the downward beams in the vertical direction. The (F=3,m_F=0) atoms continue the ballistic flight into the microwave cavity. Twice π/2-pulse microwave interactions at going upward and downward move the atom in the (F=3, mF=0) state to the (F=4, mF=0) state with following the transition probability due to the microwave frequency.

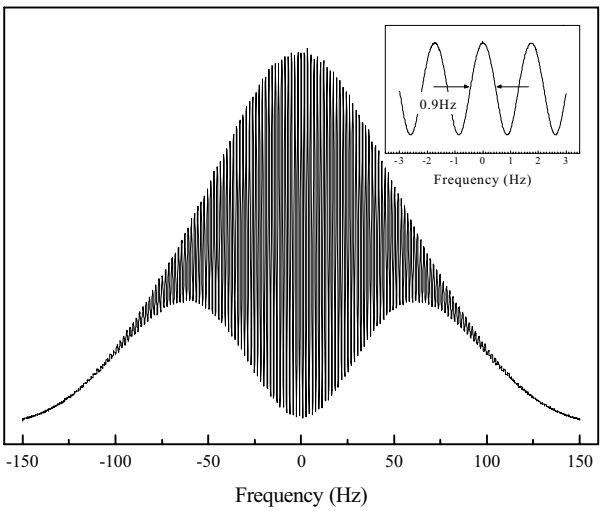


Fig. 3 Ramsey fringe (Centre 9192631770Hz)

In order to avoid the cycle-to-cycle fluctuation of the number of the launched atoms, the signal is normalized by measuring both the number of the F=4 atoms and the number of the F=3 atoms separately at the detection zone. At the upper detection port, the number of the F=4 atoms (N_4) is detected, and after detection the F=4 atoms are pushed away by the radiation pressure. The number of the F=3 atoms (N_3) is counted by the optical pumping to the F=4 atoms at the middle port and measuring their number at the lower detection port. The fraction $P=N_4/(N_4+N_3)$ indicates the normalized transition probability.

With the initial velocity of 4m/s, the atoms are launched to the apogee of 38cm above the microwave cavity, which gives the drift time of 560ms. As a result, the Ramsey fringe of less than 0.9Hz is observed. The observed Ramsey fringe is shown in Fig.3.

3.2 Frequency Stability

In order to lock the microwave frequency to the narrow atomic resonance, we adopt the frequency modulation locking method, in which the microwave is toggled between f_0-Δν/2 and f_0+Δν/2, where f_0 is the microwave frequency, and Δν is the linewidth of the Ramsey fringe, at each cycle. By comparing the signal intensities at the two frequencies, the centre frequency f_0 is detuned to make the intensity difference zero. The amount of the detuning is proportional to the intensity difference at the two frequencies. By this method, the microwave is stabilized to the narrow Cs atomic line. Fig.4 shows the allan variance of our new atomic fountain vs the hydrogen maser. At the first attempt, we have already achieved the frequency stability of $6\times10^{-13}/\tau^{1/2}$ much better than that of the previous one.

4. SUMMARY AND FUTURE PLAN

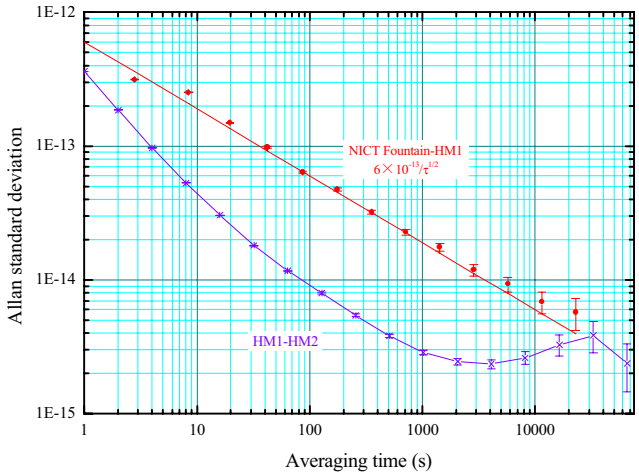


Fig.4 Allan deviation of our atomic fountain -hydrogen maser

NICT have been developing the new atomic fountain frequency standard. We constructed the new vacuum vessel, which is more compact in the size and sophisticated in the design. Furthermore, we improved the optical system and installed the stable microwave synthesizers. These improvements have produced the higher stability. So far, the frequency stability is presently $6\times10^{-13}/\tau^{1/2}$, which is better than that of the previous fountain.

We have started to estimate the several frequency shifts and their uncertainties and evaluate the reliability as the primary frequency standard. At the estimation of the frequency shifts in the atomic fountain, the instability of the reference makes the accuracy evaluation difficult. In order to overcome the difficulty, we have been developing the stable reference system based on the hydrogen maser and the Cs clocks. Furthermore, it is planned to construct the next atomic fountain and compare the data of two atomic fountains to avoid the instability of the reference.

REFERENCES

[1] K. Fukuda, H. Ito, M. Kumagai, M. Kajita, A. Hasegawa, M. Hosokawa, T. Morikawa, “Cesium Fountain Development at CRL” Proceedings of ATF2000, pp. 217-223, (2000)

[2] M. Kumagai, H. Ito, K. Fukuda, M. Kajita, M. Hosokawa and T. Morikawa: “Development of Atomic Fountain Frequency Standard at CRL” Proc. of 17th EFTF, pp. 120-122, (2003)

[3] M. Kumagai, H. Ito, K. Fukuda, M. Kajita, M. Hosokawa and T. Morikawa: “Recent Progress on Cs Atomc Fountain at NICT” Proc. of 18th EFTF, (2004)

[4] A. Sen Gupta, D. Popovic and F. L. Walls: “Cs Frequency Sthesis: A New Approach” Proc. of 1999 Joint Meeting EFTF-IEEEE IFCS, pp.615-619, (2000)

[5] A. Hasegawa, K. Fukuda, M. Kajita, H. Ito, M. Kumagai, M. Hosokawa, N. Kotake and T. Morikawa: “Accuracy Evaluation of Optical Pumped Primary Frequency Standard CRL-O1” Metrologia, pp.257-263, (2004)

[6] L.Sojdr, J. Cermak, and G. Brida, “Comparison of high-precision frequency-stability measurement systems” Proc. of 17th EFTF, pp.317-325

HIGH FREQUENCIES MONOLITHIC QUARTZ FILTERS

Bertrand d'ALBARET, Bernard WOLCOFF, Séverine MICHEL*, Philippe CAPPELLE*

AR-Electronique 15 Rue d'Estienne d'Orves 78500 Sartrouville FRANCE
* AR-Electronique Espace Valentin Nord BP 3113 – 25047 Besançon FRANCE

ABSTRACT

The place given to the components by the designers of radiocommunications and telecommunications systems is smaller and smaller and furthermore the middle frequency of the receiver is higher and higher.

In this paper, we present new filters of very small size with a big selectivity using high frequencies monolithic resonators operating in fundamental mode.

INTRODUCTION

The designers of radiocommunications and telecommunications systems are always demanding higher middle frequency for their receiver.

Furthermore they want to use components smaller and smaller.

The selectivity of the receiver is most of the time given by the selectivity of the input filter.

In addition to the selectivity, most of the receivers must transmit numerical signal, so the filter must be linear in the larger part of the bandwidth. That will say that the monolithic cells must be always wider.

AR-Electronique has studied filters with very small size and is now able to provide piezoelectric filters with an enlarged useful bandwidth by the use of monolithic resonators operating in fundamental mode.

AR-Electronique has developped special very small masks devoted to these high frequencies and has optimized a special automatic machine in order to produce easily the high frequency monolithics.

State of the art

A piezoelectric resonator is commonly represented under it's simplified scheme fig 1, where L, C, R are motional elements i.e. they transcribe mechanical vibrations of the crystal into the electrical domain.
C0 is a physical capacitor related to the inter-electrode dielectric. The resonator has 2 noteworthy frequencies :

Fs= 1/(2*PI*sqr(L*C))
Fp=1/(2*PI*sqr(L*C*C0/(C+C0)))
(The transmission and impedance response curves are represented in fig 2A and 2B.

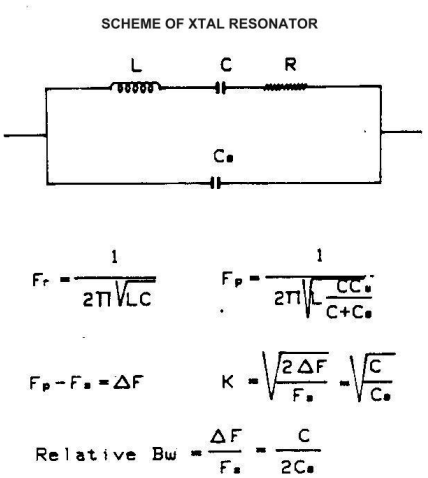


Fig 1

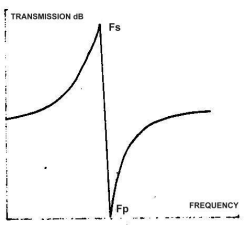


Fig 2A

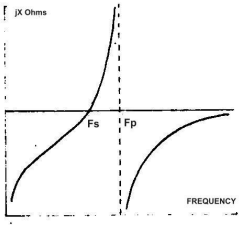


Fig 2B

One usually calls the useful bandwidth of a resonator : ΔF, the difference between the resonant and the anti-resonant frequencies.
An approximation leads to an expression for the usefull relative banwidth : ΔF/Fs ≈ C/(2*C0)
In fact ΔF has the order of magnitude of the maximum bandwidth able to be produced in a filter and this value is close to the maximum tuning range of a VCXO.
When used in overtone mode the maximum relative banwidth is divided by the square of the overtone rank.
One easily notices that wideband piezoelectric devices are limited by overtone rank
So it is very interesting to use resonators in fundamental mode.
There is 2 technologies to obtain blanks with thickness smaller that 30 µm, that will say 50 MHz.
The ion etching technology and the chemical etching technology.

In the first case, the wafer is eroded in it's center by an Argon ion beam. A glass mask cover the circumference of the blanks.
In the second case, the wafer is eroded in it's center by a chemical solution. A metallic layer protect the wafer and only the center part is eroded.
The principle is given on fig 3.

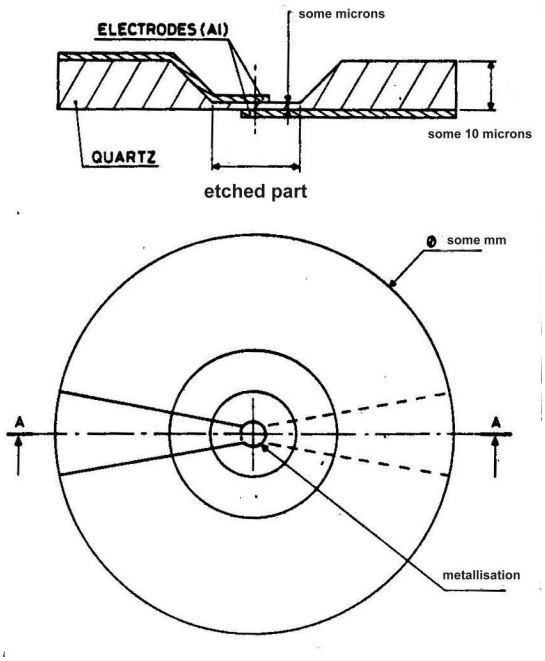


Fig 3

On this exemple, the metallisation is simple. That gives a discret resonator i.e. 1 pole only for a filter.

MONOLITHIC CELLS FOR FILTERS

In order to reduce the dimensions of the filter and to use less blanks for the same synthesis, we have studied monolithic cells, i.e. 2 resonators coupled on the same blanks.
The working principle with the equivalent scheme of an elementary cell is given fig 4.

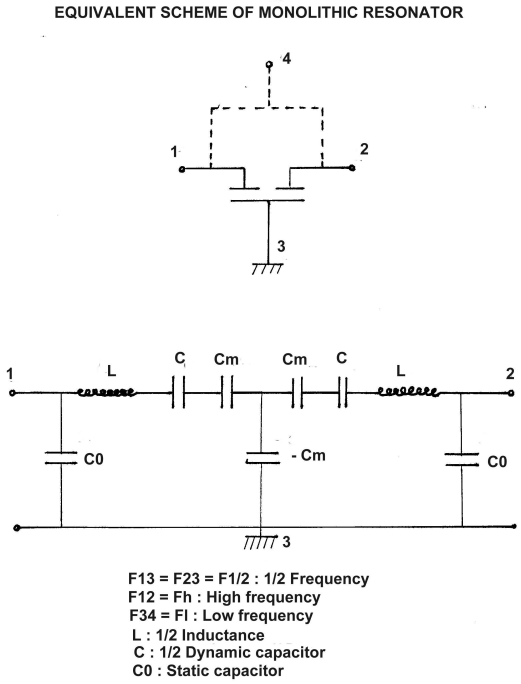


Fig 4

A special calculation program of the monolithic cell bandwidth has been developped, i.e. the difference between the 2 frequencies high and low in step with the dimension of the metallization mask and the metallization load.
The optimum dimension of the metallization mask is that of the trapped energy of the discret quartz.
We have developped very small metallization masks in order to obtain the very small motionnal inductance given by the calculation at such high frequencies.
An exemple of mask is given fig 5.

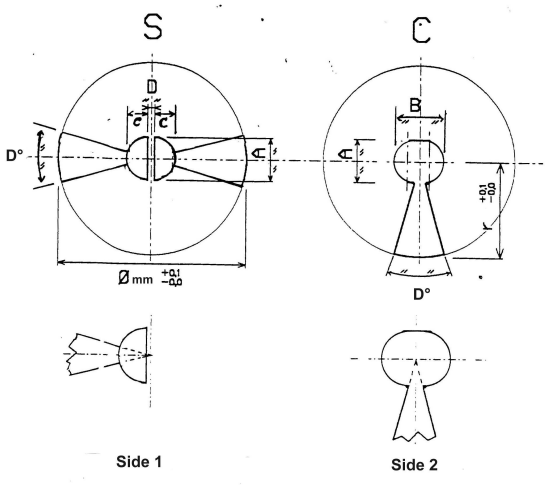


Fig 5

We propose now to analyze some filters produced with this technology.

70 MHz Type 60 kHz - 4 poles Filter

This filter has to stand in a case of 14 x 25 mm with height of 8 mm.

The specification implies 4 poles :

F0 = 70 MHz
1 dB Bw ≥ ± 30 kHz
30 dB Abw ≤ ± 100 kHz
60 dB Abw ≤ ± 200 kHz

We have made the calculation with 2 monolithic cells. The bandwidth of the cells is 56 kHz.

The theoretical curves are given fig 6.

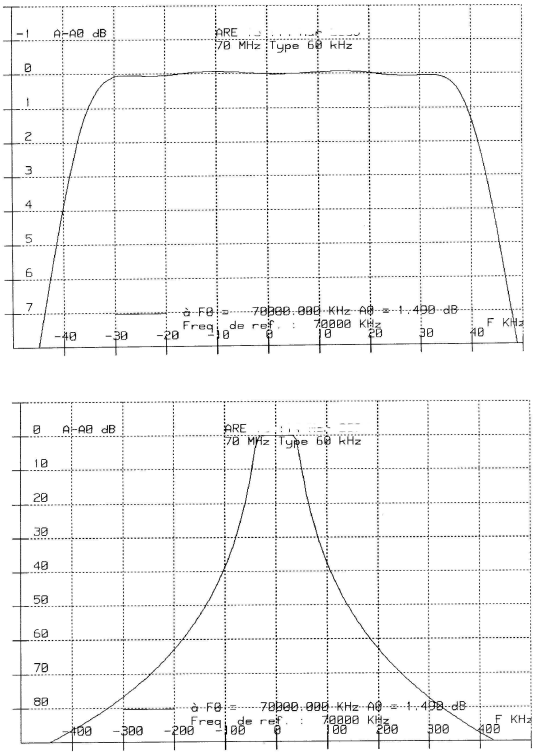


Fig 6

The calculation of the metallization load with the choosen mask is given fig 7.
A view of the metallized cell is given fig 8.
The electrical scheme of the filter is given fig 9.

The real responses curves are given fig 10, and we can see that they are very close of the theoretical..

Numéro du cache : 220001 Type : S23-C23

Self 1/2 = 3.711 mH Self Basse = 1.855 mH Self Haute = 7.422 mH
Energie piégée : Partiel 1 Charge = 793 kHz

Fréquence Centrale du Filtre : 70 MHz
Self 1/2 = 3.711 mH Self Basse = 1.855 mH Self Haute = 7.422 mH

Energie piégée : Partiel 1 Charge = 793 kHz

* CHARGE (kHz) *	LARGEUR SUR XX' (kHz) *	LARGEUR SUR ZZ' (kHz) *
* (Fu - F0) *	(Fh-Fb) *	(Fh-Fb) *
* 1500 *	41.6 *	14.8 *
* 1450 *	43.4 *	15.7 *
* 1400 *	45.3 *	16.7 *
* 1350 *	47.3 *	17.7 *
* 1300 *	49.5 *	18.8 *
* 1250 *	51.8 *	20.0 *
* 1200 *	54.2 *	21.3 *
* 1150 *	56.8 *	22.8 *
* 1100 *	59.6 *	24.3 *
* 1050 *	62.5 *	26.0 *
* 1000 *	65.7 *	27.8 *
* 950 *	69.0 *	29.9 *
* 900 *	72.6 *	32.1 *
* 850 *	76.5 *	34.5 *
* 800 *	80.7 *	37.2 *
* 750 *	85.1 *	40.2 *
* 700 *	90.0 *	43.4 *
* 650 *	95.2 *	47.1 *
* 600 *	100.8 *	51.2 *
* 550 *	106.9 *	55.8 *
* 500 *	113.6 *	61.1 *
* 450 *	120.8 *	67.0 *
* 400 *	128.6 *	73.8 *
* 350 *	137.1 *	81.5 *
* 300 *	146.3 *	90.5 *
* 250 *	156.1 *	101.0 *
* 200 *	166.3 *	113.3 *
* 150 *	175.8 *	127.6 *
* 100 *	182.1 *	143.4 *
* 50 *	173.9 *	155.4 *
* 0 *	0.0 *	0.0 *

Fig 7

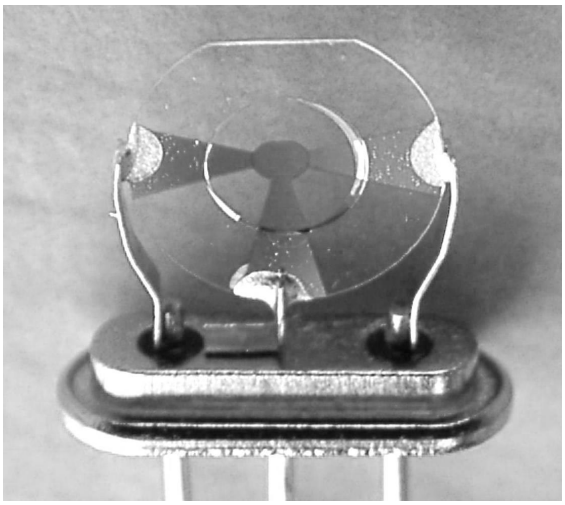


Fig 8

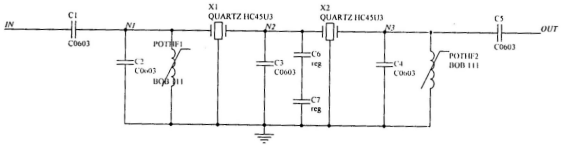


Fig 9

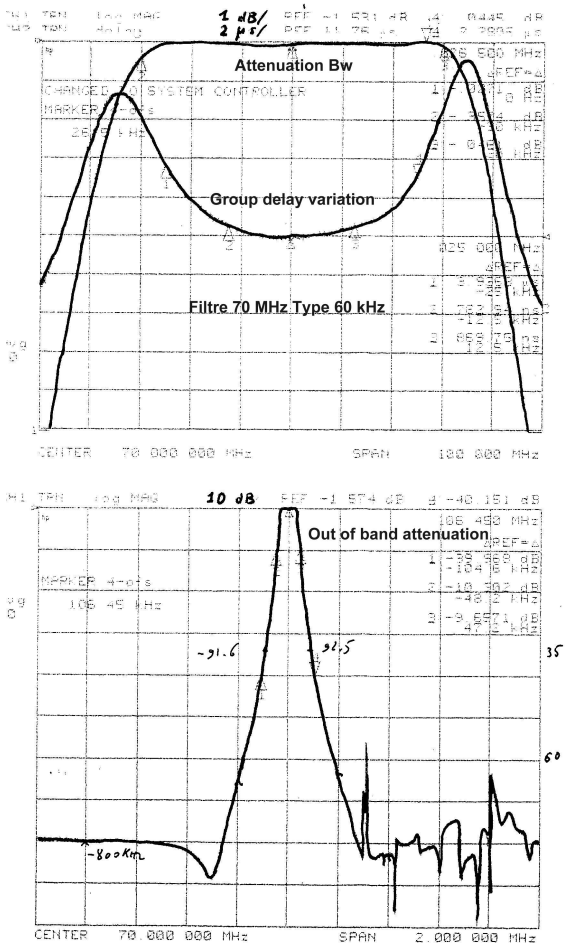


Fig 10

75 MHz Type 36 kHz - 6 poles Filter

The dimensions of the case are 16 x 25 x 10 mm.
The specification implies 6 poles :

F0 = 75 MHz
3 dB Bw ≥ ± 18 kHz
50 dB Abw ≤ ± 75 kHz
IL ≤ 3 dB

Furthermore there is a group delay condition :
ΔTpg ≤ 10 µs over ± 18 kHz

The theoretical synthesis give us a scheme with 3 monolithic resonators. We obtain 2 different bandwidths of cells : 61 and 27 kHz.
The scheme of the filter is given fig 11.

We observe on the real responses curves fig 12 that they are absolutely in concordance with the theory and that we obtain an attenuation of 90 dB including the spurious responses.

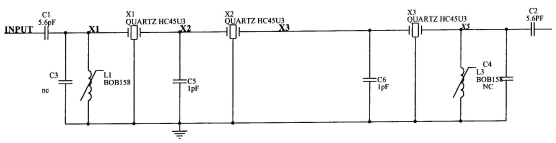


Fig 11

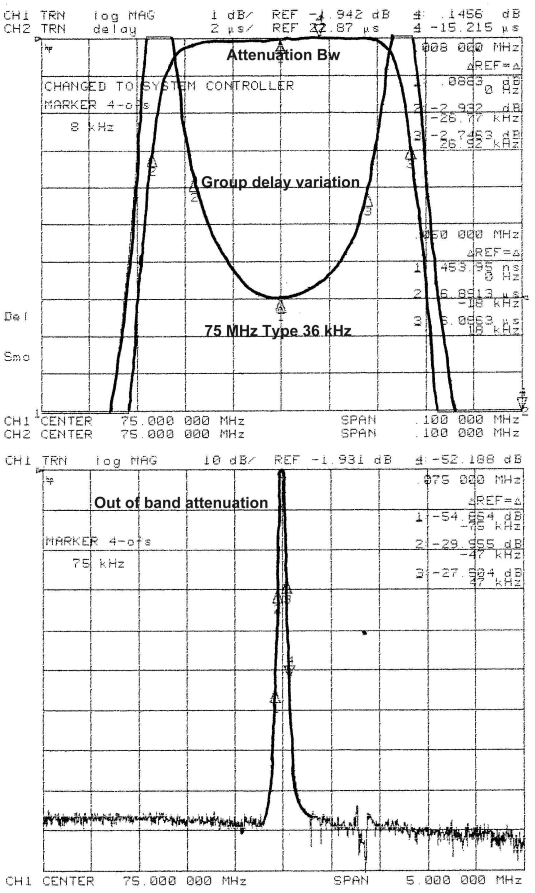


Fig 12

70.425 MHz Type 34 kHz - 8 poles Filter

In the same small case, the specification of this filter requests 8 poles :

F0 = 70.425 MHz
3 dB Bw ≥ ± 17 kHz
60 dB Abw ≤ ± 50 kHz
80 dB Abw ≤ ± 70 kHz
Out of band ≥ 90 dB

We have calculate a 8 poles filter designed with 4 monolithic resonators. We obtain 2 bandwidths of cells : 30 and 22 kHz.
The scheme is given fig 13 and the real responses curves fig 14. We can see that the out of band attenuation is remarkable : ≥ 100 dB.

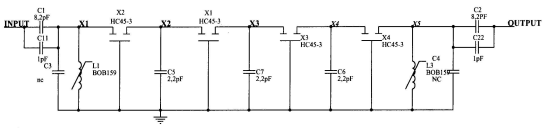


Fig 13

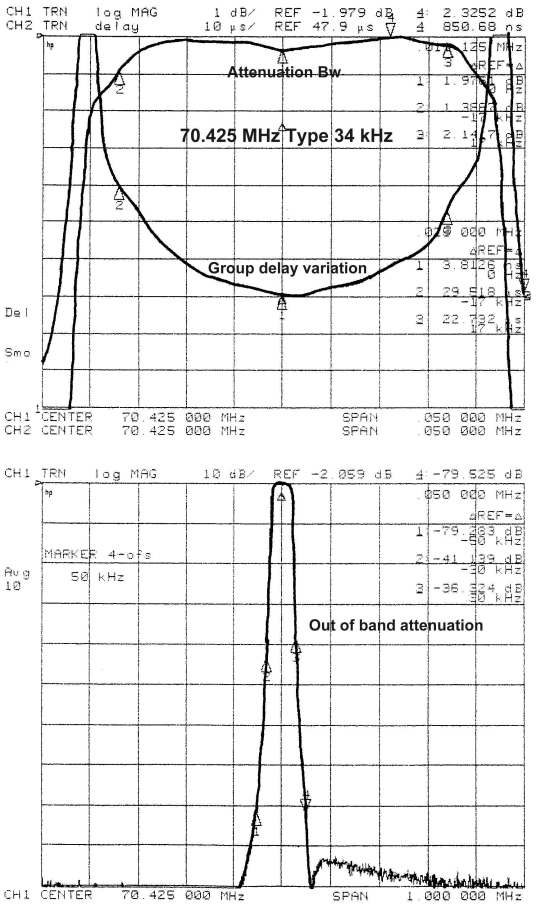


Fig 14

125 MHz Type 50 kHz - 4 pole phase linear filter

The specification of this filter is very difficult because due to the case size and price, it is only a 4 poles filter with an important out of band attenuation and furthermore a constant group delay in all the total passbandwidth.

F0 = 125 MHz
3 dB Bw ≥ ± 25 kHz
30 dB Abw ≤ ± 100 kHz
70 dB Abw ≤ ± 300 kHz
IL ≤ 3 dB
Group delay variation ≤ 2 µs over ± 25 kHz

The filter is calculated with a special synthesis type semi-linear. That give us 2 different cells with bandwidths of 45 and 72 kHz.

We have calculated 2 different masks in order to achieve the 2 different bandwidths and to minimize the level of the spurious in different bands of frequencies.
You can see the 2 different curves in transmission of the 2 cells fig 15.

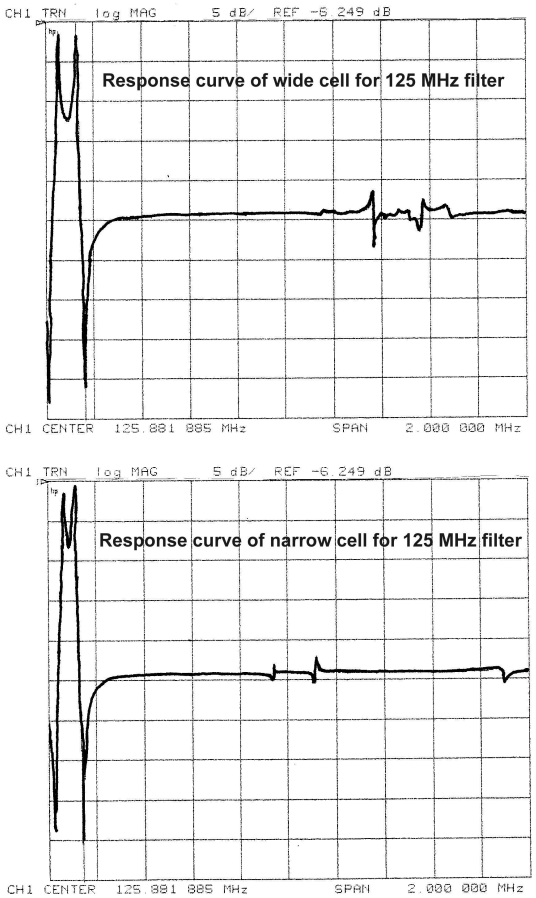


Fig 15

The real response curves of the filter are mentioned fig 16.
We can see the remarkable low variation of the group delay, and the very good attenuation ≥ 70 dB including the spurious.

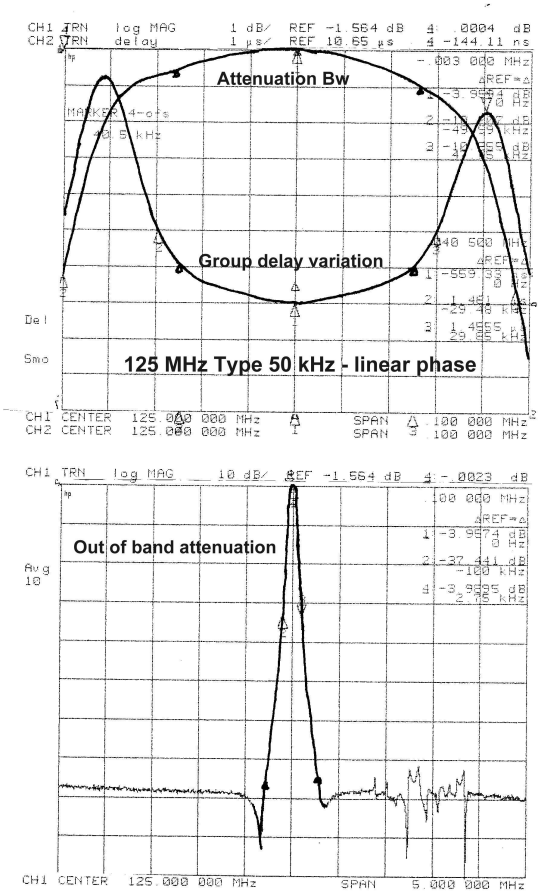


Fig 16

CONCLUSION

Currently AR-Electronique has objectively demonstrated its capacity to produce high frequency monolithic filters. Up to now the frequency range is between 50 and 130 MHz with a relative bandwidth up to 0.1 %.

Shortly, we intend to improve our performance in higher frequencies but for now we are limited by the mechanical tolerances of the masks.

With this technology we are able to offer to the telecommunication engineers a good solution in performance and size for their filtering problems in U.H.F.

AR-Electronique is now involved in industrial manufacturing process of filters using monolithic resonators in fundamental mode.

COLD ATOM ONBOARD GRAVIMETER ANALOGIES WITH ATOMIC CLOCKS

Fabien Lienhart, Salah Boussen, Alexandre Bresson, François Deyzac
ONERA, 29 avenue de la Division Leclerc, 92400 Châtillon
fabien.lienhart@onera.fr

We present an experimental set-up of a cold atom gravimeter after discussing briefly its principles. Notable features are a 780 nm, <10 kHz linewidth, fibered laser source developed using a telecom laser and a PPLN waveguide crystal, and a Titanium vacuum chamber built using an original method for sealing windows. Analogies with atomic clocks are finally explored, leading to a proposition for a new type of gravimeter accelerometer on one hand, and justifying the interest of our work on onboard constraints for atomic clocks on the other.

1. INTRODUCTION

Taking advantage of extensive experience in the Time and Frequency domain, in particular for environmental behaviour studies of USOs and atomic clocks, of the leadership in ultra precise space accelerometer development and of the solid competence in laser sources and OPO in particular, a cold atom onboard gravimeter is being developed in our laboratory with the aim of space, military and industrial applications.

The cold atom gravimeter operation was first demonstrated by Kasevich & Chu [1] in 1992. This instrument rapidly reached performances comparable to the best conventional absolute gravimeters [3], but was still a heavy and non transportable cold atom laboratory instrument. Building a cold atom absolute gravimeter for onboard use would allow for many applications such as oil prospecting, global positioning, earth tidal studies or even for military applications of navigation.

We developed laser sources and mechanical devices which satisfy onboard conditions in order to ultimately build an onboard cold atom gravimeter. As will be recalled in this paper, cold atom sensors are very similar to atomic clocks. Therefore most of the results presented here might also be useful for any onboard cold atom experiment such as atomic fountain clocks.

2. PRINCIPLE OF THE COLD ATOM GRAVIMETER

The cold atom gravimeter consists of an atomic fountain. In our case rubidium is used, and the atomic cloud is not launched vertically but is dropped in free fall along the vertical axis. During the free fall three laser pulses, tuned to the hyperfine transition of the ground state, will split ($\pi/2$ pulse), reflect (π pulse), and recombine ($\pi/2$ pulse) the atoms. The main difference from atomic clocks is that the transition produces a strong velocity recoil. This is due to the use of Raman transitions between the hyperfine sublevels of the ground state instead of radio-frequency transitions. Raman transitions are achieved by two laser beams ~1GHz red-detuned from the respective F=1->F' and F=2->F' transitions of ^{87}Rb. They lead to a recoil velocity of 12 mm/s (compared to 46 nm/s for radio-frequency transitions), and thus to a maximised spatial separation between the two interferometric paths which is responsible for the dependence of the phase difference on vertical acceleration. A matter-wave interferometer is thus achieved with the phase difference between the two paths proportional to vertical acceleration by $\Delta\varphi = k_{eff}.g.T^2$, where g is the local gravity field, T is the time between two laser pulses (~50 ms), and k_{eff} is the sum of the wave vectors of the two lasers.

After the three pulses, the phase difference is derived from the measurement of the normalised population of a hyperfine level by fluorescence.

Note that, contrary to atomic clocks, the interferometer phase difference does not depend on the frequency of the hyperfine splitting. This is due to the fact that the time T between the different pulses is kept constant, and thus both paths experience the same times in the different sublevels, which cancels out the hyperfine transition contribution to the interferometer.

3. EXPERIMENTAL SET-UP

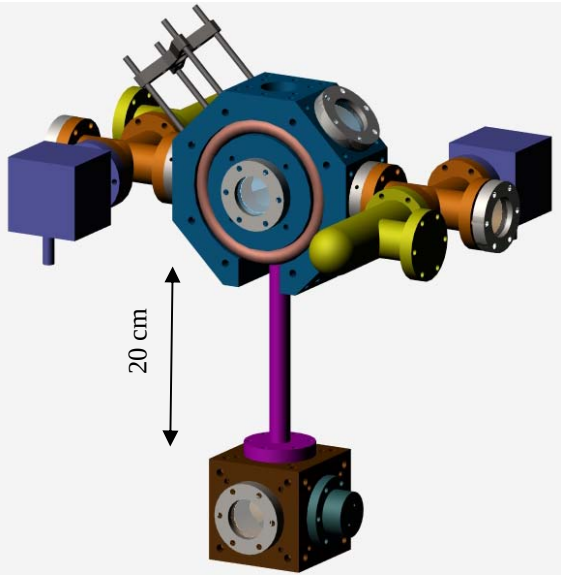


Figure 1 - Schematic of the vacuum chamber. The MOT is prepared in the top part, atoms are then dropped. The 3 laser pulses occur during the free fall in the central tube, and detection is achieved in the bottom. Two 3L/s ionic pumps surround the top part.

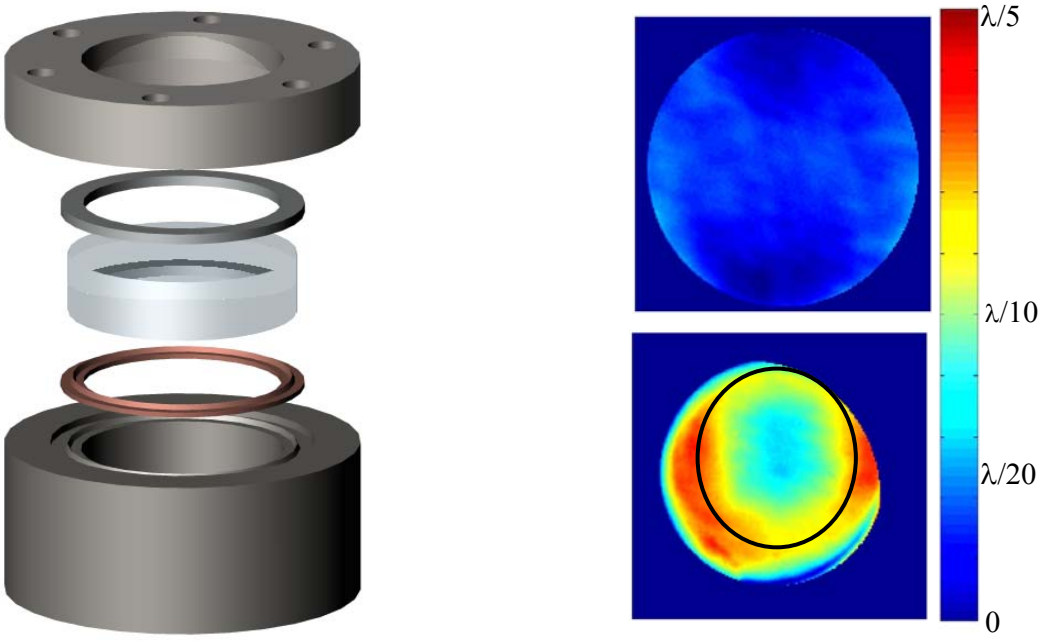


Figure 2 - Left: the 15-mm thick glass is sandwiched between the machined knife-edge copper gasket and the aluminium buffer gasket. The top flange is screwed to the bottom chamber – **Right**: measurement of the wave front deformation before (top) and after (bottom) sealing. The black circle represents the 25mm diameter area maintaining λ/10 quality.

Our set-up, sketched in Figure 1, is similar to usual Rubidium atomic fountains, except that the atomic cloud will not be launched. The vacuum chamber is mainly made of Titanium in order to minimise disturbing magnetic fields. The magneto-optical trap is loaded with Rb dispensers.

An original method was used to seal the windows. Windows are usually brazed to metal using Kovar which is a magnetic material. The method developed by Noble and Kasevich [2] consists in machining a knife edge on a copper gasket which is placed between the glass and the metal of the chamber. On top of the glass an aluminium gasket serves as a buffer, and a flange is used to fix the glass to the chamber. This method provides two advantages: the materials used are non magnetic, and the deformation of the window is minimised. Wavefront deformation measurement was performed on 40 mm diameter λ/10 quality windows before and after sealing (see Figure 2). The window kept a λ/10 quality on a 25 mm diameter area after sealing.

A laser source on D2 transition of Rb (780 nm) is also being developed. As our instrument is designed for onboard use, a laser without an extended cavity was preferred in order to be more robust to vibrations. It uses a 1560 nm fibber laser which is amplified through an Erbium Doped Fibber Amplifier (EDFA) and finally sent into a PPLN waveguide crystal. This non-linear crystal allows for second harmonic generation with conversion efficiencies as high as 12%/W/cm², thanks to its periodic poling and to the confining of light along the waveguide. Up to 70 mW of 780 nm light come out of the crystal with 1W of pump, but power rapidly decreases due to photorefractive effects. This problem comes from the low temperature at which the crystal is operated (60°C) and should be circumvented using a crystal with a different poling period.

Previously, free space crystals have been tested both in single pass and in a butterfly cavity configuration. This leads to much lower efficiencies (0.5%/W/cm^2 for PPLN) but a higher pump power could be reached without photorefractive effects. In the cavity configuration up to 1.1W of 780 nm light was produced with 2W of pump.

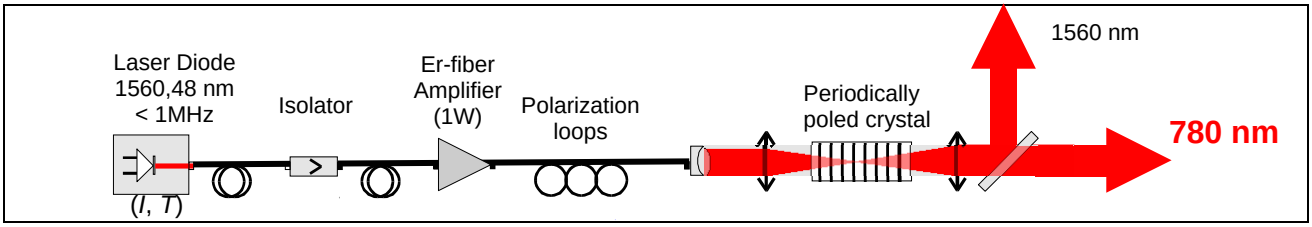


Figure 3 - 780nm laser source in the free-space crystal configuration. Laser Diode can easily be changed for a fibber laser with narrower line width (< 1kHz).

4. ANALOGIES WITH ATOMIC CLOCKS

Cold atom gravimeters and atomic clocks are in many ways similar: they both are atomic fountains, they both rely on the same principle of a matter wave interferometer using hyperfine sublevels of the ground state, and they both use a normalised detection technique to derive the phase difference. The two main differences are that atomic clocks use radio-frequency transitions, whereas our gravimeter uses Raman transitions in order to maximise the spatial separation of the two paths. For this purpose, atomic clocks use a $\pi/2$-π-$\pi/2$ interferometric sequence in order to be mainly sensitive to the hyperfine frequency, whereas our gravimeter uses a $\pi/2$-π-$\pi/2$ sequence which cancels this former sensitivity and allows for a vertical acceleration sensitivity.

The comparison of atomic clocks with cold atom inertial sensors should be profitable for both types of experiments. Inspired by the servo-control scheme of a usual cesium atomic clock, a design for a new type of accelerometer is now proposed. The main idea in atomic clocks is to drive a radiofrequency transition by a very stable oscillator with very good short term properties but which drifts in time, and to remove the drift by probing the hyperfine transition of Cs, which remains constant in time. If now the stable oscillator is replaced by an electrostatic accelerometer (with very good short term properties, but drifts in time) and the atomic clock by a cold atom accelerometer (which provides absolute measurement), this should result in a high accuracy absolute accelerometer (see Figure 4).

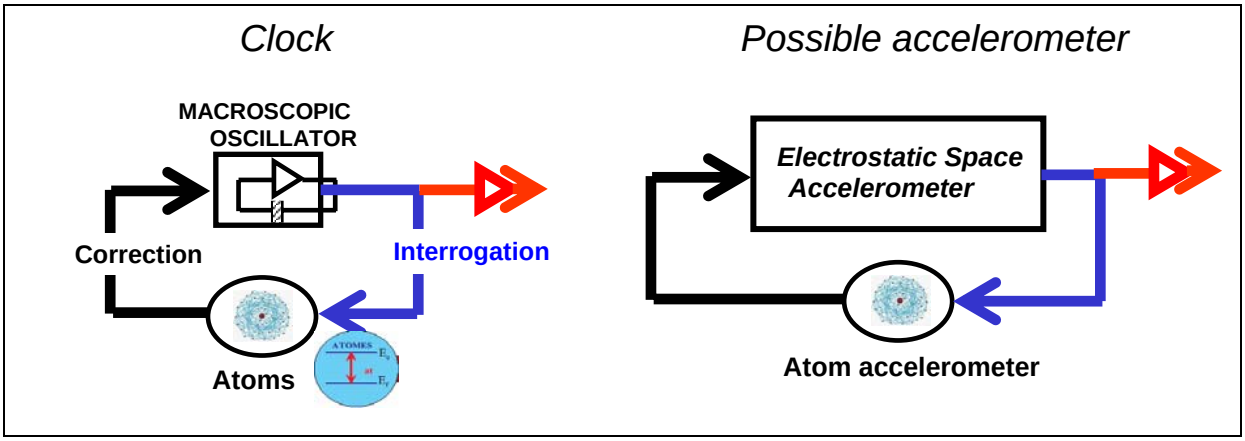


Figure 4 - Left: retroaction scheme of a conventional clock. A crystal oscillator provides the RF stability allowing for the interrogation of the atoms, which in return is used to correct long-term drift of the oscillations of the crystal. **Right**: analogous scheme for a high performance accelerometer. An electrostatic space accelerometer measures acceleration with high precision, allowing stabilization of the mirror of the atom accelerometers, which in return corrects the drift of the electrostatic device.

5. CONCLUSION

ONERA is developing a cold atom absolute gravimeter for onboard use. The fruitful comparison of atomic clocks with cold atom inertial sensors inspires a new possible accelerometer scheme with high accuracy and minimised drift. Reciprocally, our progress regarding onboard instrumentation, such as a 780 nm fibber laser source may benefit any onboard cold atom experiment.

This research is funded by the following french institutions : Conseil Régional d'Ile-de-France, Conseil Général d'Essone (SOFIA Program), National French Space agency (CNES) and Direction Générale de l'Armement (DGA).

6. REFERENCES

[1] M. Kasevich, S. Chu, Measurement of the Gravitational Acceleration of an Atom with a Light-Pulse Atom Interferometer, @ Applied Physics B: Lasers and Optics, vol 54, pp. 321-332, 1992.

[2] A. Noble, M. Kasevich, UHV optical window seal to conflat knife edge, @ Review of Scientific Instruments, vol 65, pp. 3042-3043, 1992.

[3] A. Peters, K.Y. Chung, S. Chu, High-precision gravity measurements using atom interferometry, @ Metrologia, vol 38, pp. 25, 2001.

INVESTIGATION OF SPECTRAL DEVICES SOURCE FOR MINIATURIZED RUBIDIUM GAS CELL ATOMIC FREQUENCY STANDARDS

A. B. Bassevich, O.V. Beresovskaya, G.M. Smirnova

"The Russian Institute of Radionavigation and Time" PJSC, 2 Rastrelli sq., St.-Petersburg 191124, RUSSIA
E-mail: niiss@rirt.ru

1. ABSTRACT

As a result of studies provided, it was shown that, after accelerated tests, all spectral devices maintained their working ability and had actual lifetime up to 100,000 and more hours. This allows to use such devices in small-size rubidium frequency standards for space-based systems.
As regards the spectral lamps, the filing parameters were refined and it was shown that, for supporting the above lifetime, a minimum Rb^{87} quantity should be of 0.017 mg and minimum pressure of the buffer gas (Kr) – of 2.1 Torr. The relative change in the optical pumping coefficient for tested cell-filters was about 2 % and did not exceed the measurement error. The relative change in the effective frequency value for gas cells was of $(9\pm4)\times10^{-10}$ over 100,000 hours, and a change in signal and in temperature coefficient after tests is also within limits of the measurement error for these parameters.

2. INTRODUCTION

The atomic gas-cell frequency standards with optical pumping are widely adopted in land, marine, aviation, and space equipment due to their compactness, small weight and relatively low price.
The rubidium frequency standards (RFS) belongs to the category of passive atomic frequency standards in which an atomic gas-cell frequency discriminator is used as frequency-defining component. The RFS's accuracy and operation parameters are to a large extent defined by a main unit of the atomic frequency standard – rubidium atomic discriminator (physics package).
The physics package contains a resonance emission source – an electrodelles spectral lamp in which the electrodelles RF discharge is being excited; a gas cell in which the radio optical resonance is observed; a cavity; an optical indication system; ovens, and magnetic shields. The lifetime of spectral devices affects significantly the operation parameters of physics package and device as a whole.
According to the current requirements to the space-based systems, the lifetime of RFS should exceed 10 years. In this case, it is of great importance to consider problems related to limiting the lifetime of spectral devices when their weight and dimensions become significantly lower.
This paper presents the results of studying the lifetime of spectral devices.

3. STUDYING A LIFETIME OF THE ELECTRODELLES SPECTRAL LAMPS

Problems related to the longevity of spectral lamps were considered in a number of works [1, 2, 3, 4], but, when the requirement to the lifetime achieved 100,000 hours, the additional studies were needed. In our work the problems are studied related to the longevity of RFS's electrodelles spectral lamp operating under «strained» conditions of RF discharge and high temperature
The main requirements to spectral lamps as a light source are value and stability of the emission intensity, affecting the rubidium frequency standard (RFS) stability through the light shifts mechanism. The frequency shifts should be reduced to the minimum and be less than 1×10^{-13}.
Such lamp parameters are defined their filling and theirs behavior over time, i.e. quantity of the working substance and buffer gas pressure. In this case, the reasons limiting the electrodelles spectral lamp lifetime are as follows:

- gradual decrease in the buffer gas pressure in the course of operation (this processes is known as a «hardening» process [4]);
- dependence of lamp parameters on the content of the introduced metal, its activity relative to filing and lamp balloon's material.

These processes cause frequency shifts and change in lamp's emission intensity, deterioration of its ignition and, as a consequence, restrict its lifetime. There are procedure for contact less checking these parameters [5,6] allowing to carry out studies on determining quantitative values of foregoing parameters being sufficient for supporting requirements to light sources for RFS.
In addition, emission intensity, its stability, operation reliability of spectral lamp with rubidium vapour depends also on the metal quantity (Rb^{87}). The metal quantity in electrodelles spectral lamp needed for the long-term operation was determined in [5] as early as 70s. This allowed to introduce operative checking against this parameter and, as a result, to eliminate electrodelles spectral lamp failures caused by the metal's burn-out. Later in 80s these results were confirmed by experimental studies in the framework of developing RFS for GPS in the USA [7]. All this data

were, however, obtained for spectral lamps of different design compared to current ones which operated under another conditions. At present, due to the decrease in rubidium frequency standard dimensions, the spectral lamp dimensions decreased also. Therefore one should conduct studies for selecting optimal filling for up-to-date electrodelles spectral lamp. In addition, these studies should substantiate and support the possibility to improve a lifetime for spectral lamp of this design up to 100,000 hours needed for space-based RFS.

The electrodelles spectral lamps being studied were spherical glass balloons of 8 mm diameter filled by a mixture of Rb^{87} and Kr. The lamps were placed into RF-oscillators of elevated power for accelerating aging processes [1].

In the course of the test, the buffer gas pressure was measured at specified time intervals, t_i, using radiospectroscopic technique [6]. Measurement results are presented in Fig. 1 and Table 1.

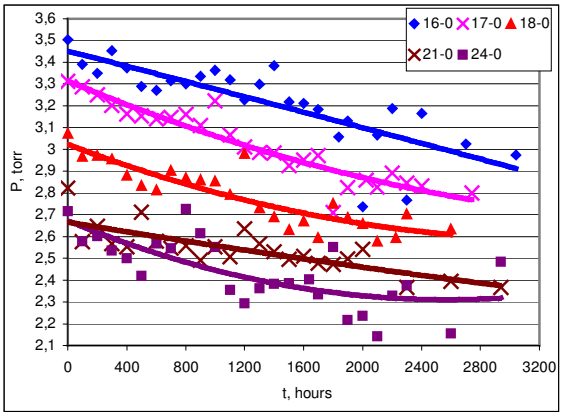


Fig. 1 Change in buffer gas pressure in the course of aging

Table 1.

Change of buffer gas pressure before and after test

№ lamp	Training time (t_{tr}), hours	Buffer gas pressure before test, torr	Buffer gas pressure after test, torr
16-0	3.040	3.50±0.01	2.97±0.01
17-0	2.740	3.31±0.01	2.80±0.01
18-0	2.600	3.07±0.01	2.63±0.01
21-0	2.940	2.82±0.01	2.37±0.01
24-0	2.940	2.72±0.01	2.49±0.01

The tests were being conducted under conditions of spectral lamp's accelerated aging where training time (t_{tr}) and routine operating time (T_{ro}) were related through following expression [4]:

$$T_{ro} = t_{tr} \times K_{acc}, \qquad (1)$$

where K_{acc} – acceleration coefficient for aging processes in spectral lamps during training which depends on the discharge power in the lamp:

$$K_{acc} = \left[\frac{P_{tr}}{P_{ro}}\right]^{1,72}, \qquad (2)$$

where P_{tr} and P_{ro} – discharge power in the lamp during training in the course of accelerated test and when operating in the routine mode, respectively. These powers were determined using the calorimetric technique [2].

Proceeding from the obtained data, the rates were calculated for decreasing a gas pressure within studies spectral lamps, $\Delta P/\Delta t$. The results are presented in Table 2.

Table 2.

Change velocity of gas pressure in the spectral lamps

№ lamp	$\Delta P/\Delta t$, 10^{-6}, torr/hour
16-0	6.8
17-0	7.4
18-0	6.6
21-0	6.1
24-0	3.1
$\overline{\Delta P/\Delta t}$	6±2

The minimum buffer gas pressure which should remain after ending of electrodelles spectral lamp operation time and at which the stable discharge is maintained should be no less than 1.3 torr [6]. In this case, for providing a lifetime of 100,000 hours, the initial gas pressure in the lamp should be no less than 2.1 torr.

In the context of the experiment being conducted, a quantity of the metal was measured in the lamp before and after the test using optical-thermic technique [5]. Table 3 presents the data on changing in the quantity of metal (M) in electrodelles spectral lamp in the course of aging obtained during the accelerated test.

Table 3.

Change of quantity of metal before and after test

№ lamp	t_{tr}, hour	M before test, mg	M after test, mg
16-0	3.040	0.016±0.002	0.016±0.002
17-0	2.740	0.017±0.002	0.016±0.002
21-0	2.940	0.017±0.002	0.014±0.002
24-0	2.940	0.014±0.002	0.015±0.002

The rate of metal quantity changing was defined using a following formula:

$$v = \frac{M_n - M_k}{T_{ro}}, \qquad (3)$$

where M_n и M_k – quantity of metal measured before and after test and T_{ro} – time interval over which this change would take place in lamp operating in the routine mode (1). The mane rate of the metal quantity change was of $(4\pm3)\times10^{-8}$ mg/hour.

The minimum metal quantity which should remain when electrodelles spectral lamp works out its operating time should be no less than 0.01 mg [4]. In this case, when a lamps' lifetime of 100,000 hours should be provided, one can determine a lower limit for a metal quantity which ESL should have:

$$M_{\min} = 0.01 + v \times 100{,}000 = 0.017 \text{ mg}$$

4. STUDYING A LIFETIME OF THE FILTER-CELL

The short-term stability of RFS is defined by a signal-to-noise ratio which in the limit of the shot noise depends on the spectral composition of the emission traversing through a cavity with the gas cell (under condition of optimal parameters of microwave field and detection system).

In a perfect case, for achieving the maximum signal-to-noise ratio, the emission falling on the gas cell should consist of two HF components of D_1 and D_2 – lines of the Rb^{87} resonance doublet. The low-frequency components should totally be eliminated.

For supporting a pumping process in RFS, one utilizes the optical filtering of hyperfine components using a filter-cell in a form of a glass cylindrical balloon with transparent end planes filled by Rb^{85} vapour and inert gas, argon is most frequently used. The gas is introduced to provide a better coincidence of Rb^{85} and Rb^{87} atoms' hyperfine components.

The main parameter of the filter-cell is a coefficient of optical pumping efficiency, ($\boldsymbol{k_{opt.pump}}$), characterizing the increase in the optical resonance signal when installing a filter-cell to the RFS's optical path. $\boldsymbol{k_{opt.pump}}$ is defined as a ratio of the second harmonic amplitude of the error signal measured at the output of the matching amplifier with and without the filter-cell.

As for this work, the small-size filter-cell of 14 mm diameter and 7 mm thickness were studied which were filled by Rb^{85} vapour and Ar at a pressure of 150 torr.

The filter-cells were being placed to the oven with the elevated temperature, and $\boldsymbol{k_{opt.pump}}$ was periodically measured. The data on the behavior of coefficient of optical pumping efficiency over time is given in Table 4.

Table 4.
Coefficient of optical pumping efficiency over time

$\boldsymbol{k_{opt.pump}}$				
T_{ro}, *hour* *№ FC*	0	59683	104,557	Relative change over T_{ro}, %
18-4	30.7	25.9	30.0	2.3
26-4	35.3	28.2	34.7	1.7
20-3	25.0	26.8	23.6	5.6
34-3	28.8	30.5	30.6	5.9

The coefficient of scaling the time, K_{acc}, for the operation in the normal mode was calculated using relations given in [4].

The relative error at which the optical pumping coefficient was measured was of 6.4%. Therefore, proceeding from the data of Table 4, one can conclude that the relative change in the optical pumping coefficient in filter-cell does not exceed the measurement error after their artificial aging for 100,000 hour operation.

5. STUDYING A LIFETIME OF THE GAS-CELL

One the main physics package component is a gas-cell which defines to a large extent its accuracy parameters. The main gas-cell parameter is a frequency of the Rb^{87} atomic transition which change during the RFS lifetime is a subject of studies.

The main reasons which cause a drift of the gas-cell frequency and, as a consequence, limit the gas-cell lifetime are following processes being relatively complicated to analyze and all the more to imitate:

- instability of the gas mixture in the gas-cell during its operation due to the permeation of the atmospheric gases to the cell and the leakage of the buffer mixture components from the cell;
- diffusion of gases dissolved in the cell's glass envelope to the cell volume as well as adsorption and desorption of the gas mixture at its inside surface;
- interaction of the rubidium with residual gases in the cell which appearance in the cell is inevitable during its manufacture;
- "migration" of the rubidium along the inside gas-cell surface and change in the absorption/desorption rate due to the change in the inside surface area.

A reason for frequency deviation in the gas-cell is also the sensitivity to temperature variations. Such a type of the instability is characterized by the temperature frequency coefficient (TC). The change in this parameter with gas-cell aging contribute also to the frequency drift and limits its lifetime. The procedures for checking these parameters are described in [4] thus allowing to conduct studies on determining their quantitative values.

The gas-cell being studied were spherical glass balloons of 12 mm diameter filled by Rb^{87} and Ar and Ne buffer gas mixture. The cells were placed to the oven at the elevated temperature to accelerate aging processes [4]. In the course of their running in, at specified time intervals, the frequency was measured using a laboratory breadboard of the rubidium frequency standard.

Taking into account aging processes' acceleration at the elevated running-in temperature, one can calculate a relative change of frequency as:

$$V_F = \frac{F_{before_test} - F_{after_test}}{F_{before_test}}.$$

These calculations are presented in Table 5.

Table 5.

Relative change of gas-cell frequency

№ gas-cell	V_F
22-0*	5.346×10^{-10}
1-90	5.555×10^{-10}
4-3	5.130×10^{-10}
25-95	4.573×10^{-10}
10-97	1.000×10^{-09}
6-98	2.707×10^{-10}
10-98	1.929×10^{-09}
11-98	2.535×10^{-09}
13-98	1.094×10^{-09}
14-98	1.380×10^{-09}
18-98	9.303×10^{-10}
24-98	3.101×10^{-09}
14-99	8.892×10^{-10}
15-99	1.020×10^{-09}
Mean Drift over 140,057 hours	$(1.2\pm0.5)\times10^{-09}$

* - reference cell not exposed to the elevated temperature

In such a way, a relative change in effective frequency value over 100,000 hours was $(9\pm4)\times10^{-10}$, thus corresponding to the monthly frequency drift of $(6\pm3)\times10^{-12}$.

The gas-cell TC measurement conducted in the course of these studies has shown that the relative change was within the limits of the measurement error and its value was of ~ $(3 - 5)\times10^{-11}$ 1/°C.

REFERENCES

[1] G. Smirnova, V. Khutorschicov. Accelerated test of spectral lamp and their longevity.// Thes. Repts. 6th All-Union Conf. "Metrology in Radioelectronics". Moscow, VNIIFTRI, 1984, – p. 201 (in Russian).

[2] G. Smirnova. Candidate's dissertation. – Leningrad, LNIRTI, 1985 (in Russian).

[3] Very long lifetime expectancy design / Temex time & frequency – http://www.temex-components.com

[4] A.S. Kazantsev, V.I. Khutorshicov. Practical spectroscopy of high-frequency discharge. Plenum Press: New Jorke and London, 1998.

[5] T. Nikolayeva, V. Semenov, V. Khutorschicov. On measurement of the metal quantity in spectral devices.// Problems in Radioelectronics. Ser. General Technique. Issue 4. 1974, – p. 164 – 167 (in Russian).

[6] V. Khutorschicov. On measurement of pressure in electrodelles spectal lamps.// Problems in Radioelectronics. Ser. General Technique. Issue 2. 1975, – p. 110 – 114 (in Russian).

[7] Cook R.A., Frueholz R.P. An improved Rb consumption model for discharge lamps used in Rb frequency standard // Proceedings of 42nd Annual. Control. Sympos., 1998, pp. 525 – 531.

AN INFORMATION AND REFERENCE SYSTEM FOR ACOUSTOELECTRONIC DEVICE DEVELOPMENT

O.L. Balysheva, L.P. Konovalova, S.V. Kulakov, Yu.G. Smirnov

St.-Petersburg State University for Aerospace Instrumentation,
67, B. Morskaya, St.-Petersburg, 190000,
E-mail: weconf@home.ru

ABSTRACT

One of the first and main stages in acoustoelectronic device design is the selection of substrate material. For the applications in acoustoelectronic devices, two most important parameters are the velocity of acoustic waves and the electromechanical coupling coefficient. In addition, the following material parameters should be known for the acoustoelectronic device design: temperature coefficients (delay and velocity), dielectric constant, attenuation constants, anisotropy parameter, the power flow angle, material density, electrode pair capacity.

New requirements to the performance of telecommunication equipment have led to the fact that new materials are to be used. New data on various materials and their characteristics are constantly published in contemporary scientific literature. This information requires systematization, storage, and convenient access to it should be provided. A database, a part of which is the information on the main parameters of piezoelectric materials, is under development at the Department of physical and theoretical fundamentals of radioelectronics of the St.-Petersburg State University for Aerospace Instrumentation.

1. INTRODUCTION

Acoustoelectronic devices occupy the leading position among ad-hoc electronic devices. They sufficiently comply with the modern requirements to functional integration, miniaturization, and constant increase of operation frequencies. Sharp development of acoustoelectronic devices, in particular, surface acoustic wave (SAW) devices in the recent decades is connected with the development of telecommunication systems, primarily mobile telephony. Replacing their possible competitors, acoustoelectronic devices have firmly took this niche showing their undisputed advantages.

The principle of operation of acoustoelectronic devices is based on generation, propagation, reflection, and reception of acoustic waves in solids. The study and development of such devices are connected with such fields of science as solid-state physics, acoustics, crystallography, radio engineering network theory, and microelectronics.

One of the first and the main stages of the development of acoustoelectronic devices is substrate selection. At the same time, substrate is the base of the entire device, converts electric power into acoustic one, and is a medium where acoustic waves propagates. Characteristics of the waves and thus the performance of acoustoelectronic devices being developed depend on the characteristics of the medium. Recently, materials for acoustoelectronic devices have attracted more attention, due to the fact that new materials can both enhance the performance of the devices and widen their functional capabilities and the area where they can be used. Both new materials are being synthesized, and optimal directions of wave propagation are being searched for in already known materials. Special search techniques and objective functions are being developed [1].

2. THE DATE FOR THE INFORMATION AND REFERENCE SYSTEM

The main parameters of piezoelectric materials used for acoustoelectronic device manufacturing are [2, 3, 4]:

- phase velocity of acoustic waves (surface, leaking, etc., ones) on a free and metallized surface;
- electromechanical coupling coefficient;
- material's decay constants for acoustic waves
- temperature coefficients of delay, frequency, velocity;
- anisotropy coefficients;
- power flow angle;
- effective dielectric permittivity;
- capacitance/finger pair/unit length;

Also are often required data on material's density, hardness, and other reference information. These main parameters not only allow the developer to estimate the possibility to use a certain material in devices with specified parameters, they can also be further used in mathematical simulation. The main problem is that all those data are scattered over many sources, or even incomplete. Data may be either evaluated or measured.

3. NEW MATERIALS FOR ACOUSTOELECTRONIC DEVICES

In last decades, quartz, lithium niobate, and lithium tantalate were the crystals most often used in acoustoelectronic devices [3, 5]. But new requirements to technical characteristics of telecommunication equipment demand that new materials are to be used.

Recently, so-called “quartz-like” materials attract much attention. Those materials have temperature characteristics similar to those of quartz, but higher values of electromechanical coupling coefficient provide higher relative bandwidths. SAW velocities, lower than those of quartz, allow the developers to manufacture smaller and thus less expensive devices. Among those materials, langasite, langanite, langatate, and gallium orthophosphate show particularly good results [6…9]. Four new oriented crystals from the langasite family, CNGS, CTGS, SNGS, STGS, have additional advantages [10, 12, 13], showing a good future in intermediate frequency circuitry in various processing systems. Currently, new crystals from the langasite family are being actively investigated [10, 11].

New data on various materials and their characteristics are constantly being published in modern scientific literature. Taking into account a constant growth of information flows and the power of modern computing, it is necessary to systematize, store, and provide a convenient access to various reference data. The development of the database and its control system solves these tasks.

4. THE INFORMATION AND REFERENCE SYSTEM FOR ACOUSTOELECTRONIC DEVICE DESIGN

An information and reference system for acoustoelectronic device design has been developed on the Department of Physical and Theoretical Fundamentals of Radioelectronics in the St.-Petersburg State University for Aerospace Instrumentation. The core of this system is a database containing information on materials and literature. The Section “Materials” contains the parameters of materials used as substrates for acoustoelectronic devices. The database allows its users to quickly and conveniently search and retrieve necessary data on material’s parameters. It also provides information about publications containing data on the material of interest. Thus the user can obtain information not only on the parameters but also on the source of the information. Fig. 1 shows the result of database work in the “Materials” Section. The result of a search for the values of velocity and electromechanical coupling coefficient is shown.

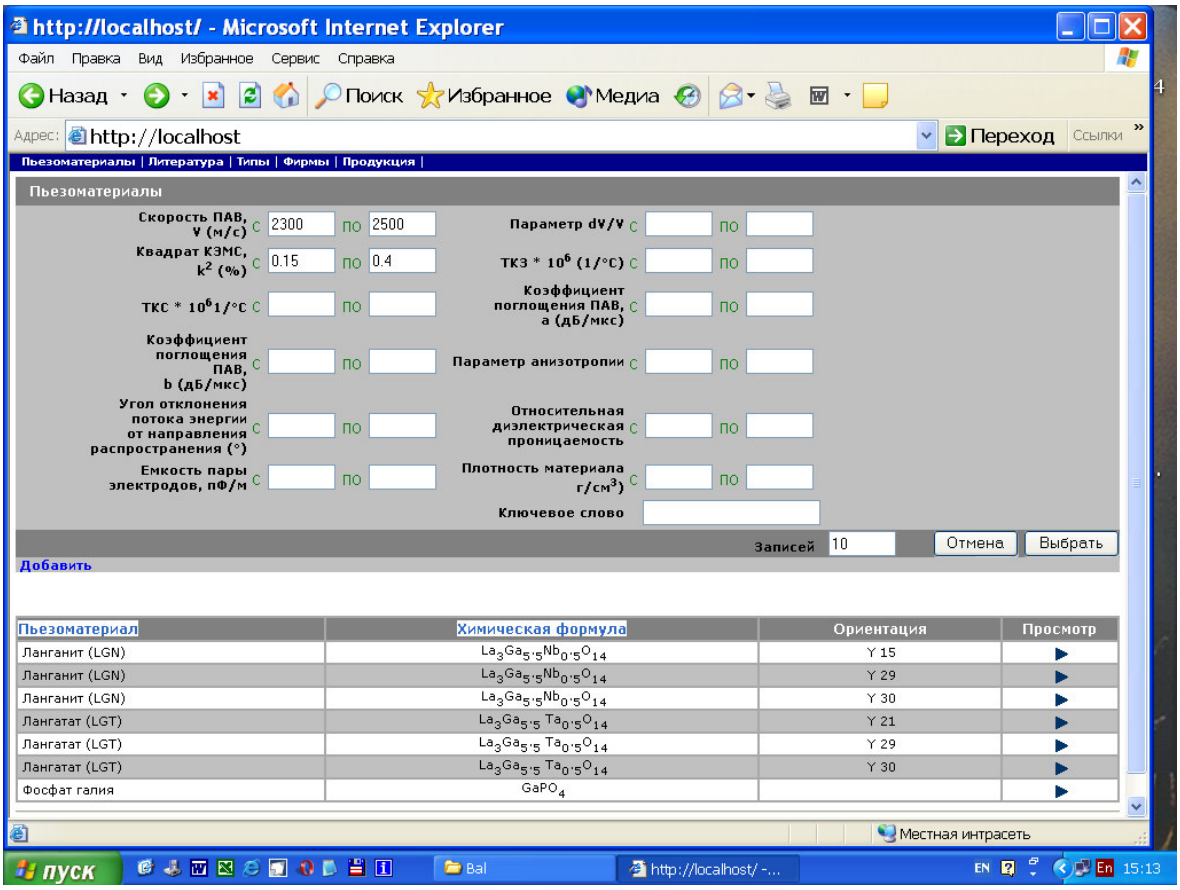


Fig. 1. Example of the Database work. Result on the search for the values of velocity and electromechanical coupling coefficient.

In the "Literature" Section, information on books, reference books, articles, proceedings and transactions of symposiums and conferences on acoustoelectronic devices is systematized. Brief summaries and keywords are included allowing for faster search for information in the Database. For some publications, the Database contains full texts in the Adobe Acrobat format (PDF). Fig. 2 shows an example of the Database work in the "Literature" section.

The software of the Database is developed using the client-server model, the client part being a program written on Perl. The data on the server are stored in MySQL tables. The Database is very resource-friendly. The following minimum hard and software are required: a PC with at least Pentium II or alike, at least 128 MB of RAM, Windows NT/2000/XP installed, at least 500 MB of HDD free space.

5. CONCLUSION

The developed Database is modular and can be expanded by adding new sections, and new information to the existing ones. The information in the Database can be searched for, retrieve, added, and edited. The Database is intended for specialists in acoustoelectronics, students, and post-graduates.

This work is financially supported by Russian Fund for Basic Research, Grant № 03-07-90176, and is at its disposal.

ACKNOWLEDGMENTS

The authors would like to thank the Russian Fund for Basic Research for financial support and Dmitry Buchilov for the help of database software development.

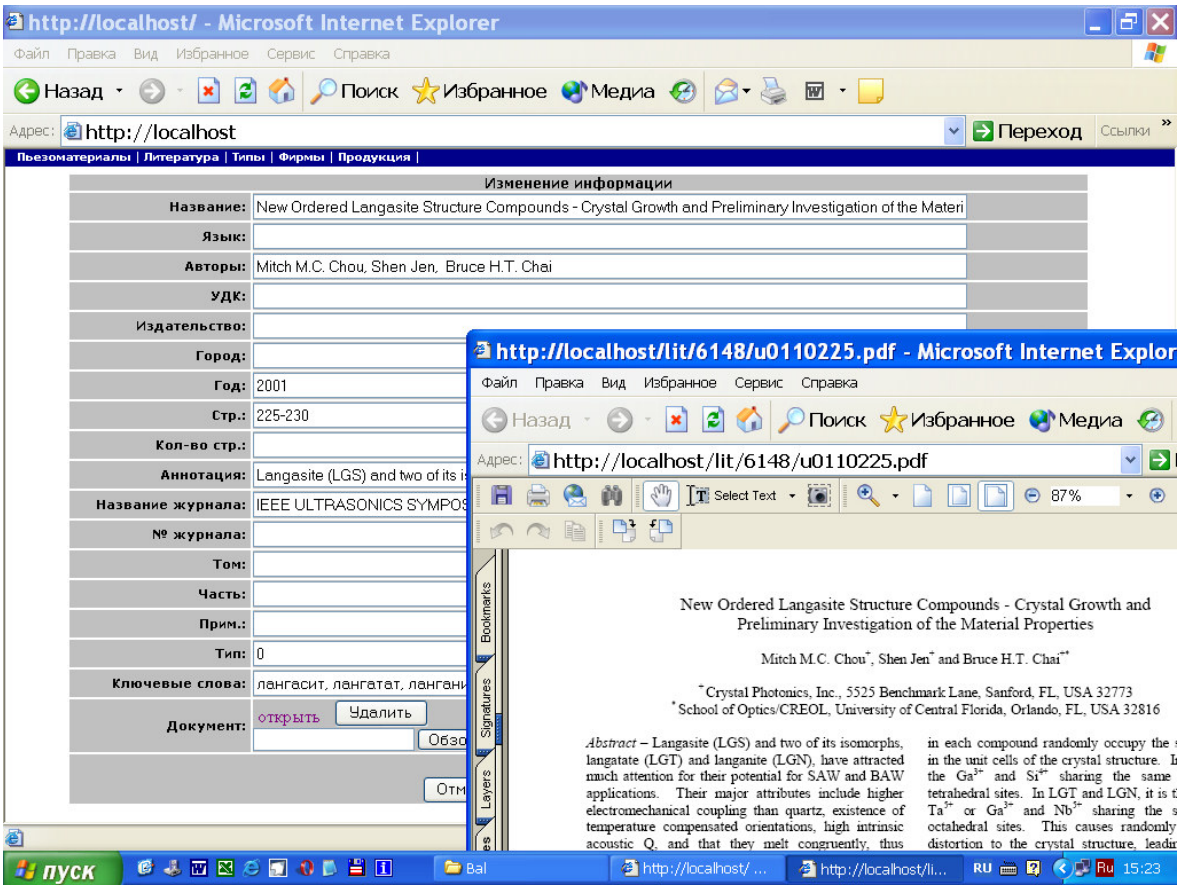


Fig. 2. Example of the Database work in the "Literature" section.

REFERENCES

[1] M.Yu. Dvoeshorstov, S.G. Petrov, V.I. Cherednik, A.P. Chirimanov. New Optimal Orientations for Surface Acoustic Paves in the Piezoelectric Crystals of Langasite, Langanite, and Langatate. @ Journal of Theor. Phys., vol. 72, #8, pp.103-108, 2002 (in Russian).

[2] D.P. Morgan. Surface-Wave Devices for Signal Processing. Amsterdam: The Netherlands: Elservier, 1991.

[3] Ken-ya Hashimoto. Surface Acoustic Wave Devices in Telecommunicatios: modeling and simulation. Springer, 2000 (Engineering online library).

[4] C.K. Campbell. Surface Acoustic Wave Devices for Mobile and Wireless Communications. Academic Press. Inc. 1998.

[5] C.C. Ruppel, L. Reinld, R. Weigel. SAW Devices and Their Wireless Communication Applications. @ IEEE Microwave Magazine, pp. 65-71, June 2002.

[6] D.C. Malocha, H Francois-Saint-Cyr, K. Richardson, R Helmbold. Measurements of LGS, LGN, and LGT Thermal Coefficients of Expansion and Density. @ IEEE Transactions on UFFC, vol. 49, pp.350-355, No.3, January 2002.

[7] D.C. Malocha, M.P. Cunha, D. Puccio, K. Casey. Investigations of Langanite and Langatate materials for use in SAW device applications@, in Proceedings of the IEEE Ultrasonic Symposium, 2001, pp. 231-234.

[8] M.P. da Cunha, S.A. Fagundes. Investigation on Recent Quartz-like Materials for SAW Applications. @ IEEE Transactions on UFFC, vol. 46, pp. 1583-1589, no.6, November 1999.

[9] M. P. da Cunha, T.B. Pollard, H. Whitehouse, P.M. Worsch. $GaPO_4$ SAW Devices: Measured and Predicted Propagation Properties@, in Proceedings of the IEEE Ultrasonic Symposium, 2003, pp. 110-113.

[10] M.M.C. Chou, S. Jen, B.H.T. Chai. New Ordered Langasite Structure Compounds - Crystal Growth and Preliminary Investigation of the Material Properties@, in Proceedings of the IEEE Ultrasonic Symposium, 2001, pp. 225-230.

[11] D. Puccio, D.C. Malocha, and N. Saldanha. Investigations of New Materials, CTGS and CNGS for SAW Application@, in Proceedings of the IEEE Ultrasonic Symposium, 2003, pp. 96-99.

[12] E. Chilla, R. Kunze, M. Weihnacht, J. Bohm, R.B. Heimann, M. Hengst, U. Straube. Acoustic Waves Measurements on SNGS Crystals and Determination of Material Constants@, in Proceedings of the IEEE Ultrasonic Symposium, 2003, pp. 92-95.

[13] D. Puccio, D.C. Malocha, M.M.C. Chou. Investigations of STGS, SNGS, CTGS & CNGS Materials for Use in SAW Applications@, in Proceedings of the IEEE International Frequency Control Symposium, 2003, pp.627-630.

LONGITUDINALLY-COUPLED BULK ACOUSTIC WAVE RESONATORS FILTER FOR 5 GHZ APPLICATIONS

A. Reinhardt[1], F. Martin[2], R. Lanz[2], P. Muralt[2], V. Laude[1], S. Ballandras[1], W. Steichen[3]
[1] Institut FEMTO-ST, département LPMO, 32 avenue de l'Observatoire, 25044 Besançon Cedex, France
[2] Ecole Polytechnique Fédérale de Lausanne, Ceramics Laboratory, Lausanne, Switzerland
[3] TEMEX Microsonics, 399 route des Crêtes, 06904 Sophia Antipolis Cedex, France
Corresponding author : alexandre.reinhardt@lpmo.edu

ABSTRACT

Wireless applications require wide bandwidth and high selectivity filters in the Gigahertz range. We explore the possibility to implement such filters by coupling longitudinally Thin-film Bulk Acoustic wave Resonators (FBAR) in so-called Coupled Resonator Filters (CRF). We describe the simulation method and design rules employed. We then focus on the fabrication before considering the measurements and the analysis of a preliminary fabrication result.

1. INTRODUCTION

Wireless applications require wide bandwidth and high selectivity filters in the Gigahertz range. With the trend in increasing operating frequencies, new filter structures need to be explored to provide alternative or complementary solutions to standard Surface Acoustic Wave (SAW) ladder filters. A key technology could be thin film bulk acoustic wave resonators (FBAR) exploiting the thickness-extensional vibration mode of piezoelectric thin films [1]. These resonators are most of the time electrically coupled to form a ladder-type filter. Another way of achieving a filter, fulfilling more easily the ever more stringent size requirements, is to longitudinally couple two or more resonators, separated by a set of coupling layers to control the spacing between resonances, in the so-called coupled resonator filters (CRF) [2].

We have used a plane-wave model to describe the behavior of electromechanical waves in multilayered structures. This model is based on a Fahmy-Adler formulation for piezoelectric materials or metals, and uses a scattering matrix algorithm to calculate electromechanical fields in a stack of materials. We have adapted it to the calculation of the electrical response of CRF's. This has allowed us to design theoretically a filter for 5 GHz applications. We have then fabricated a first set of demonstrators.

In a first part, we briefly describe the principles and the simulation of the filter, and give its theoretical performance. In a second part, we focus on the fabrication process. Finally, we present measurements of devices obtained after a first run. Their response is analyzed to show the influence of some technological factors, as well as of some design points.

2. SIMULATION AND DESIGN OF THE FILTER

We use a model that calculates the propagation of waves in a multilayer structure to design the CRF. First, we recall the principles of this model, and then the fundamental concepts of coupled resonator filter operation.

2.1. Planes waves in stratified structures

We use a Fahmy-Adler formulation [3] to study waves in a structure made of thin film layers. The electromechanical fields in a medium are considered as a linear superposition of partial waves (eight for a piezoelectric or a dielectric, six for a perfectly conducting metal, four for a fluid, such as air): two of these partial waves are purely electrostatic in an insulating medium, while the remaining others are purely mechanical, or electromechanical in case of piezoelectrically coupled partial waves in a piezoelectric. Their polarization and propagation characteristics are obtained by solving an eigenvalue problem depending on material constants. It is this possible to calculate the fields in each layer in a matrix form:

$$\begin{aligned}\mathbf{h} &= \left(u_1 \quad u_2 \quad u_3 \quad \Phi \quad T_{21} \quad T_{22} \quad T_{23} \quad D_2 \right)^T \\ &= F\Delta\left(x_2\right)\mathbf{a}, \end{aligned} \tag{1}$$

where $\mathbf{u}$ is the displacement vector, Φ the electric potential, T_{2i} the vertical stress components, according to the conventions depicted in Fig. 1 and D_2 the vertical component of the electric displacement vector. In the second line, F is a matrix containing the partial waves polarizations, Δ a diagonal matrix containing the propagation terms along the vertical direction, and $\mathbf{a}$ the vector containing the amplitudes of the partial waves.

Due to the linearity of our model and of classical network theory, it is convenient to describe a multiport device made of a stack of materials by its admittance matrix, defined as

$$\begin{pmatrix} I_1 \\ \vdots \\ I_n \end{pmatrix} = Y \begin{pmatrix} V_1 \\ \vdots \\ V_n \end{pmatrix}, \tag{2}$$

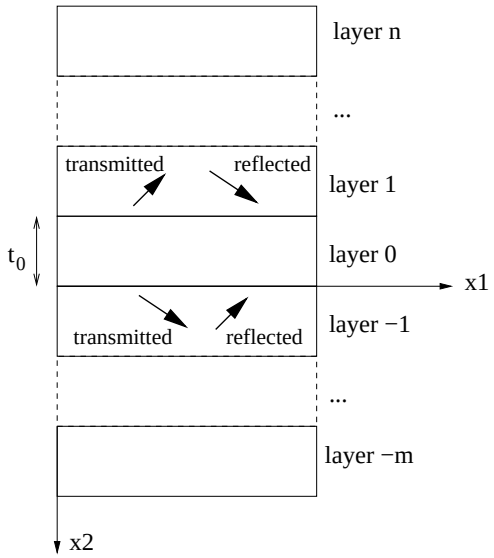


Fig. 1. Conventions adopted for representing the propagation of plane waves in a multilayer.

where I_i and V_i are respectively the current and voltage applied at the port i. This means that the components of the admittance matrix can be determined by columns, setting for example a port j at a reference potential while grounding all others and calculating the currents generated by this excitation at all ports. These currents are obtained by calculating the charge accumulated on the metal layers, which is equal to the value of the electric displacement at the interface between an insulating medium and the metal. As a consequence, the current generated in a layer, with index 0 in Fig. 1, is obtained by

$$I_0 = j\omega \left(D_2 \left(x_2 = t_0 \right) - D_2 \left(x_2 = 0 \right) \right). \qquad (3)$$

To improve the stability of the original Fahmy algorithm, we use this formulation within a scattering matrix method [4]. Partial waves are splitted into two groups, depending on their propagation direction. These groups are termed either "transmitted" or "reflected" with respect to the source electrode position (layer 0 in Fig. 1). The partial waves amplitudes in these two groups are related by the means of a reflexion matrix. At the beginning of the algorithm, reflexion matrices are initialized on the structure boundaries, where stress free and optionally charge free conditions are applied, providing total reflexion of waves, or where an infinite medium is considered, providing no reflexion at all. The matrices are then transfered to the other sides of the layers by considering the propagation of all partial waves through the layer and its consequences for the reflexion matrix. By writing that the **h** vector is continuous across the interface between two adjacent layers, it is possible to determine how partial waves are converted through the interface, and thus how to transfer the reflexion matrix from one layer to the other. The only exception is when dealing with the interface between a metallic layer and an insulating one: the electric displacement is not continuous any more, but equals the electric charge accumulated at the electrode boundary and the electric potential has to be set to 0. But these are the only differences, and this does not modify the physical meaning of the operation [5]. This scheme is repeated until the reflection matrices on both sides of the structure are transfered to the source electrode. This way, all electric or mechanical boundary conditions, as well as the electromechanical properties of all layers are reported to the excitation layer.

Given the reflexion matrices on both sides of the source electrode and also the mechanical properties of the metal and of the neighboring materials, it is possible to write a relation that links the amplitudes of the partial waves within the -1-th, 0-th and 1st layers to the applied voltage. Amplitudes in the other layers can then be recursively retrieved by using once more the relations obtained when considering the mode conversions at the interfaces needed in the calculation of the reflexion matrices. Thus, the calculation of the electromechanical fields is obvious from Eq. (1), and then the components of the admittance matrix are obtained from the electric displacement fields using Eq. (3).

2.2. Filter design

Now that we have described the tool used for simulating the electric response of a stratified structure, we focus on the design of a demonstrator. The basic structure from which we start is the Stacked Crystal Filter (SCR) introduced by Ballato *et al.* [6] in which two piezoelectric resonators are in contact in order to be mechanically coupled, as shown in Fig. 2. Such a device basically exhibits three resonances, corresponding respectively to the whole stack behaving like a classical resonator, to both resonators vibrating at their resonance frequencies, and to the third harmonic of the fundamental resonance. Usually, only the second one is considered when looking for transmission characteristics. But the use of this structure is limited to very narrow bandwidths.

In order to increase the bandwidth, the degree of coupling between the two resonators has to be decreased.

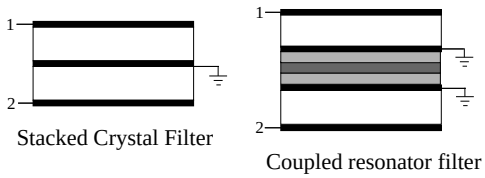


Fig. 2. Topologies of filters exploiting the longitudinal coupling of bulk acoustic wave resonators.

Lakin *et al.* [2] proposed to insert a set of quarter-wavelength layers forming a partial Bragg mirror between the two resonators, to form a Coupled Resonator Filter (CRF). This takes benefit from the fact that such a structure exhibits symmetric and antisymmetric thickness vibration modes, corresponding to Lamb waves at their cutoff frequency. When these resonances come close enough, the transmission spectrum of the device exhibits a bandpass behavior. The gap between the cutoff frequencies of two modes can be adjusted by controlling the number of coupling layers, as shown in Fig. 3. The rule is that the gap between the two Lamb modes decreases with the number of coupling layers. In order to precisely control the filter bandwidth, it is also possible to adjust the thickness of the electrodes located between the resonators and the coupling layers.

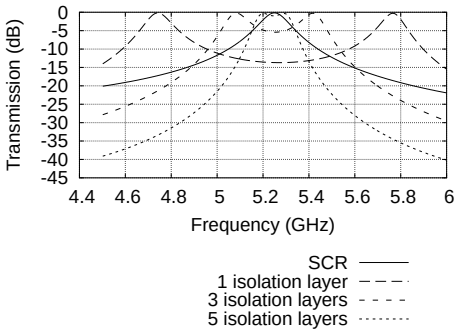


Fig. 3. Narrow band response of a CRF as a function of the number of coupling layers.

We have noted that the SCR has three fundamental resonances. A solution to reject the two unwanted resonances is to use the fact that an acoustic Bragg mirror can be fabricated by periodically piling up quarter wavelength layers of alternatively low and high impedance materials, and that such a mirror has a high reflexion coefficient only within a given frequency window. So, by adjusting the properties of the mirror, which is achieved by modifying the thicknesses of the layers it is made of, the rejection of unwanted resonances can be optimized. At the same time, the mirror provides an acoustic isolation between the bottom resonator and the substrate. The effect of this mirror is depicted in Fig. 4.

Finally, the sharpness of the rejection is controlled by the number of resonators included in the filter. In the basic configuration, the filter function corresponds to a two-pole theoretical filter response. If more selectivity is needed, a function with more poles is needed. This means that more resonators are required. A first way to add resonators is to couple not only two, but three resonators longitudinally. This makes it necessary to deposit and pattern more layers, but the lateral size of the device remains broadly the same. Another way is to electrically cascade two sections. This method ensures

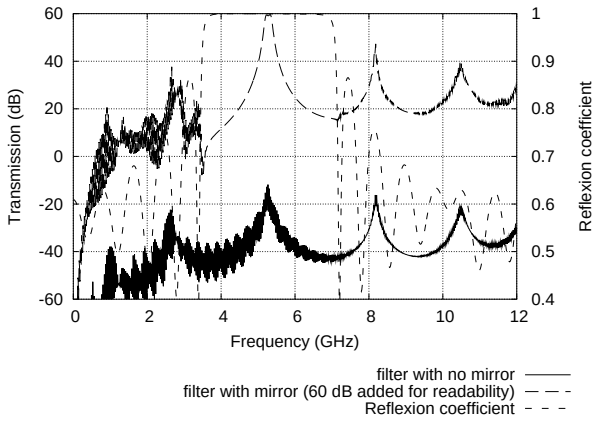


Fig. 4. Effect of the Bragg mirror on the transmission response.

that no more layers are needed for the fabrication, but increases the size of the final device. However, as the dimensions of one section are very small, the addition of another section is the best solution, as it also allows input and output electrodes to be located at the same level.

3. FILTER FABRICATION

Considering the design rules presented in the previous section, we have set up a fabrication process for a demonstrator in the 5 GHz band. Materials are chosen to be Aluminum Nitride (AlN) for the resonators, grown on Platinum in order to obtain well textured piezoelectric layers. The coupling layers and the Bragg mirror are made of silica for the low impedance material, and once again of AlN for the high impedance material. Simulations using previously fitted material constants were performed. In order to limit the mass loading of the resonators, the electrodes must only be 60 nm thick, for piezoelectric films of 435 nm. Two sections are needed to obtain a selectivity that reaches the common specifications for filters in this range. For a 50 Ω adaptation, an active area of 50×60 μm^2 per section is required.

AlN films are deposited by reactive pulsed DC magnetron sputtering, using a 200 mm-diameter, 99.999 % Al target. The system is pumped down to a base pressure of less than 2×10^{-7} mbar. The substrate is then heated to 300°C in an atmosphere containing N_2 at a pressure of 4×10^{-3} mbar. Where piezoelectric properties are of importance, AlN is grown on (111)-oriented Pt electrodes, which are themselves prepared with a 100 Å Ti adhesion layer. More details about the properties of these layers have already been presented in other papers [7], [8]. Other AlN layers are grown directly on the 525 μm thick Si (100) substrate, for the very first AlN layer in the Bragg mirror, or on silica films, for the

layers in the mirror or in the set of coupling layers.

The acoustic mirror and the piezoelectric layer forming the bottom resonator are left unpatterned on the whole substrate surface, whereas all electrodes and other layers, including the coupling layers and the top resonator piezoelectric layer, are patterned by dry etching in an ICP reactor.

As we have already shown, electrodes are very thin, compared to the insulating layers. Early tests have shown that the metallic strips may not continuously cover the steps obtained after patterning the insulating films, which could degrade the electric connection, as appears from Fig. 5. So, it is necessary to add an amount of metal to cover these steps: a supplementary thick layer is deposited by a lift-off process over the existing metalizations, but care is taken to avoid overloading the active areas.

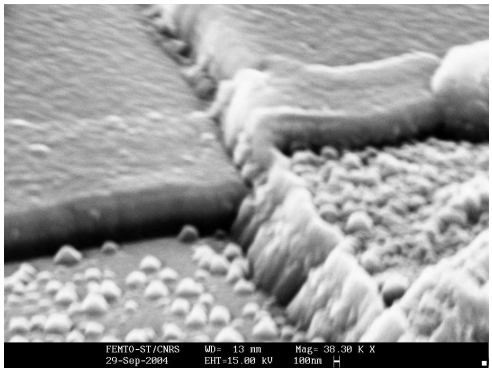


Fig. 5. SEM view of a metallic strip trying to follow the step corresponding to the boundary of the coupling layers.

Fig. 6 shows a SEM view of the device implementation at the end of the fabrication process. The whole device holds in a 600×700 μm^2 area, with most of its surface dedicated to the layout of the electric connections. Fig. 7 shows a cross section of the wafer, revealing the structure of the deposited layers. AlN films exhibit a colonnar structure, with the column size depending on the nucleation surface. Clearly, layers grown over silica have larger columns than those grown over Pt or Si. This is because silica layers provide a much rougher surface than Pt or Si.

4. CHARACTERIZATION

The devices were measured using two coplanar probes and a network analyzer. The transmission is shown in Fig. 8. This latter seems quite different from the theoretical response shown in Fig. 4. A first analysis revealed that the design of our filter is practically not adapted to a 50 Ω load. A second reason accounting for the high insertion losses is that at an early stage of the fabrication,

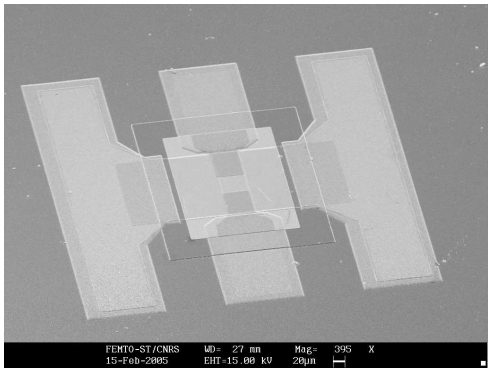


Fig. 6. SEM view of the filter implementation.

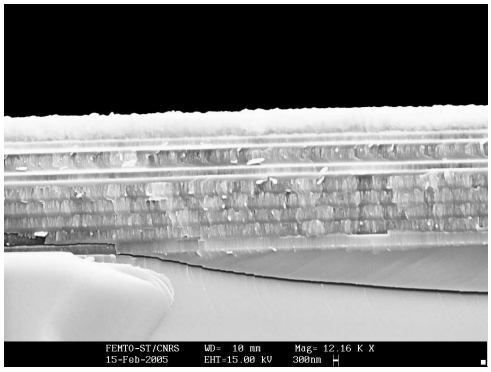


Fig. 7. SEM view of a cross section of the wafer.

we got problems in stripping the photoresist after a dry etching of Pt electrodes. A chemical reaction seems to occur in the ICP reactor between the chlorine species in the plasma and the organic compounds in the resist, catalyzed by platinum. This causes the resist to become very hard to strip, and a long descum step is required to get totally rid of it. At one step in our process, some photoresist has remained on the electrodes, which has caused all layers deposited afterwards to grow on an unflat surface. Because of this, the properties of these layers need to be degraded. The actual material parameters have been obtained by measuring and fitting test devices also located on the wafer. Taking these corrections into account, but without modifying the theoretical design to account for technological dispersions in the thicknesses of deposited layers, we obtain the second curve shown in Fig. 8.

To some extent, the layout of the device is responsible for some characteristics of the filter transmission. We have investigated some effects that can arise from the geometry of the electric strips and from the parasitics that may appear but that we had not considered in the first design step. All subsequent simulations were performed using the Advanced Design System software, from Agilent Technologies. Some lumped elements were added to the acoustic simulations obtained using the

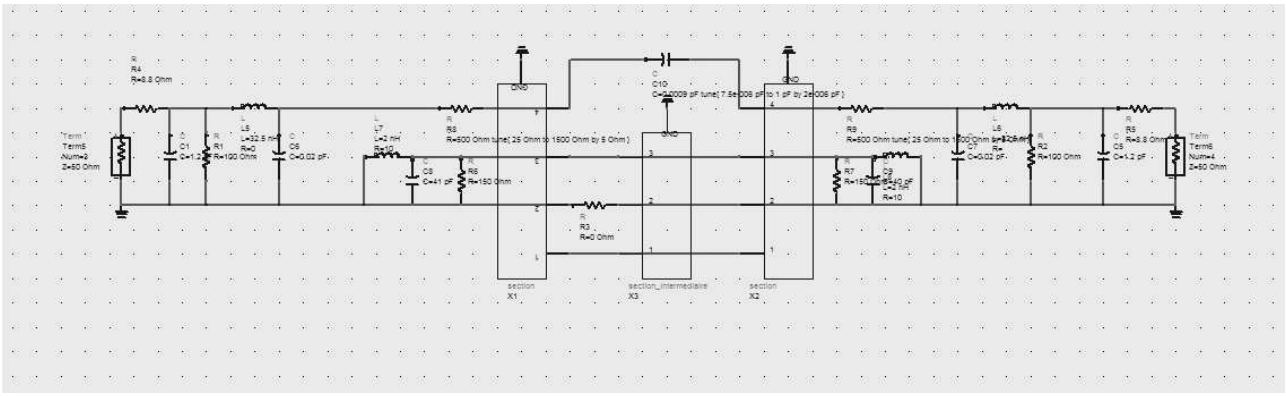


Fig. 9. Representation of parasitics by adding lumped elements to the acoustic simulation.

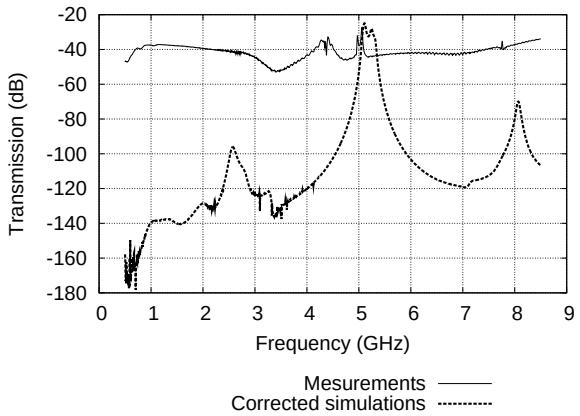


Fig. 8. Measured transmission compared to the theoretical response including corrections for material properties.

model described in section 2. The complete circuit is shown in Fig. 9.

The considered parasitic effects are: (i) two transmission lines, corresponding to the lines joining the pads were the input and output probes are connected during measurements and the active areas of the device, need to be considered. These transmission lines are not adapted to a 50 Ω load and thus cause reflexions to occur at the input and the output of the filter.

(ii) another section of transmission lines needs to be added, to account for the fact that when the input and output metallic strips step over the coupling layers or the top AlN layer, the characteristics of the transmission lines are modified. In particular, as the gap between the strips is slightly increased, and the equivalent dielectric constant of the insulating layers on which they are deposited is decreased (the connexion pads lay on an AlN layer, whereas at the end, the strips lay on an $AlN/SiO_2/AlN/Si0_2$ multilayer), the capacitance of the line is reduced, whereas the inductance is increased. These two transmission lines are used to fit the reflection coefficient. Although they explain very well the high frequency behavior of the input and the output, some low frequency parasitics in the reflexion are not yet explained and need further investigations.

(iii) the initial design staassumed the ground electrodes of the two resonators in a section were connected together, so that the coupling layers get shorted, which prevents the AlN layer located between the resonators to vibrate because of an applied electric field. However, in the implementation, the first electrode is deposited, then the coupling layers are deposited and patterned, and finally the second ground electrode is fabricated. Connexion between them is done by overlapping the two metallic strips. The last that has been deposited has to follow the step caused by the coupling layers. As a consequence, the connexion may not be perfect. Moreover, parts of these electrodes are partially covered with insulating films, so that they cannot benefit from the additional metalization that is added at the end of the process. For this reason, electrodes exhibit a high resistivity. This is modelled by inserting an inductance and a resistance to one of the ground paths. As the electrodes may finally not be at the same electric potential, a capacitance is added, corresponding to the area where both ground strips are facing each other on each side of the coupling layers, not including the active area, which is included in the acoustic simulation. These elements influence mostly the phase of the transmission coefficient, and allow to fit the phase of the measurements.

(iv) a capacitance is added between the input of the first section and the output of the second one. This is intended to simulate the capacitance that may exist between the input and the output electrodes. This capacitance induces an electrical resonance with elements from the transmission lines, and is responsible for the resonance that appears between 4 and 5 GHz in the measurements. The level of this parasitic is extremely sensitive to the value of the capacitance, showing that care must be taken to minimize its value in future

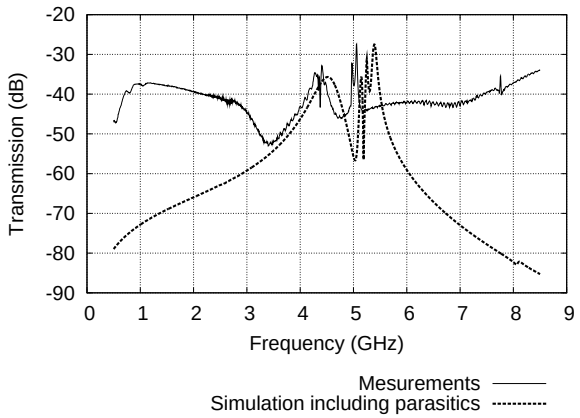


Fig. 10. Measured transmission compared to the theoretical response including some parasitic elements.

designs. The reason is that it provides a direct path between input and output in a given frequency range. The sensitivity can be reduced by considering ohmic losses at the end of the input and output paths. But even with high resistivity (500 Ω are added between the input and the active area), even a low value of the input to output parasitic capacitance (9×10^{-4} pF in our fits) has a considerable influence on the transmission.

With all these elements added, the final simulated response is shown in Fig. 10. The in-band response and the transmission between 3 and 5 GHz are quite well explained. But there remains to explain the rejection floor at low and high frequency.

5. CONCLUSION

We have written a simulation program that is able to simulate the electric response of a device implemented as a stratified structure. This simulation tool is able to handle multiport devices. We have used it to design a coupled resonator filter in the 5 GHz area. We have performed a first run for the fabrication of this filter. Measurements have revealed some problems associated with the layout of the device, especially the influence of parasitic capacitances, the need to perform electromagnetic simulations to ensure that the input and output are well adapted to a 50 Ω measurement system, and eventually the choice of another material than Platinum which enables to grow highly textured AlN layers but with lower resistivity. Thus, a redesign of the layout of the filter should be performed before another fabrication run is performed.

Acknowledgments

The authors want to thank the staff of the Centre de Microtechniques from the Ecole Polytechnique Federale de Lausanne for their help during the fabrication. They are very grateful to the design team of Temex Microsonics for their help in the measurements and the analysis of the parasitics. They also thank Dr. Sébastien Euphrasie for electromagnetic simulations, Yannick Gruson for help with ADS simulations, and Sarah Benchabane, Dr. Abdelkrim Khelif and Dr. Jean-Yves Rauch for enlightening discussions.

REFERENCES

[1] K. M. Lakin, G. R. Kline, R. S. Ketcham, A. E. Landin, W. A. Burkland, K. T. McCarron, S. D. Braymen, and S. G. Burns, “Thin film resonator technology,” in Proceedings of the 41st Annual Frequency Control Symposium, 1987, pp. 371–381.

[2] K. M. Lakin, J. Belsick, J. F. McDonald, and K. T. McCarron, “High performance stacked crystal filters for GPS and wide bandwidth applications,” in Proceedings of 2001 IEEE Ultrasonics Symposium, 2001, pp. 833–838.

[3] A. H. Fahmy and E. L. Adler, “Propagation of acoustic surface waves in multilayers : a matrix description,” Applied Physics Letters, vol. 22, pp. 495–497, 1973.

[4] T. Pastureaud, V. Laude, and S. Ballandras, “Stable scattering-matrix method for surface acoustic waves in piezoelectric multilayers,” Applied Physics Letters, vol. 80, no. 14, pp. 2544–2546, avril 2002.

[5] A. Reinhardt, T. Pastureaud, S. Ballandras, and V. Laude, “Scattering matrix method for modeling acoustic waves in piezoelectric, fluid, and metallic multilayers,” Journal of Applied Physics, vol. 94, no. 10, pp. 6923–6931, novembre 2003.

[6] A. Ballato and T. Lukaszek, “A novel frequency selective device: The stacked-crystal filter,” in Proceedings of 1973 IEEE Frequency Control Symposium, 1973, pp. 262–269.

[7] M.-A. Dubois and P. Muralt, “Properties of aluminum nitride thin films for piezoelectric transducers and microwave filter applications,” Applied Physics Letters, vol. 74, no. 20, pp. 3032–3034, May 1999.

[8] F. Martin, P. Muralt, M.-A. Dubois, and A. Pezous, “Thickness dependence of the properties of highly c-axis textured AlN thin films,” Journal of Vacuum Science and Technology A, vol. 22, no. 2, pp. 361–365, March-April 2004.

THIRD-ORDER ELASTIC CONSTANTS FOR SI FROM FIRST PRINCIPLES

Kirill I. Volianskiy
St.Petersburg State University of Aerospace Instrumentation
67, Bolshaya Morskaya Str., St.Petersburg, Russia, 190000
E-mail: weconf@home.ru

ABSTRACT

A method of calculating the third order elastic constants with using the local density approximation of the density-functional theory is presented. In this work, the method is applied for calculating the third-order elastic constants of Si. ABINIT package is used for the calculation. The stress tensor at defined strains is computed with using the package and then the third-order elastic constants are calculated by finite-differences. The calculated data are in good accordance with experimental data.

1. INTRODUCTION

During designing and analysis of some acoustoelectronic sensors it is necessary sometimes to take into account the nonlinear elastic properties of materials of the sensors. For example, these properties are significant in designing and analysis of torque sensors [1]. Third-order elastic constants are used for describing such properties. These constants were already defined for some materials, but the process of their experimental defining is rather complex one requiring high precision equipment. Therefore it is not always affordable for researchers. In this connection the possibility to define the third-order elastic constants by computations with minimal set of initial experimental data, which are available for all, is actual. Local density approximation (LDA) of density-functional theory (DFT) [2], as implemented in the ABINIT package [4], is applied for computing the constants in the present paper. The theory was created with using quantum mechanics principles The peculiarity of methods based on quantum mechanics principles or ab initio principles is that they allow to define large variety of material properties with basing on its internal structure.

2. COMPUTATIONAL TECHNIQUE

Third-order elastic constants can be expressed through the thermodynamics tensions t_{ij} [3]

$$c_{ijklmn} = \frac{\partial^2 t_{ij}}{\partial \eta_{kl} \partial \eta_{mn}}, \qquad (1)$$

where η_{ij} – Lagrange strain.

Since ABINIT package [4], containing an algorithm for computing stress σ_{ij}, was used in the present paper it is necessary to express the elastic constants through the stress. For this purpose, it is possible to use the connection between the thermodynamics tensions and the stress [4]

$$t_{ij} = \frac{1}{J} \frac{\partial a_i}{\partial x_k} \frac{\partial a_j}{\partial x_m} \sigma_{km}, \qquad (2)$$

where J – the ratio of deformed element volume to undeformed one, $\partial a_i/\partial x_k$ – the strain gradients in spatial coordinates. The deformed element volume V is connected with the undeformed one V_o by the following expression

$$V = (1 + \eta_{ij}\delta_{ij}) V_o . \qquad (3)$$

As a result

$$c_{ijklmn} = \left(\frac{\partial^2 \left((1 + \eta_{ij}\delta_{ij}) \frac{\partial a_i}{\partial x_s} \frac{\partial a_j}{\partial x_t} \sigma_{st} \right)}{\partial \eta_{kl} \partial \eta_{mn}} \right) . \qquad (4)$$

There are six independent third-order elastic constants for Si [5]: c111, c112, c123, c144, c155, c456. It is possible to expand the expression (4) for the constants (in reduced indexes) with taking into account the isotropy of the strain:

$$c_{111} = \frac{\frac{\partial^2 \sigma_1}{\partial \eta_1^2}}{1 + \eta_1} - 2 \frac{\frac{\partial \sigma_1}{\partial \eta_1}}{(1 + \eta_1)^2} + 2 \frac{\sigma_1}{(1 + \eta_1)^3}$$

$$c_{112} = \frac{\partial^2 \sigma_2}{\partial \eta_1^2} (1 + \eta_1) + 2 \frac{\partial \sigma_2}{\partial \eta_1}$$

$$c_{123} = (1 + \eta_2)(1 + \eta_3) \frac{\partial^2 \sigma_1}{\partial \eta_2 \partial \eta_3} +$$

$$+ (1 + \eta_2) \frac{\partial \sigma_1}{\partial \eta_2} + (1 + \eta_3) \frac{\partial \sigma_1}{\partial \eta_3} + \qquad (5)$$

$$+ \sigma_1$$

$$c_{144} = \frac{\partial^2 \sigma_1}{\partial \eta_4^2}$$

$$c_{155} = \frac{\partial^2 \sigma_1}{\partial \eta_5^2}$$

$$c_{456} = \frac{\partial^2 \sigma_4}{\partial \eta_5 \partial \eta_6}$$

In the present paper the derivatives to strain were computed with finite-differences. The strain difference was $4*10^{-5}$. Teter Pade parametrization [6] was used for the exchange-correlation functional. The all-electron potentials were replaced by norm-conserving

pseudopotentials [7]. Monkhorst-Pack grid [8] 6x6x6 was used for the sample in the Brillouin zone. Converged results were reached with 65 hartree kinetic energy cutoff for the plane waves and with the cutoff smearing 0,5 hartree.

3. RESULTS

Si is the cubic crystal of m3m class [10]. The elementary cell of Si was presented by two atoms with the reduced internal coordinates [9] Si_1 = (0, 0, 0) и Si_2 = (0.25, 0.25, 0.25).

The results of third-order elastic constants computation in compare with experimental data [10] are presented in the table 1.

TABLE 1. Third-order elastic constants c_{mnp}, 10^{11}N/m^2

Constant	Present work	Experiment [10]	Experiment [10]
c111	-6.14	-7,95	-8,25±0,10
c112	-3.93	-4,45	-4,51±0,05
c123	-0.84	-0,75	-0,64±0,10
c144	-1.94	0,15	0,12±0,25
c155	-2.68	-3,10	-3,10±0,1
c456	-1.06	-0,86	-0,64±0,2

As can be seen in the table, the computational results based on LDA give rather good approximation to the experimental data. But they have 16% discrepancy from the experimental ones in average excluding c144. The discrepancy can be produced by the approximation accuracy and by difference of temperature of material, because the computations were done for 0ºK and the experiments are usually done at about 293ºK.

4. CONCLUSION

In this paper, the results of the DFT-LDA application for computing third-order elastic constants of Si are presented. Although there are rather large discrepancies of computational and experimental data, the data are in good accordance. It is possible to get more precise computations with using more accurate approximations and with taking into account the non-zero temperature of crystal.

REFERENCES

[1] J.Beckley, V.Kalinin, M.Lee, K.Volianski, Non-contact Torque Sensors Based on SAW Resonators, in Proceedings of 2002 IEEE International Frequency Control Symposium and PDA Exhibition, pp.202-213.

[2] X.Gonze, J.-M.Beuken, R.Caracas, F.Detraux, M.Fuchs, G.-M.Rignanese, L.Sindic, M.Verstraete, G.Zerah, F.Jollet, M.Torrent, A.Roy, M.Mikami, Ph.Ghosez, J.-Y.Raty, D.C.Allan, First-principles computation of material properties: the ABINIT software project. Computational Materials Science 25, 478-492 (2002).

[3] R.N.Thurston in Physical Acoustics, edited by W.P. Mason (Academic Press Inc., New York, 1964), Vol. 1A.

[4] The ABINIT code is a common project of the Université Catholique de Louvain, Corning Incorporated, and other contributors (URL http://www.abinit.org).

[5] R.F.S.Hearmon, 'Third-Order' Elastic Coefficients, Acta Cryst. (1953), 6, 331-340.

[6] S. Goedecker, M. Teter, J. Huetter, Phys.Rev.B54, 1703 (1996)

[7] N.Troullier, J.L.Martins, Efficient pseudopotentials for plane-wave calculations, Phys. Rev. B43, 1996-2003 (1991)

[8] H.J. Monkhorst and J.D. Pack, Phys. Rev. B13, 5188 (1976)

[9] R.W.G.Wyckoff, Crystal Structures, 2nd ed. (Interscience Publishers, New York, 1963)

[10] Акустические кристаллы. Справочник/под ред. М.П. Шаскольской. – М.: Наука, 1982.

FERROBIELASTIC BEHAVIOUR OF 32 CLASS CRYSTALS: LGS, LGT and $GaPO_4$

M. ASSOUD[1, 2], J. J .BOY[2, *], K. YAMNI[1], A. ALBIZANE[1]

[1]Laboratoire des Matériaux Catalyse et Environnement, FST Mohammedia BP 146 - 20650 Mohammedia - Maroc
[2]LCEP, ENSMM & FEMTO-ST – 26, Chemin de l'Epitaphe- 25000 Besançon - France
*Email : jjboy@ens2m.fr

ABSTRACT:
It is well known that the quartz crystal is ferrobielastic and that the induced twinning is due to the movement of SiO_4 tetrahedra from one Dauphiné twin state to the other by applying a suitable field of mechanical stresses. The growth of one Dauphiné twin state out of another can be explained by the ferroic theory which has induced classification of ferroic crystals based on the nature of material properties that differs from one domain to the other.
The existence of the two domains in a crystal sample is translated by the existence of boundaries between these domains and modification of the frequency thickness relationship or sometimes disappearance of the piezoelectric activity in the crystal resonator. Furthermore, the stress level needed to complete the volumetric transformation to the alternate twin state is associated with the crystallographic orientation and also with the content of impurity-related point defects randomly distributed in the quartz lattice, without forgetting the temperature which plays an important role on the threshold values.
So, to realize piezoelectric high pressure and/or high temperature sensors in new piezoelectric materials, we have to clarify and to characterize the possible ferrobielastic behaviour of the crystal used to manufacture the resonator.

The aim of this paper is to study the ferrobielastic behaviour of Langasite (LGS, $La_3Ga_5SiO_{14}$) which is a material belonging to the crystal class 32 and Gallium orthophosphate ($GaPO_4$). For that, we have applied uniform stresses on parallelepipedic X-cut samples having a face defined by a third angle ψ, chosen equal to -45°. Our polished samples are prepared with the same orientation and same dimensions that the previous ones, built in different kind of quartz crystals [1]. Associated to theoretical thermodynamic considerations, we describe here the experimental observations made on the stress-strain curves obtained at ambient temperature.

INTRODUCTION:
According to ref [2], ferrobielastic switching stimulated by a uniaxial static pressure in certain specially chosen directions can occur in class 32 crystals, which are ferrobielastic. The transition of a sample as a whole into a twinned state has been observed repeatedly in quartz single crystals by optical or acoustic methods [3, 4]. The switching effect has also been observed in $GaPO_4$ [5] and LGS [6]. This paper reports new observations of switching in these last new materials to compare precisely their properties to those of quartz crystal. Indeed, in the outlook of the realization of pressure or temperature sensors for very severe environments (high pressure and/or high temperature), we could consider to replace the quartz crystal by new piezoelectric materials.
If previous papers aiming the use of quartz crystal for Bulk Acoustic Waves resonators in pressure gauge [7, 8] confirm that the quartz gauges exhibit good long term stabilities and high accuracy up to 100 MPa and 175°C, we assume that hysteresis, then twins and at least fractures can occur at higher pressures and temperatures…

BACKGROUND OF THE FERROBIELASTIC TWINNING:
At the microscopic level, the α-β phase transformation in quartz is a first order transition with a symmetry breaking order parameter described by the Landau theory and, at the microscopic level, it is a displacive transition produced by small atomic displacements of silicon (0.34 Å) and oxygen (0.81 Å) atoms. In the β-phase of point symmetry 622, the atoms are placed in more symmetric positions than in a α-phase of point symmetry 32. When β-quartz is cooled below 573°C, the SiO_4 tetrahedra rotate by an angle of $\approx$ 16° about the twofold axes. Two rotation senses are possible leading to a domain structure known by mineralogists under the name of Dauphiné twins. The domains differ in the sign of the tilting angle and are structurally related by a rotation of 180° about the threefold axis. The atomic shifts needed to go from

one domain to the other involve motions of less than 0.3 Å and require no breaking of Si-O bonds.
According to Aizu classification, quartz is a second order ferroic crystal that belongs to the class 622F32. The ferroic state shifts of 622F32 crystals are accomplished by a combination of a mechanical stress and an electric field or by a mechanical stress alone. The response in the elastic strain is different from one domain to another due to the differences in the piezoelectric and elastic compliances tensors. The spontaneous electric displacement vector and the spontaneous strain tensor are equal to both domains. For these reasons, is simultaneously classified as a ferroelastoelectric and a ferrobielastic crystal. Neglecting high order terms then the second one, the Gibbs free energy difference between two domains, S_1 and S_2, may be determine as follow:

$$G - G_0 = -\frac{1}{2} s_{ijkl} T_{ij} T_{kl} - d_{jk} T_{ij} E_k - \frac{1}{2} \chi_{ij} E_i E_j$$

where E_i and T_{ij} represent the electric field and the mechanical stress tensors, respectively. To evaluate the differences between S_1 and S_2, in the piezoelectric and elastic compliance tensors, it is necessary to apply the symmetry operation relating S_1 and S_2 on each independent piezoelectric and elastic compliance coefficient. Furthermore, if we consider the simple case when $E_i = 0$ and an uniaxial and homogeneous compressive stress (P) is acting along X''_3, an arbitrary chosen direction of a rotated set of axes $X''_1X''_2X''_3$, by the conventional angles φ, θ and ψ, the Gibbs free energy difference between the 2 Dauphiné twin states is given by:

$$\Delta G = 2 \, . \, s_{14} \, . \, F(\varphi,\theta,\psi) \, . \, P^2$$

where only the s_{14} change its sign from S_1 to S_2. If the sign of s_{1123} is assumed positive and if the F function is negative, the ferrobielastic twinning is thermodynamically favorable because the free energy of the alternate twin state S_2 is lower is lower than that of the initial one.
For our experiences, we have chosen to work on rectangular X-cut samples on which the quasi-uniaxial force is applied on a direction defined by a ψ-angle equal to +/- 45°.

PREVIOUS EXPERIMENTS:

1. Quartz samples:

The Figure 1, extracted from [1], illustrates the ferrobielastic switching when a pressure is applied in a X''_3 direction defined by an angle of -45°. We have also shown that the threshold level (called too "coercive stress") differs from one sample to the other, depending on the growth quality (natural or synthetic) and on the content of impurity-related point defects.

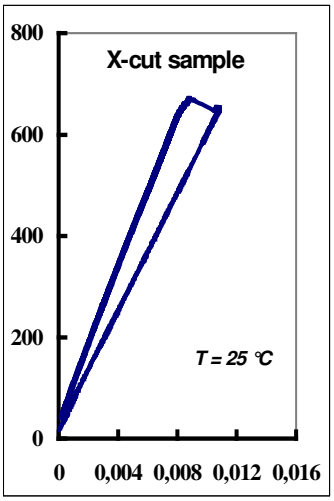


Fig. 1: example of stress-strain curve for a X-cut quartz sample

The figure below, extracted from [9], presents the strong variation of the coercive stress with the temperature up to the α-β phase transformation in which the ferrobielastic switching is close to zero.

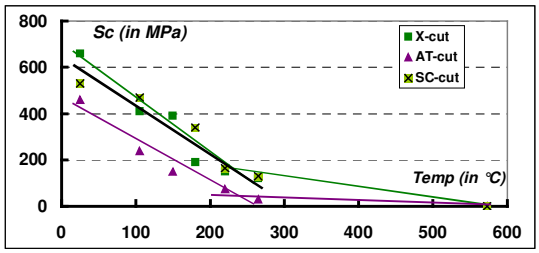


Fig. 2: evolution *with the temperature of the coercive stress for X- and AT-cut samples*

2. $GaPO_4$ samples:

Previous experiments have been made at ambient temperature on X-cut samples of $GaPO_4$ and $AlPO_4$ in 1997 with the same experimental arrangement than this used for quartz. The following picture summarized the stress-strain behavior on which the ferrobielastic switching is clearly observed in quartz (jump in point A). In contrary, the dashed lines shown at the end of the curves 2 and 3 denote that the samples were broken before removing the pressure. Nevertheless, the stress-strain curve of the sample 3 presents little jumps identical to these observed previously for quartz and revealing probably ferrobielastic switching. Furthermore, the development of twin domains was optically registered always on the $GaPO_4$ samples but not on the samples of berlinite.

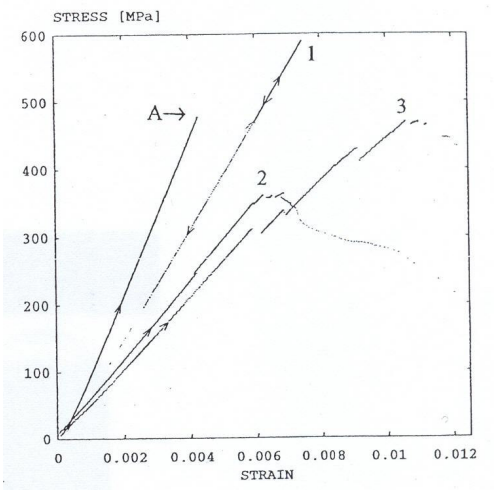


Fig. 3: comparison of stress-strain behavior for quartz (1), $GaPO_4$ (2) and $AlPO_4$ (3).

3. Experiment on LGS:

As the Langasite, its ferrobielastic switching was observed by measuring the change in the frequency of the compound acoustic cavity made from a Langasite single crystal, under uniaxial pressure. The effect was detected from the jump in the cavity frequency at the point of switching.

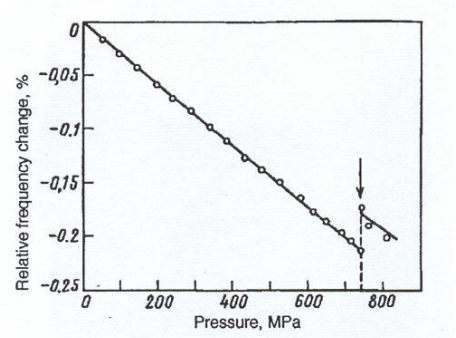


Fig. 4: relative change in the frequency of a compound acoustic cavity vs Papplied. The arrow marks the point of switching

4. A few comments:

If the threshold in pressure for the complete switchover has not been observed for $GaPO_4$ samples, twin boundaries have appeared at pressures near 350 MPa. For quartz, the stripe-shaped domains disappeared at about 500 MPa and an abrupt jump was observed in T-S dependence: it is the value of the coercive stress for quartz. The alone observation made on samples of LGS at ambient temperature gives a coercive stress near 750 MPa, much higher than in quartz. Nevertheless, this value has to be confirm…

CONCLUSION:

New experiments are planed with similar samples cut in each crystal (see fig. 6) of SiO_2, $GaPO_4$ and LGS / LGT. We had to be reviewed the stressing apparatus and so the results will be published in the next forum.

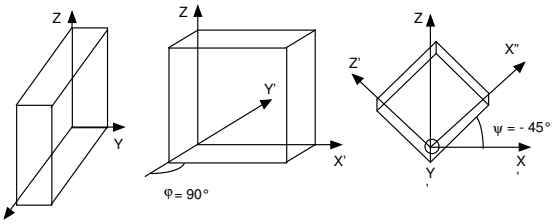


Fig. 6: sample orientation

REFERENCES:

[1]. P.L. Guzzo, J.J. Boy: "The role of crystallographic orientation and crystal origin in ferrobielastic twinning of quartz under uniaxial stress" Eur. Phys. J. AP 12, 17-30 (2000).

[2]. K. Aizu: "Possible species of ferromagnetic, ferroelectric and ferroelastic crystals", Phys. Rev. B 2, 754-772 (1970).

[3]. H.Klapper: "X-ray topography of twinned crystals" Prog. Crystal Growth and charact. 14, 367-401 (1987).

[4]. G.D.Mansfled, A.D. Freik, B. Kutur: "A high overtone BAW resonator as a tool for acoustic characterization of thin films and layers" EFTF, 7, 261-266 (1993).

[5]. G.D. Mansfled, P.L. Guzzo, J.J. Boy, O. Zvereva: "Observation of ferrobielastic twinning in GaPO4 under uniaxial stress" IEEE Trans. On Ultra. Ferro. and Freq Control, vol.44, n°3 722-724 (1997).

[6]. G.D. Mansfled, J.J. Boy: "Observation of ferrobielastic switching in langasite crystals under a uniaxial static pressure" JETP Lett.. vol. 66, n°5 362-364 (1997).

[7]. N. Matsumoto, Y. Sudo, B. Sinha, M. Niwa: "Long-term stability and performance characteristics of crystal quartz gauge at high pressures and temperatures" – IEEE Trans. on UFFC, vol. 47, n°2, 346, 354 (2000).

[8]. E. P. EerNisse, R.W. Ward: "A reduced hysteresis extended range quartz pressure transducers" Annual Symposium on Frequency Control 41, 344-349 (1987).

[9]. K. Yamni, J.J. Boy, A. Yacoubi: "Temperature effect on the ferrobielastic behaviour of the quartz crystal" Ann. Chim. Sci. Mat. 26, 19-22 (2001).

INVESTIGATION OF FREQUENCY STABILITY OF ACTIVE HYDROGEN MASER WITH AUTONOMOUS CAVITY AUTOTUNING USING FREQUENCY SWITCHING METHOD

N. A. Demidov, V.A. Logatchev, V.S. Gorev, V.G. Vorontsov

Federal State Unitary Enterprise
Institute of Electronic Measurements KVARZ (IEM "KVARZ")
176, Gagarin Prospect, Nizhny Novgorod, 603009, Russia
Fax: (8312) 66-55-62; e-mail: nnipi_kvarz@sinn.ru

ABSTRACT

The description of the Active Hydrogen Maser developed by the IEM "KVARZ" with the autonomous cavity autotuning using frequency switching method is given. Frequency stability of $1 \cdot 10^{-13}$/s, $5 \cdot 10^{-16}$/day and temperature frequency coefficient of $8 \cdot 10^{-16}$/°C are achieved.

INTRODUCTION

The IEM "KVARZ" of Nizhny Novgorod has been developing Hydrogen Masers for nearly 40 years. In this time 3 generations of hydrogen masers have been developed. The last development of the IEM "KVARZ" is the Hydrogen Maser with autonomous automatic cavity tuning, using frequency switching method which does not require the second analogous Maser as a reference one. Such a system for the first time was proposed by B. Gaigerov and G. Elkin (Russia) [1] and then was successfully used by H. Peters (USA) [2].

DESIGN OF THE CH1-75A HYDROGEN MASER WITH AUTONOMOUS CAVITY AUTOTUNING

The external view of the CH1-75A Hydrogen Maser with the autonomous cavity tuning is shown in Fig. 1, the physics package schematic – in Fig. 2.



Fig. 1. CH1-75A Hydrogen Maser

The main features of the CH1-75A Hydrogen Maser are the following:
– high-precision Maser cavity thermostats with temperature stability of 10^{-4} ° C;
– efficient and light magnetic cavity shields;
– light and compact vacuum system;
– technology for fluoroplastic coating of the storage bulb providing small frequency offset, small offset drift and high stability of coating parameters with time;
– the Hydrogen Maser electronics (PPL system) having low noise (and as a result low frequency instability contribution) and low temperature coefficient of phase change;
– the autonomous (without the 2nd Maser) cavity autotuning system.

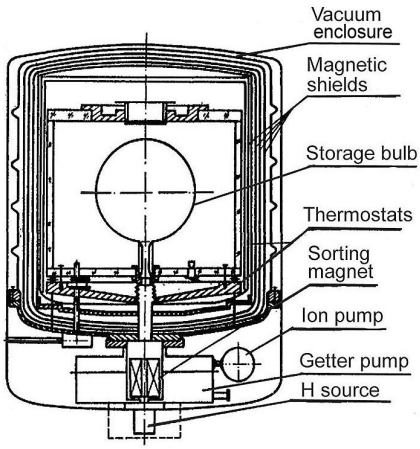


Fig. 2. Hydrogen Maser physics package schematic

A two-stage multizone structure has been chosen in the process of the cavity thermostats development. The whole cavity volume subjected to a temperature control is divided into a number of zones: the lateral cylindrical cavity surface, the upper cover and the lower cover. In addition the vacuum tube connecting the vacuum enclosure, in which the cavity is placed, with atom beam source and the ferrite isolator connecting the Maser output with the microwave signal receiver input are temperature controlled separately. At the same time

the cylindrical surface and the cavity covers have been temperature controlled in two steps. As a result of choosing such a construction, it becomes possible to the achieve temperature stability on the cavity surface (at small gradient) of~10^{-4} °C. The metal shells with the oven heaters windings and the magnetic shields are coated with a thin aluminium layer using a vacuum deposition technique, which provides a low power (~2 W) consumption of the ovens.

The cavity, as in the earlier developed Hydrogen Maser models , is made of sitall – the material having a very low (<$2 \cdot 10^{-7}$/ °C) temperature coefficient of linear expansion. Such a construction of the cavity and the ovens allows to obtain the Maser frequency temperature coefficient of $2 \cdot 10^{-15}$/ ° C.

In order to reduce magnetic shields weight and to increase a shielding factor, the 81 HMA perm alloy having high magnetic susceptibility is used for the shields manufacture. The cavity magnetic shield has four layers of 0.5 mm width each. Three layers are located inside the vacuum enclosure and the fourth layer is outside. The technology of shields annealing in a vacuum furnace has a great importance for obtaining high shield efficiency. All these measures have resulted in achieving the magnetic shielding factor of ~500000, which has led to a weak dependence of the Maser frequency on external magnetic field fluctuations, in particular, the Earth magnetic field (<$1 \cdot 10^{-14}$/ Gauss).

Titanium, light and strong material, having excellent vacuum properties is used as a basic construction material for the vacuum system manufacture. The light getter pumps, in which elements are made of compacted titanium chips, are used for hydrogen pumping out. The small ion pump having a capacity of 2 L/s is utilized for pumping out residual gases. The heater located inside the getter pump allows to activate the pump in vacuum at its initial start-up. The vacuum system of the Hydrogen Maser consists of two separate chambers. In the first chamber, comprising the storage bulb and the beam forming system, the vacuum of ~10^{-8} mm Hg is provided with the help of the getter pump and the small ion pump. In the second chamber, comprising the cavity, ~10^{-3} mm Hg vacuum is provided by the low-capacity ion pump. The lifetime of the getter pump is more than15 years.

The hydrogen used for Hydrogen Maser operation is located in a special vessel as a component of $LaNi_5H_n$ hydride. A partial pressure of saturated hydrogen vapours in the volume, where hydride is located at 50 ° C, is 5 atm. 250 grams of hydride at normal pressure contain 18 litres of hydrogen, which provides the Maser operation for a period of more than 10 years.

The main distinctive feature of the new Maser model is the autonomous cavity autotuning (CAT) system. The operation principle of the CAT system is similar to that described in [1, 2]. The block-diagram of the Maser is shown in Fig. 3.

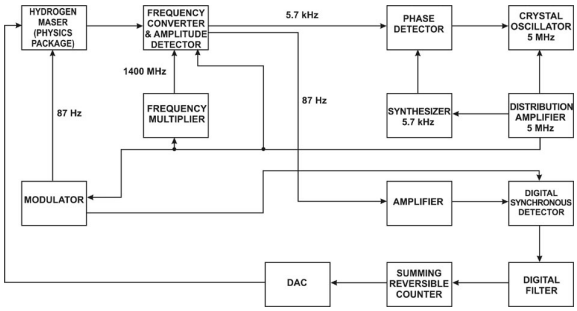


Fig. 3. CH1-75A Hydrogen Maser block diagram

Hydrogen Maser cavity frequency is switched with the help of the modulator that results in amplitude modulation of the Maser generation signal. The amplitude modulated signal is selected by the receiver amplitude detector then is amplified and applied to the synchronous detector. The synchronous detector signal depends on the cavity frequency offset from an atom line (it equals to zero at zero offset). This signal is used for the cavity frequency tuning.

MODULATION FREQUENCY AND CAVITY FREQUENCY DEVIATION SELECTION

The modulation frequency is selected from the condition:

$$\Delta f_l << f_m << \beta, \qquad (1)$$

where f_m– modulation frequency;
β – cavity band half width;
Δf_l – atomic emission line half width.

In our Maser the modulation frequency f_m is selected equal to 87 Hz. The increase of the modulation frequency favorably influences the decrease of lateral components in the Maser output signal spectrum but can cause frequency shifts due to the limited receiver bandwidth.

The cavity frequency deviation (Δf) selected from the max signal value at the receiver amplitude detector output equals to the cavity band half width β, i.e. ≈ 15 kHz. But signal dependence on Δf near optimum value is weak and at $\Delta f = \frac{\beta}{2}$ the signal value is decreased only by 20 %. In our Maser the cavity frequency deviation is

selected equal to 10 kHz.

EXPERIMENTAL RESULTS

The CH1-75A metrological characteristic were experimentally investigated.

The experimental curve of frequency instability dependence of two Hydrogen Masers with autonomous cavity tuning is shown in Fig. 4.

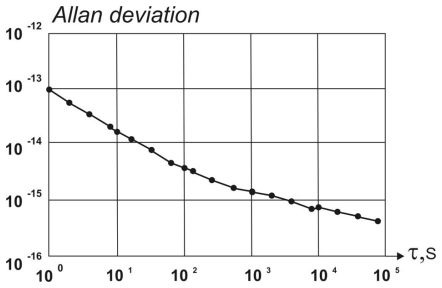


Fig. 4. CH1-75A Allan deviation

The frequency instability equals to $1 \cdot 10^{-13}$/1 s, $5 \cdot 10^{-16}$/day.

The further increase of the Maser frequency stability may be achieved in case of enhancement of an environmental temperature keeping thanks to an air conditioner operation. Periodic switching of air conditioners negatively influences a frequency stability. The best results may be achieved when the Maser will operate in the room where the constant temperature is provided without air conditioners (for example in an underground basement).

The frequency systematic drift at the Maser operation start is $2 \cdot 10^{-15}$/day; after one year of continuous operation the drift usually decreases to $(2–3) \cdot 10^{-16}$/day.

Fig. 5 shows the dependence of the frequency of the Hydrogen Maser with automatic cavity tuning on time (a) and the environment temperature variation on time (b). From that we may define the Maser temperature coefficient of frequency of $8 \cdot 10^{-16}$/ °C.

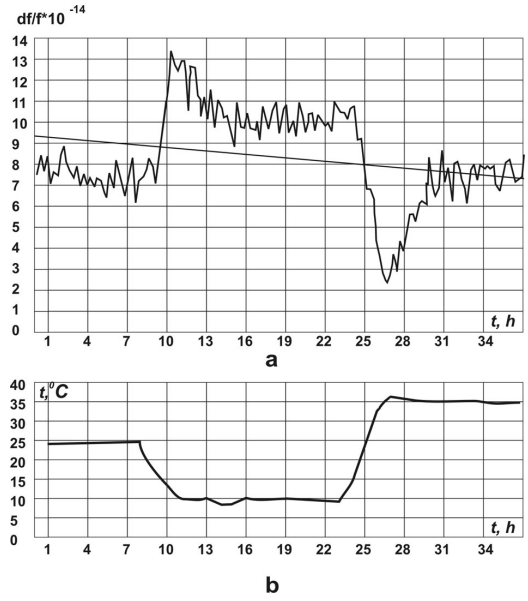


Fig.5. CH1-75A frequency versus time (a) and environmental temperature versus time (b)

Fig. 6 shows spectral power density of the CH1-75A Hydrogen Maser. We can see that the components with modulation frequency of 87 Hz and its harmonics are attenuated in relation to the carrier by more than 140 dB.

The first models of the Masers with autonomous automatic frequency tuning were shipped to laboratories of a number of countries (Canada, Finland, Sweden, Korea and others).

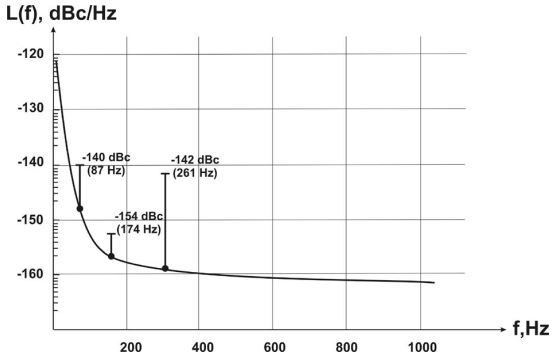


Fig. 6. Spectral power density of CH1-75A Hydrogen Maser

CONCLUSION

So the IEM "KVARZ" has developed the Hydrogen Maser with autonomous automatic cavity tuning having frequency stability $1 \cdot 10^{-13}$/1 s, $5 \cdot 10^{-16}$/day and temperature coefficient of frequency $8 \cdot 10^{-16}$/°C. Further frequency stability increase may be achieved at environmental temperature keeping accuracy increase.

REFERENCES

[1] B. Gaigerov, G. Elkin "Cavity autotuning in hydrogen frequency reference", Measurement equipment, 1968, No 6.

[2] H.E. Reters "Design and performance of new Hydrogen Masers using cavity frequency switching servo", Proc. of the 38th Annual Frequency Control Symposium, USA, 1984, p.p. 420–427.

SIGNAL PROCESSING PRINCIPLE IN A MICROWAVE AUTODYNE SENSOR

Boloznev V.V., Safonova E.V., Sultanov F.I., Stanchenkov M.A.
Kazan State Technical University, Russia
e-mail: ksafonova@yandex.ru, kate@c-tel.net

ABSTRACT

In this paper authors suggest to use the quantity of extracted information in a unit of time as a criterion of sensors' effectiveness for technological processes diagnostics. We are going to present the results of the first stage of investigation, dedicated to proposed index estimation in application to the sensors of combustion process. During our investigations information density, i.e. derivative from required quantity by frequency, has been found. Two main features in measuring signal forming: correlation interconnections of required parameter with noise, transfer characteristic nonlinearity have been considered.

Preliminary estimations of the sensor's perfecting opportunities by circuit technique and design methods have been done.

1. INTRODUCTION

Since the famous work of C.Shannon [1] was published, information capacity step-by-step became an integrating quality index of radio systems. It has displaced traditional indexes such as radio coverage, power consumption, carrier frequency and others. The term (information capacity) itself and its importance became clear even to non-specialists including consumers. It happened mostly because computing operations were applied into signal processing.

Similar occurrences didn't happen in measuring technique and a number of quite successful scientific attempts didn't find support and understanding. Traditional set of errors is still dominates in this field. In a case of technological process control these errors linked up to the current time in an awkward way. Partly, it caused by some features in application of Shannon's information approach to analysis of the measuring system first element - sensor.

Not pretending for complete coverage of the problem, let's list the titled features in application to the microwave autodyne sensor of the system of combustion process monitoring [2]:

- Initial process, measuring signal and noise have nonuniform (nonwhite) spectrum,
- Signal and noise have been generated by one physical process and therefore correlated,
- Sensor and some nodes of measuring system are nonlinear.

Analytical solution of the task is complicated by these circumstances. It doesn't have any principal difficulties, but it is too big for one article. Authors decided to do publications by parts.

Purpose of the first part of our investigation is: estimation of information quantity, extracted by the sensor in a unit of time and in a unit frequency range. In the technical literature [3] this value calls differential entropy difference; it can be total or conditional. We are going to use a term "information density".

2. ELECTROPHYSICAL MODEL OF THE MEASURING PROCESS

Model presented on the figure 1a includes a path of measuring signal forming: controlled technological parameter (CTP) of flame; indirect parameter – electron density N_e; directly

measurable electrophysical parameter – plasma permittivity of flame $\mathcal{E}$; and sensor, transforming $\mathcal{E}$ into oscillation frequency ω, which is a probing frequency also. A number of physicochemical effects, caused by technological process and external factors (1, 2, 3, …T) work as a noise origin. Temperature is a main one; it has a strong influence to the measuring errors.

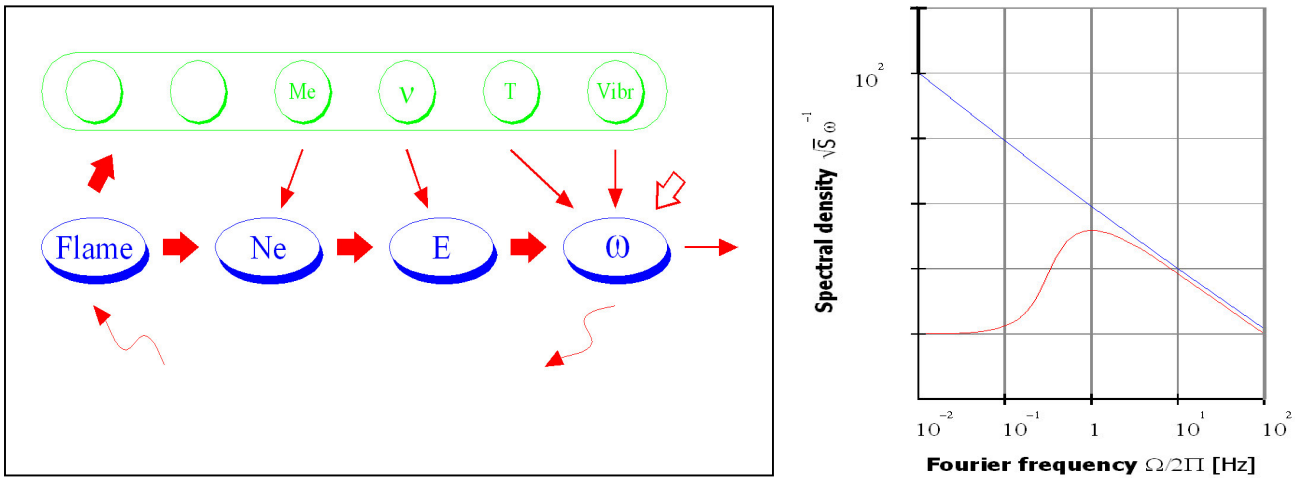


***Figure1** 1a -Physical processes of measuring signal and interference factors forming in transport thermal power engines (TPE). 1b - Spectral densities of temperature fluctuations in flame in natural conditions (curve 1) and in TPE (curve 2).*

Sensor's antenna emits probing signal into half-space of combustion chamber, filled by material medium (plasma of a flame) with following permittivity

$$\varepsilon(t) = 1 - \frac{bN_e(t)}{\omega^2(t) + \nu^2(t)}, \qquad (1)$$

where $b = 80 m^3 Hz^2$ - coefficient, connected with mass and electron charge, ν - electron frequency of collision with neutral particles (unburned fuel, soot, products of combustion chamber erosion). Functional interconnection of N_e with CTP stays outside authors' competence, but it is known for a number of application tasks in electrophysics of combustion.

Transformation $N_e(t) \to \varepsilon(t)$ is linear. Nonlinear and noise components connected with probing frequency ω and collision frequencies ν. Corresponding errors (that cause an information loss from our point of view) are called truncation and implementation errors.

Measuring signal forming, i.e. permittivity $\mathcal{E}$ transformation into frequency ω, and this frequency dependence from the main origin of noise – temperature, has been seriously investigated in [2]:

$$\omega[\varepsilon(t)] \approx \omega_0 \cdot \left[\frac{\Gamma_A(\varepsilon)}{2Q} \sin \varphi(\varepsilon) \right], \qquad (2)$$

where Γ_A, φ - amplitude and phase of antenna's reflectance, Q-resonator's Q-factor. In technical literature similar dependences called - autodyne response. It has nonlinear character and we will present it by polynomial.

Temperature dependence is linear [4]:

$$\delta\omega(T, \Omega) = K_T(\Omega) \cdot \delta T(\Omega), \qquad (3)$$

symbol Ω shows processes and linear operator $K_T(\Omega)$ dependence from the frequency of temperature fluctuations, moreover in comparison with measuring signal spectrum, operator is quite narrowband. In the works [4, 7] this operator, presenting contribution of three factors: materials and

sizes of the sensor's body and also antenna's influence, shown as a sum $K_T(\Omega) = K_p(\Omega) + K_L(\Omega) + K_A(\Omega)$.

Let's take into account also collision frequencies variations, which are an inaccessible for monitoring, and noise corresponded to them. As the result signal structure of an autodyne sensor has a following view:

$$\Delta\omega = \omega[\varepsilon(t), T(t), \nu(t)] - \omega_0 = \Delta\omega + \delta\omega = \{a \cdot \Delta\varepsilon(t) + L[\Delta\varepsilon(t)]\} + \delta\omega_T[T(t)] + \delta\omega_\nu[\nu(t)], \quad (4)$$

where linear and nonlinear signal parts $(L[\varepsilon(t)] = (a_2\Delta\varepsilon^2 + a_3 \cdot \Delta\varepsilon^2 + ...)$ are presented by polynomial in braces, the rest are noise components. Vibrations, which are also the noise origin, are still stay outside our consideration.

3. INFORMATION DENSITY

Let's consider $N_e(t), T(t), \nu(t)$ as the stationary random processes with corresponding mathematical expectations and spectral densities. With influence on the sensor they cause a response (4) with mathematical expectation m_c and spectral density

$$S(\Omega) = S_N(\Omega) + S_T(\Omega) + S_\nu(\Omega) = S_e(\Omega) + S_n(\Omega), \quad (5)$$

where index "n" represents noise.

There is an interconnection between electron density N_e and temperature T. For its description we present the noise component in (4) as

$$\delta\omega_n = r(\Omega) \cdot \delta\omega_s + K_T(\Omega) \cdot \delta T + K_\nu(\Omega) \cdot \delta\nu = r(\Omega) \cdot \delta\omega_s + \delta\omega_H, \quad (6)$$

i.e. in a temperature noise we extract a component, which is linearly connected with a useful signal. Where index "s" represents signal. $r(\Omega) = 2q(\Omega)\sqrt{S_T/S_N}$, $q(\Omega)$ - correlation coefficient. Now spectrum of signal and noise mixture has a following view:

$$S(\Omega) = S_N(\Omega) \cdot [1 + q^2(\Omega)] + S_H(\Omega) = S_s(\Omega) + S_H(\Omega), \quad (7)$$

moreover $S_H(\Omega)$ is a spectral density of independent noise.

According to the posed purpose and follow the Shannon methods, supplemented in [5], let's do following operations:

- signal normalization (4) with a scale $\delta_H(\Omega) = \sqrt{2\pi S_H(\Omega)/e}$, e - natural logarithmic base [5]. Let's present scaled signal as $x(t) = y_s(t) + z_H(t)$;
- finding of distribution density $W(y)$ and $W(x)$;
- finding of conditional density $W(x|y)$;
- determination of differential entropies $H(x)$ and $H(x|y)$;
- determination of information density as differential entropies difference.

Determinations were made in a first approximation by nonlinear parameter.

Let's present an information density value in two forms:

$$I_1(\Delta\omega, \Omega) \approx 0{,}5\left\{ \log_2\left(1 + \frac{S_s}{S_H}\right) - æ\left[\log_2 \frac{2\pi}{e}(S_s + S_H) + \frac{1}{2}\log_2 \frac{2\pi}{e} S_H\right] \right\}$$

$$I_2(x, \Omega) \approx 0{,}5(1 - æ)\log_2\{1 + R_s(\Omega)[1 + r^2(\Omega)]\}, \quad (8)$$

where $æ = a_2 m_s / a$, m_s - mathematical expectation of the signal, $R_s(\Omega) = S_N(\Omega)/\sigma_H$. Both forms are completely equivalent each other, first form (without scaling) is more convenient for the physical interpretation, second one (after scaling) - for calculations.

On the figure 2a, 3a, 4a an approximate view of the signal spectral densities $S_s(\Omega)$ and noise have been presented as a sum of $S_T(\Omega)$, $S_v(\Omega)$ and $S_n(\Omega)$. $S_n(\Omega)$- is a well known heat noise (on the figures it is shown as a shaded area), which is, of course, not relating to the investigated heating. envelope curve form $S = A/\Omega$, which is usual for the natural and man-caused processes, in application to the $S_T(\Omega)$ is disfigured by two factors: low-frequency fluctuations ($\leq 1Hz$) suppressed by TPE control system, high-frequency fluctuations ($\geq 1Hz$) suppressed by the temperature waves' attenuation in the sensor's body [4]. Quantitative values here stipulated by temperature conductivity, linear expansion coefficient and antenna's construction (mono- and bimetallic).

Dependences $I_1(x,\Omega)$ have been given for the several modifications of the sensor construction.

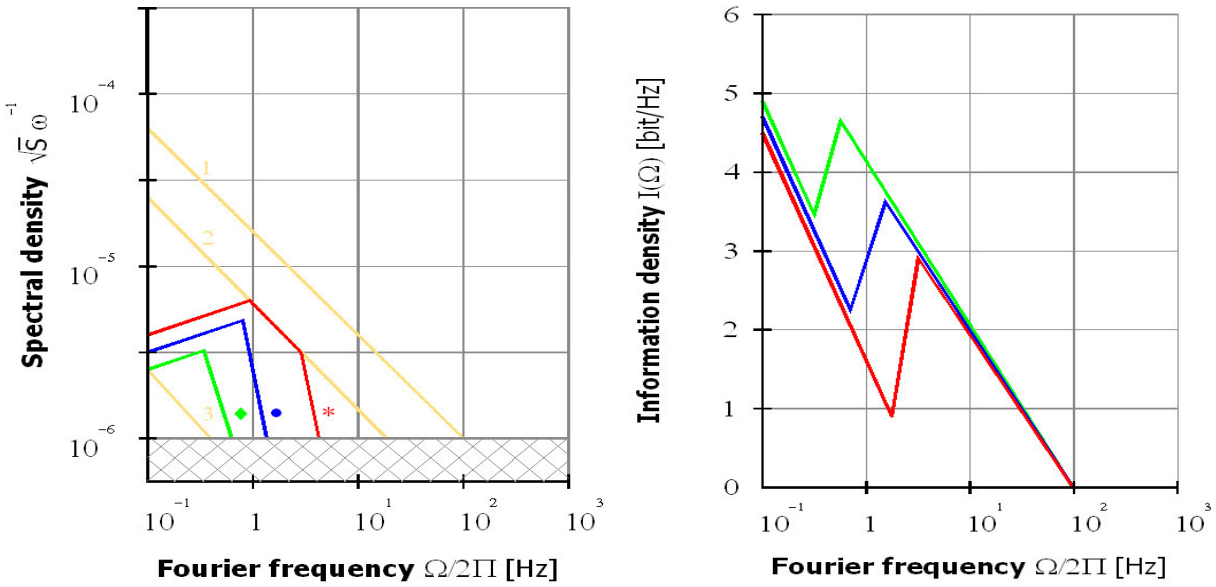


***Figure3** Spectral and information densities of the signal (line 1), collision interference (line 3) and temperature interference in sensors made from different materials: titanium (∗), melted (•) and porous (◊) quartzes. Shaded area is a spectral density of the heat noise.*

4. DISCUSSION

First summand in $I_1(\Delta\omega,\Omega)$ (8) is a linear approximation, which coincides with the Shannon's result. Information loss is presented by the second term with coefficient æ. Its redistribution in spectrum depends on tree factors. Their result quantitatively depends upon the form of initial signal and noise spectrums.

Two of these factors - nonlinearity a_2/a and working point (m_s) shift are the terms of coefficient æ. They have an effect on result of the signal and noise multiplication (i.e. on the logarithms sum) and cause the growth of the last one. Information loss caused by this fact is irreversible, to be more precise, its partial restoration needs complicated reprocessing.

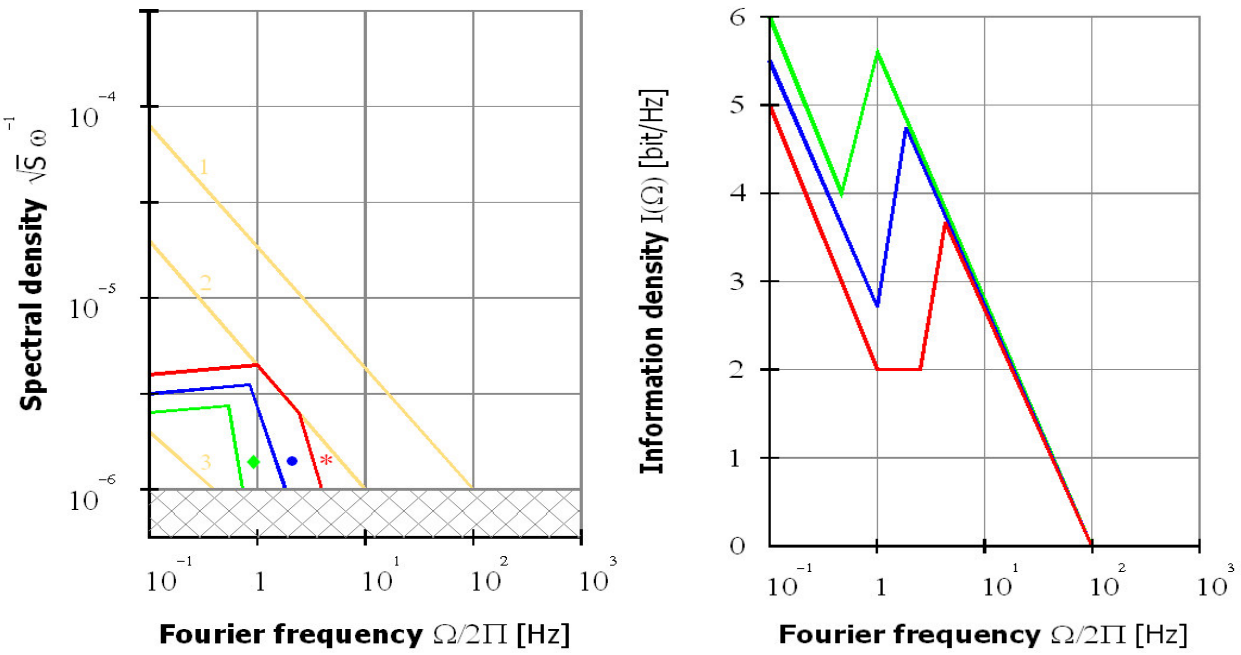


***Figure 4** Spectral and information densities of the signal (1), collision interference (3) and temperature interference in multi-link sensors with a different length L=33mm (∗), L=66mm (line •), L=100mm (line ◊). Shaded area is a spectral density of the heat noise.*

Third factor in expression (8) is not directly seeable, it is a term $æ \cdot S_s^2(\Omega)$. Its absence indicates that enclosed information has not been lost, but has been shifted into different spectrum area, extended by nonlinear transformation. It is not too much complicated to restore this information (i.e. to give it an initial structure). It is often does in practice, but under the condition of measuring path band extending. Consequence of this is a heat noise band extension.

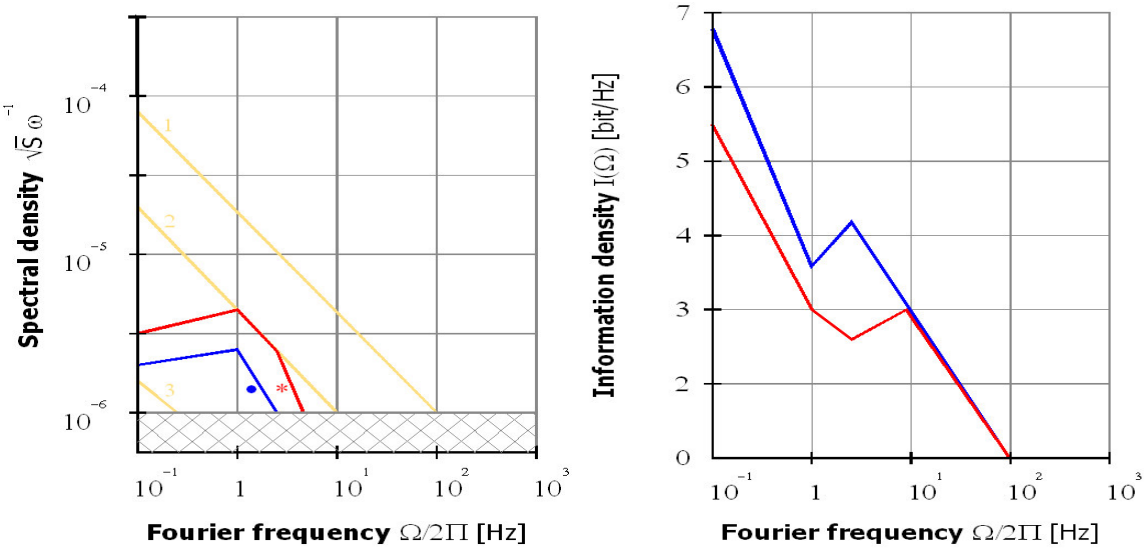


***Figure 5** Spectral and information densities of the signal (1), collision interference (3) and temperature interference (2) in the sensors with monometallic (∗) and bimetallic (•) antennas. Shaded area is a spectral density of the heat noise.*

On the figures 2, 3, 4 is shown, that in a frequency band of several tens of hertz information density reaches (4…7) bit/Hz. Power engineering specialists are interested exactly in this band, because by hypothesis, first signs of emergency mode is appear here. Received result proves that accurate information about these kinds of emergency mode changes can be detected by the sensor in tens milliseconds.

Now let's estimate an opportunities for information density increase. In parallel with well known methods (autodyne response growth, transfer characteristic linearization) some design measures could be very effective here: sensor's elongation [2] and using materials with low thermal conductivity, for example, porous glasses and ceramics [6] in its construction. Corresponding changes in spectral densities and information densities have been shown on figures 2, 3, 4. However it is not possible to reach complete suppression of the fluctuating heating consequences without temperature-compensated antenna using. This antenna should be made of materials with a different thermal expansion.

We are not going to give any quantitative estimation before we figure out, how spectrums deform under nonlinear transformation and before calculating an information density. This material hopefully will appear in future publications.

5. CONCLUSION

In this paper we've presented the first part of our investigation. It is aimed at information opportunities discovering in control systems of combustion modes in TPE, using autodyne sensors. By the quantity of extracted information, and therefore, by accuracy and operation speed these sensors are perspective, especially for the accidents prevention.

Has been noted an essential increase of information volume in comparison with previous constructions of the sensors [8] because of new circuit technology methods (multi-link resonator application) and design measures such as dielectric body and bimetallic antenna.

REFERENCE

1) C.E.Shannon " A mathematical Theory of Communication", The Bell System Technical Jornal, Vol.27, pp.379-423, 623-66, July, October 1948.
2) V.V. Boloznev, E.V. Safonova "Diode microwave active oscillator with a multi-link resonator", Proc. International Symposium on acoustoelectronics, frequency control and signal generation, pp. 139-143, St.Petersburg, Russia, 1998.
3) E.S. Ventsel "Theory of probability", Moscow, High School, 2002.
4) V.Boloznev, E.Safonova, "Microwave oscillator as a sensor of electrophysics flow parameters", Proc. EFTF 13th, pp.1062-1066, Besancon, France, April 1999.
5) V.V. Boloznev, N.Z. Safiullin, "Information capacity of the light-nonlinear measuring system" Radioelectronic devices and systems, inter-institute proceedings, pp.87-94, Kazan, Russia, 1996.
6) E.V. Safonova "Microwave generator as a sensor of electrophysics parameters", Proc. III All-Russian scientific conference: distant radiophysics methods in applications to the natural environment investigations, pp.199-200, Murom, Russia, 1999.
7) E.Safonova "Minimization of an autodyne sensor's error, stipulated by the antenna's heating", Proc. EFTF 14th, pp. 241-245, Turin, Italy, March 2000.
8) Safonova E., Boloznev V. "Autodyne microwave sensor of combustion process: behavior under fast heating", proc. EFTF15th, pp.88, Neuchatel, Switzerland, 2001.

FREQUENCY CALIBRATION TESTS OF NEW INDUSTRIAL PASSIVE HYDROGEN MASER FREQUENCY AND TIME STANDARDS VCH-1006

A. Belyaev, S. Medvedev, Y. Pavlenko, B. Sakharov, P. Smirnov,
"Vremya-CH", N. Novgorod, Russia
M. Bloch, O. Mancini "Frequency Electronics Inc.", New York, USA

Russia, 603105, Nizhny Novgorod, Osharskaya street 67, "VREMYA-CH" JS Company,
Tel.- Fax 007 (8312) 35 42 94, E-mail: admin@vremya-ch.com

ABSTRACT

Frequency calibration tests of Passive hydrogen maser frequency and time standard VCH-1006 (produced by Russian company Vremya-CH) were performed in the time period from Aug.2003 till March 2004 by request of Frequency Electronic Inc.
Frequency stability tests and long term frequency behavior at different environmental conditions have been performed by U.S. Naval Research Laboratory Washington DC, (Dr. J. White) and by National Institute of Standards and Technology, Boulder, CO, (Dr. Thomas E. Parker).
Two ordinary industrial devices (serial numbers 020 03 and 022 03) have been tested both in stable environmental conditions and at changing operating temperature, humidity and external magnetic field.

SHORT TERM STABILITY

Figure 1 shows the results of short term stability measurement (Allan deviation) made by NRL. The 5 MHz outputs were measured against the NRL maser N1 using a Timing Solutions model 5110.

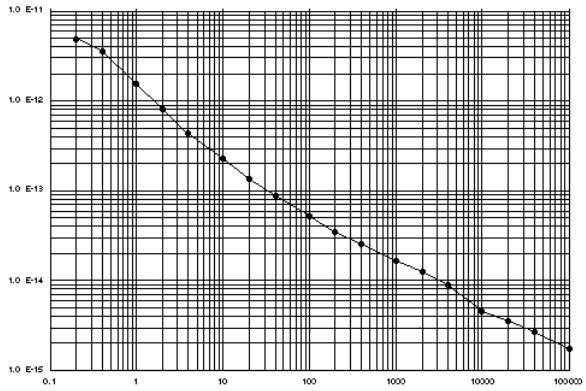


Figure 1. Short term stability (NRL)

The data for averaging times beyond 40 seconds has a slope of about $5 \times 10^{-13}/\sqrt{\tau}$. The data at less than 40 seconds shows a steeper slope. This may be due to the measuring equipment.
In NIST during all of tests the 5 MHz output from the VCH-1006 was compared to a 5 MHz signal from AT1, which is the NIST free running time scale that is based primarily on an ensemble of active hydrogen masers. The stability of AT1 is better than the instrument under test for all averaging times at or above 2 seconds. Below 2 seconds the NIST reference may corrupt the measurements.
Figure 2 shows the Allan deviation as a function of averaging time (τ) in seconds for the range of τ from 0.01 seconds to 4×10^6 seconds. The data represented by hollow circles were obtained with the TSC 5110A Time Interval Analyzer during the last six days of testing in NIST. Similar results were obtained in August when the maser first arrived. The data represented by diamonds and dots were obtained over the 95 days the maser was in the NIST time scale measurement system. The data recording interval here is 2 hours. The diamonds show the Allan deviation with frequency drift included and the dots show the stability with drift removed. Confidence limits are shown by the error bars. With both the TSC 5110A data and time scale measurements the last few points were calculated using the TOTAL deviation. Below $\tau \sim 2$ seconds the NIST frequency reference may corrupt the measurements by a small amount.

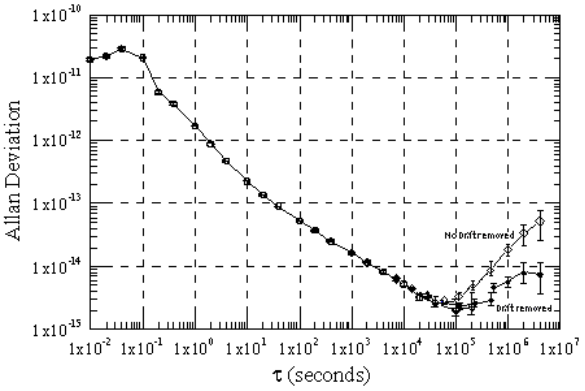


Figure 2. Short term stability (NIST)

Test results show excellent maser's frequency stability performance:
Allan variance is about $5 \times 10^{-13}/\sqrt{\tau}$ for time average τ from 10 up to 10000 second and $(2 - 3) \times 10^{-15}$ per day.

LONG TERM STABILITY

Long term stability measurements were made both by NRL and NIST.
5 MHz outputs against the NRL maser N1 using a Timing Solutions model 5110 data were collected at 5 minute interval on NRL's dual mixer data system. As shown in Figure 3, the residual phase after removal of phase and frequency offsets indicates the presence of a low level of frequency drift.

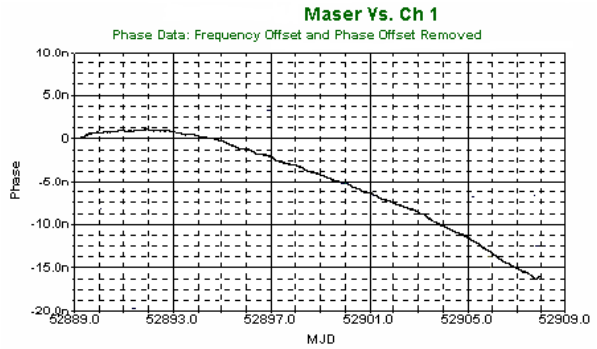


Figure 3. Phase plot (NRL)

The calculated frequency drift rate is about $1x10^{-15}$. At that level, it could be attributed to either the test clock or the reference clock. Removing the measured drift produces the residuals shown in Figure 4. The residuals have a peak-peak span of less than 2 nanoseconds over the data period of 19 days.

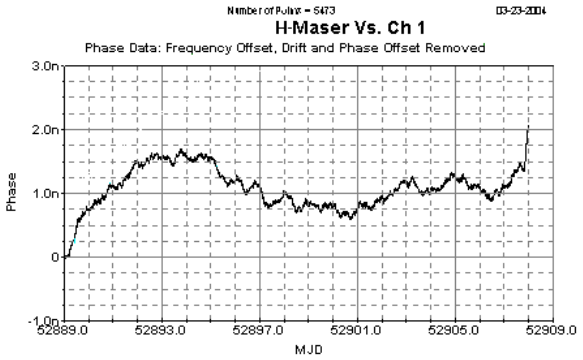


Figure 4. Phase residuals (NRL)

Calculating the Allan Deviation on this span of data produces the results shown in Figure 5. As in Figure 1, there is a distinct break in the slope of the curve. However, in this case, the break is at around five thousand seconds and is the point where the measurement system noise curve meets the clock noise curve.
Figure 6 shows the fractional frequency shift relative to AT1 (NIST) that occurred after October 3, 2003 when the standard was place in one of the environmental chambers used for clocks in our time scale ensemble.

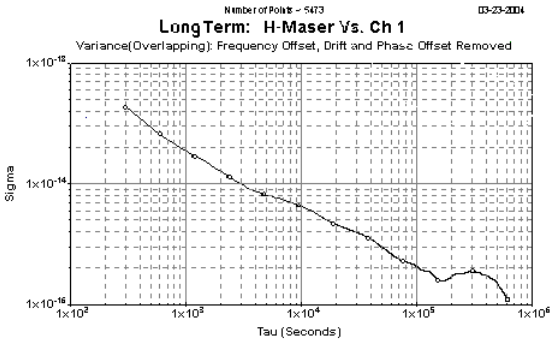


Figure 5. Allan deviation with removed drift (NRL)

In this chamber the temperature and relative humidity were held constant at 23°C and 45% respectively. The maser clearly shows a nonlinear frequency drift, however it was becoming more linear the longer it ran.

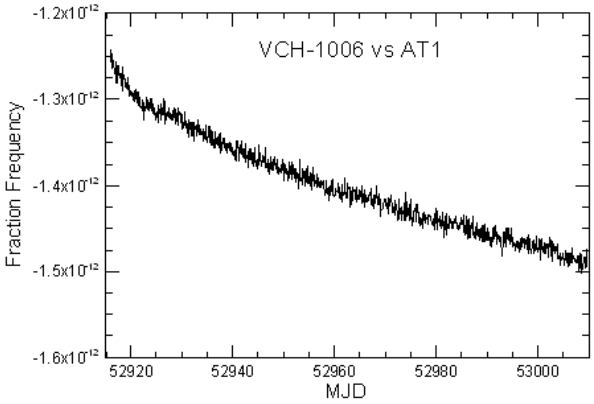


Figure 6. Fractional frequency drift with a constant temperature and humidity (NIST)

The average drift over this entire period was $-2.1x10^{-15}$ per day. During this period the frequency drift of AT1 was less than $5x10^{-17}$ per day.
Apparently from the given results, the device shows fine characteristics:
average frequency drift is less than $(1 – 2) \times 10^{-15}$ a day in a 3 months time interval.

ENVIRONMENTAL EFFECTS

The measurements of Magnetic Field Susceptibility were made in NRL. The maser was placed in Helmholtz coil to determine its sensitivity to external magnetic fields. The maser was aligned as shown in Figure 7.

Magnetic field testing was done by applying a field in a single direction and then measuring the frequency of the maser over a period of at least 24 hours. The strength of the applied external field was determined by

with the clock out of the coils. Table 1 shows the measurement sequence. A value of + or – 1 indicates a 1 Gauss field applied in either the positive or negative direction.

Date	Time	X	Y	Z	X	Y	Z
		applied			measured		
10/9/03	12:00	0	0	0			
10/14/03	09:00	+1	0	0			
10/15/03	09:00	-1	0	0	1.22	-0.12	0.22
10/17/03	14:37	0	-1	0	-0.48	-1.99	0.22
10/20/03	08:56	0	+1	0	1.67	2.43	0.52
10/24/03	08:11	0	0	-1	0.00	-0.72	-1.67
10/27/03	11:26	0	0	+1	-2.08	1.16	2.29
10/29/03	08:49	0	0	0	-1.09	0.25	0.37
10/31/03	12:42	0	+1	0	-1.67	2.43	0.51
11/3/03	11:06	0	-1	0	-0.48	-1.07	0.23
11/4/03	08:11	0	+1	0	-1.68	2.49	0.51
11/7/03	07:39	0	-1	0	-0.50	-1.90	0.23
11/10/03	07:57	0	+1	0	-1.67	2.44	0.51

Table 1. Magnetic field test sequence

The influence of external magnetic fields on maser output frequency is shown on Figure 8.

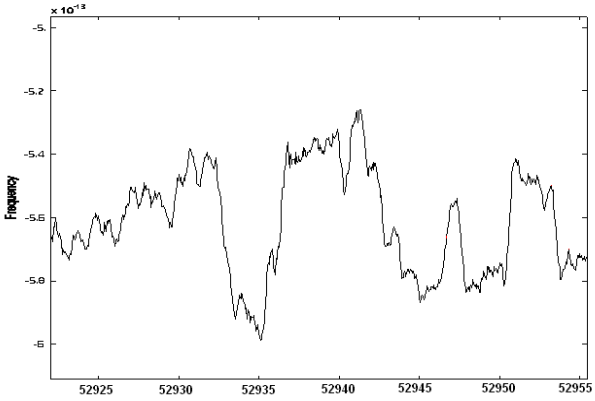


Figure 8. Magnetic field test data (NRL)

The effects in the x and z directions appear to very small and hard to quantify. However, the y direction shows a sharp response to the change in fields, moving about $3x10^{-14}$ for a two Gauss change, see Figure 9.

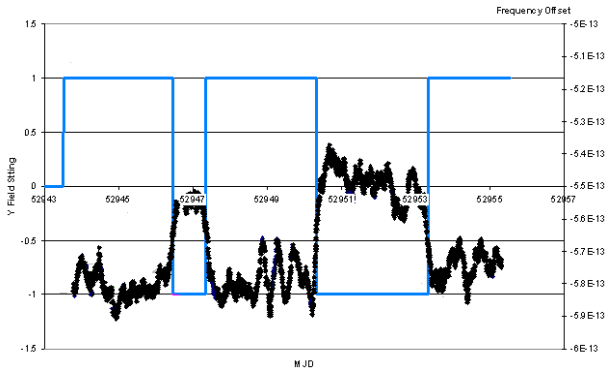


Figure 9. Y axis magnetic sensitivity (NRL)

The y direction result is unexpected. Typically, a hydrogen maser will show the greatest sensitivity along the axis of the shields. In this case, that would be the z direction. Similarly, one would expect that the sensitivity normal to the shields would the same. Here the y direction shows a sharp response but there is no corresponding effect seen in the x axis. The apparent curvature in response that appears in the graph is an artifact of the data averaging. The frequency points are done as a 12 hour running average.

It's important to note that even at this level of sensitivity, this clock is still relatively insensitive to external fields. For example, a commercial cesium clock is specified at $\pm 2x10^{-13}$ for the same two gauss field.

TEMPERATURE SENSITIVITY

The maser temperature sensitivity measurements were made by NIST.

Figure 10 shows the fractional frequency changes when temperature steps were applied. A frequency drift of $-4.0x10^{-15}$ per day has been removed for clarity, and an arbitrary frequency offset has also been removed. The temperature steps were applied in 5°C increments and ranged from 15°C to 30°C as shown in the figure. The air temperature in the chamber (shown in the figure) changed in only a few minutes, though it obviously took the standard much longer to come to equilibrium.

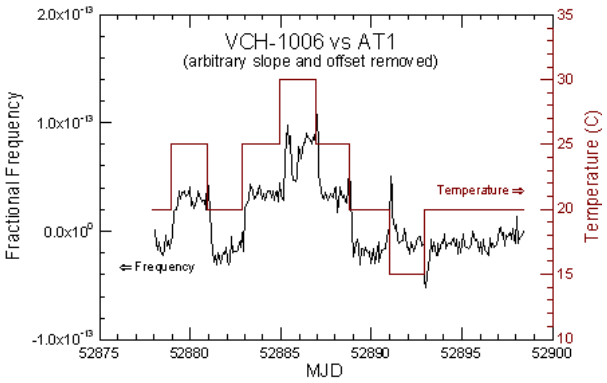


Figure 10. Fractional frequency changes caused by temperature steps (NIST)

Each temperature was maintained for approximately 2 days. There is evidence for both static and dynamic temperature coefficients.

Between 20 and 25°C the static temperature coefficient was found to be $+1.0(\pm0.1)\times10^{-14}$/°C. From 25 and 30°C it was also close to $+1.0\times10^{-14}$/°C, though the frequency was somewhat unstable at the higher temperature. Between 15°C and 20°C the static coefficient was much smaller at about $-1(\pm1)\times10^{-15}$/°C. For these temperature steps the dynamic temperature effect is very visible (frequency transients of -3 and $+6\times10^{-14}$) since the static effect is so small. A dynamic effect is probably also present for the other temperature ranges but the static effect hides it. The apparent time delay between the temperature steps and the frequency steps for the 20, 25, and 30 °C ranges is consistent with the dynamic effect observed for the 15°C to 20°C step. In other words the dynamic effect counteracts the static effect for an hour or two. In all of these tests the relative humidity was held constant at 25%, except for the 15°C temperature where the chamber could only hold 30%.

HUMIDITY SENSITIVITY

The maser humidity sensitivity measurements also were made by NIST.

Figure 11 shows the affect of relative humidity steps of 40% (25 % to 65% and back) with the temperature held constant at 20°C. The elevated humidity level was held for about 8 days. The frequency drift and offset have been removed in this figure.

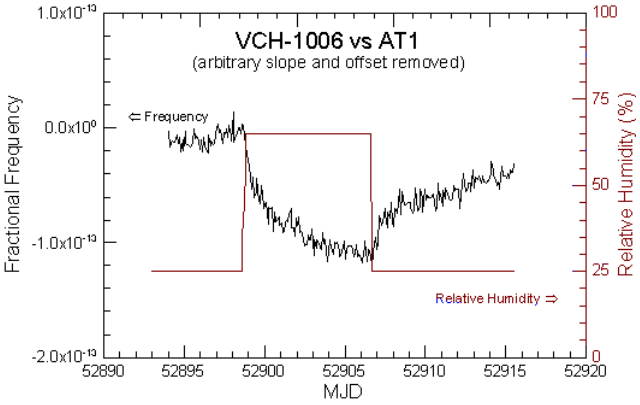


Figure 11. Fractional frequency changes caused by steps in relative humidity (NIST)

The frequency of the maser is clearly sensitive to relative humidity, but with a long time constant. After 8 days the frequency had still not completely stabilized.

There may in fact be two different time constants as it appears that a relatively fast frequency change occurs immediately after the step, and then a slower process seems to set in. The fast response could possibly be a temperature effect due to a change in convection currents, whereas the longer response may be a real humidity effect. Confirming the presence of two different time constants would require a more intensive investigation. In any case a total frequency shift of about 7 to 10×10^{-14} occured with a 40% change in relative humidity. Note that the step up in humidity occurred over a period of 6 to 7 hours, while the step down in humidity took about 3 hours.

CONCLUSION

Test results show excellent maser's frequency stability performance:

- Allan variance is about $5 \times 10^{-13}/\sqrt{\tau}$ for time average τ from 10 up to 10000 second and $(2-3) \times 10^{-15}$ per day
- average frequency drift is less than $(1-2) \times 10^{-15}$ a day in a 3 months time interval.

Environmental effect tests also give quite good results:

- temperature frequency coefficient value is less than 1×10^{-14} / ^{O}C
- magnetic sensitivity is less than $1{,}5 \times 10^{-14}$ / Gauss.

Nowadays VCH-1006 H-maser is successfully used as a local frequency reference at Frequency Electronic Inc., and the company is planning to distribute in future as a precision frequency standard in USA.

Final test results fully confirm the previous ones performed by Neuchatel Observatory (2002) and Russian Metrology Test Institute (2002), as well as Vremya-CH instrument specification and show a good perspective to its improvement.

THE EFFICIENCY INVESTIGATION OF ^{87}Rb ATOMIC BEAM LASER PUMPING FOR DESIGNING A QUANTUM DISCRIMINATOR FOR HIGH-PERFORMANCE SPACE-BORNE ATOMIC BEAM FREQUENCY STANDARD

A. Besedina, A. Gevorkyan, and V. Zholnerov

"The Russian Institute of Radionavigation and Time" PJSC, 2 Rastrelli sq., St.-Petersburg 191124, RUSSIA
E-mail: office @ rirt.ru

1. ABSTRACT

The theoretical study of optical pumping and detection efficiency in the ^{87}Rb atomic beam is performed. Using "rate equations" approximation, the systems of kinetic equations are solved for populations of eight sublevels of the ^{87}Rb atom ground state in a cycle of one- and two-frequency hyperfine pumping and the results as regards the calculation of pumping efficiency for all schemes are given. A number of fluorescent photons of ^{87}Rb atoms is calculated when exciting a beam at hyperfine transitions of D lines of various polarization. The basis of correctness for the adopted model of calculation is given and the borders are shown for practical applicability of results.
The efficiencies of main schemes for pumping and detection in ^{87}Rb and ^{133}Cs beam are compared. The advantages of using ^{87}Rb as a working substance for both process are shown.

2. INTRODUCTION

The ^{87}Rb atomic beam frequency standard with laser excitation is an interesting alternative for the conventional ^{133}Cs frequency standards with optical pumping. It is caused both by a simpler structure of ^{87}Rb atom energy levels and by a possibility to use an atomic beam of larger cross section and density at its high collimation and significantly lower effect of spin-exchange interactions. In connection with it, the theoretical study of optical pumping/detection efficiency in the ^{87}Rb atomic beam is topical.
The objective of studies is to determine the optimal variants of pumping/detection in the ^{87}Rb atomic beam with respect to the discriminator, thus allowing to make a rational and substantiate selection of a principal block diagram for the Rubidium frequency standard with laser pumping. The criteria for selecting optimal, the most efficient pumping/detection schemes for creating the rubidium discriminator are as follows:

a) maximum difference between populations of the "0-0" clock transition sublevels, $\Delta n_{o\text{-}o}$, achieved when using a selected optical pumping scheme;
b) maximum fluorescent photon yield, β_{max}, when detecting a clock resonance signal at the selected hyperfine transition for probing;
c) maximum signal-to-noise ratio when detecting a magnetic resonance signal with a selected principal block diagram of the Rubidium discriminator.

3. OPTICAL PUMPING

Using a formalism of "rate equations", the calculation set is performed concerning evolution in time of Zeeman sublevels' populations of ^{87}Rb atom ground state in a cycle of one- and two-laser hyperfine pumping. Composed and numerically solved are kinetic equations set for 64 possible (according to the selection rules) pumping schemes, distinguished by selected hyperfine transitions of the resonance D - lines and polarization of pumping light as a function of unified parameter, τ_{eff}, taking into account intensity and spectral width of the laser line, absorption line width in the atomic beam, degree of emission and absorption contours overlapping and time of beam atoms interaction with light-wave. When composing systems of equations, such processes were taken into account as all those permitted by selection rules processes of absorption and spontaneous emittance between Zeeman's sublevels of hyperfine structure of $5\ ^2S_{1/2}$ and $5\ ^2P_{1/2;\ 3/2}$ states (Fig.1).

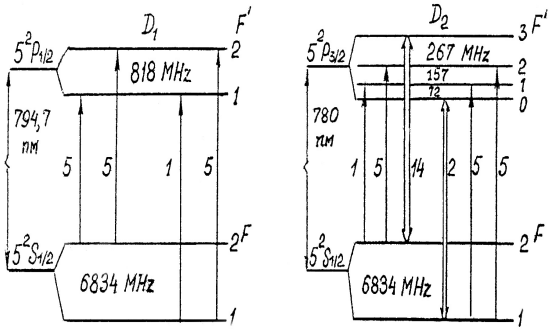


Fig. 1. Energetic levels of ^{87}Rb atom and the relative intensities of the hyperfine transitions

The systems of equations for beam pumping look like

$$dn_k(t)/dt = \Sigma_J \Sigma_{Fs} \Sigma_\mu \; P_{\mu k}^{JFs}\, n_\mu - n_k \cdot \Sigma_J \Sigma_{Fs} \Sigma_\mu \; P_{k\mu}^{JFs} \qquad (1)$$
$$n_k(0) = n_0/[2(2I+1)]$$

Here n_k, n_μ – populations of sublevels of the $|\, F_s, m_{Fs} >$ ground state with indices $k, \mu = 1, 2 \ldots 8$;
$P_{\mu k}^{JFs}$ – rate of pumping at subsequent absorption and re-emittance of I(ν) photons of given polarization.

$$P_{\mu k}^{JFs} = S_\mu^{JFs} \cdot \Sigma_F \rho_{\mu m}^{JFs} \,|\langle m_F | \mathbf{e}_\lambda \mathbf{d} | \mu \rangle|^2 \cdot |\langle m_F | \mathbf{d} | k \rangle|^2, \qquad (2)$$

where $S_\mu^{JFs} = \pi r_o c (f_J/N_J) \cdot (\Delta\nu_L/\Delta\nu_{abs}) \cdot I^{JFs}(\nu_0)$ – a number of photons being absorbed by the atom over the time unit when $\rho = 1$; $I^{JFs}(\nu_0)$ – a spectral density in the center of the hyperfine component of the radiation line, phot/(cm^2·sec·Hz); $r_o c = 8.45\times10^{-3}$ cm^2/sec; f_J - oscillator force; N_J –number of excited state's sublevels; $\mathbf{e}_\lambda$ - vector of light wave's polarization; **d** – operator of atomic electrical dipole moment; J, F, m_F and F_s, m_{Fs} - quantum numbers of the excited and ground states; $< m_F | \mathbf{d} | k >$ - angular part of atomic dipole moment ($N_J^{-1} \Sigma_k \Sigma_F | < m_F | \mathbf{d} | k> |^2 = 1$); ρ_F^{JFs} - coefficient characterizing an overlapping degree of radiation and absorption line's contours when pumping atomic beam by the resonance light; $\Delta\nu_L$ - width of radiation line's contour; $\Delta\nu_{abs}$ - width of absorption line's contour in the atomic beam.

The interaction parameter, τ_{eff}, is numerically equal to the number of photons absorbed by the atom over t time:

$$\tau_{eff} = 2\,(\pi \ln 2)^{1/2} r_0 c\, f_J\, L/N_J \cdot I_0 \rho\, t/\Delta\nu_{abs} = S_\mu^{JFs} \cdot t \approx$$
$$\approx 10^{-3} I_0 \rho t/\Delta\nu_{abs}, \qquad I_0\ [\text{phot}/(\text{s}\cdot\text{cm}^2)] \qquad (3)$$

For finding $P_{\mu k}^{JFs}$ - values, using decomposition of the atomic wave function by functions of the week field, such parameters were calculated as matrix elements of dipole moment and probabilities of electrical dipole transitions between $5\,^2S_{1/2}$ and $5\,^2S_{1/2;\,3/2}$ states.

The obtained results should be treated as maximum possible for ^{87}Rb atoms. Composed kinetic equations and obtained coefficients' matrix are the kernel for getting more precise solutions taking into account the coherences of various kinds, effects of adjacent transitions and laser line's width, interrelation between intensity and excitation line's width.

3.1 One-frequency Pumping

The relative differences of $|\Delta n|/n_0$ populations of |1,0> and |2,0> clock transition sublevels of Rb atom for all one-frequency pumping schemes by hyperfine components of D_1 and D_2 lines with the various polarization are presented in Fig. 2, 3 (σ – linearly polarized emission in which $\mathbf{e}_\lambda \perp \mathbf{B}$; π - linearly polarized emission in which $\mathbf{e}_\lambda \parallel \mathbf{B}$). The **B** vector (quantization axis) is direction of a weak magnetic field, B, created within optical regions of atomic beam's excitation for destruction Zee-man's coherence effects.

The maximum pumping efficiency for |1,0> and |2,0> sublevels achieves 41.3 % when exciting at 2-2' **σ** (D_1 or D_2) transition (in comparison with 15.5 % for ^{133}Cs atoms at 4-4'**σ,** D_2 transition). Assume that $\rho = 1$, $\Delta\nu_L = \Delta\nu_{abs}$, $t_{tr} = 3\cdot10^{-5}$ s, $\tau_{ef} = 70$, then we obtain $I_0(\nu_0) \approx 2\cdot10^9$ phot/(s·cm^2·Hz). Pumping schemes 2-2'σ, π (D_2 or D_1) are distinguished by a short pumping time ($\tau_{eff} \sim 30$). Practically, it corresponds to the excitation by the light of *mixed (σ, π) polarization* which agrees closely with our measurement condition [1]. In practice $\Delta n/n_0 \approx 32$ % in this scheme.

Our studies of Zeeman CPT effects carried out with Rb beam and the calculation of the overlapping contours of neighbouring hyperfine transitions have shown the borders for practical applicability of calculation results. The adopted calculation model can be considered adequate when the field eliminating a coherent effect in optical regions is no less than 0.8 G and the excitation line's width < 40 MHz when using $|F_s=2>$- $|F'>$ transitions and $\Delta\nu_L < 3$ MHz when using $|F_s=1>$- $|F'>$ ones.

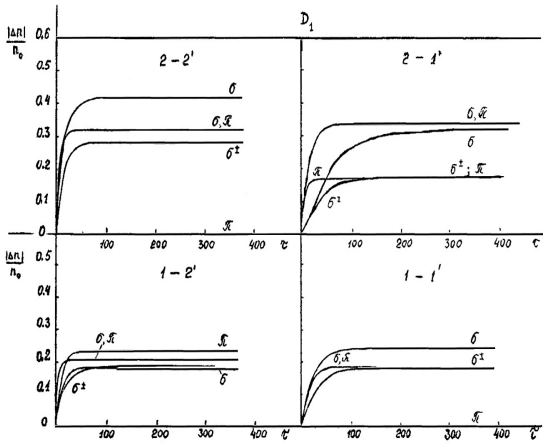


Fig.2. Pumping efficiency of ^{87}Rb atom in the beam excited by D_1 line's hyperfine components

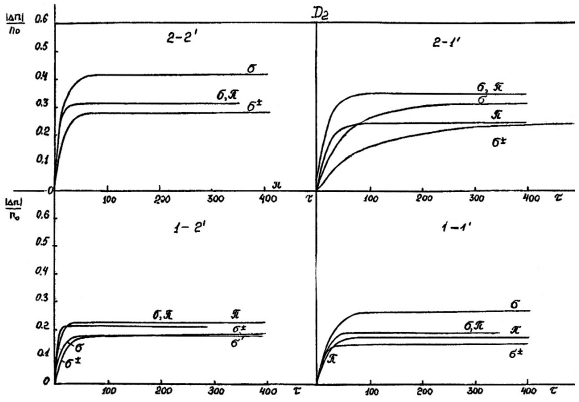


Fig.3. Pumping efficiency of ^{87}Rb atom in the beam excited by D_2 line's hyperfine components

The efficiency of one-frequency pumping is limited since a part of atoms does not interact with the light due to the hyperfine splitting of the ground state.

Table 1 gives stationary values of maximum efficiency of one-frequency pumping for $|F_s = 1, m_{Fs} = 0>$ and $|F_s = 2, m_{Fs} = 0>$ sublevels of the ground state of ^{87}Rb atom being achieved in the saturation behavior and corresponding minimum values of the parameter τ_{eff}.

Maximum efficiency of one-frequency pumping for $|F_s = 2, m_{Fs} = 0>$ and $|F_s = 1, m_{Fs} = 0>$ sublevels of the ^{87}Rb atom by D - lines' hyperfine components, fluorescence yield and $|\Delta n_{o\text{-}o} \cdot \beta|$ values

Table 1

^{87}Rb							
		D_1			D_2		
ε_λ	F_s - F'	$\lvert\Delta n\rvert$, %	β_{max}	$\lvert\Delta n\rvert \cdot \beta$	$\lvert\Delta n\rvert$, %	β_{max}	$\lvert\Delta n\rvert \cdot \beta$
σ	2-2'	41.3 (τ_{ef} = 70)	2 (n = 6)	82.6	41.3 (τ_{ef} = 70)	2 (n = 6)	82.6
	2-1'	34.4 (τ_{ef} = 230)	6 (n ~24)	**206.4**	31.2 (τ_{ef} = 250)	1.2 (n = 2)	37.4
	1-2'	17.5 (τ_{ef} = 40)	2 (n = 6)	35	17.5 (τ_{ef} = 40)	2 (n = 6)	35
	1-1'	23.8 (τ_{ef} = 100)	1.2 (n = 2)	28.6	26.2 (τ_{ef} = 80)	6 (n ~24)	**157.2**
π	2–2'	0	2 (n = 6)	0	0	2 (n = 6)	0
	2–1'	16.9 (τ_{ef} = 30)	1.2 (n = 6)	20.3	24.1 (τ_{ef} = 100)	0.7 (n = 1)	16.1
	1–2'	22.7 (τ_{ef} = 30)	2 (n = 6)	45.4	22.7 (τ_{ef} = 30)	2 (n = 6)	45.4
	1–1'	0	1.2 (n = 2)	0	17.8 (τ_{ef} = 70)	6 (n ~24)	106.5

3.2 Two-frequency Pumping

A great advantage of optical methods for creating a population inversion is the opportunity of accumulating all beam's atoms at a single magnetic sublevel ($|F_s, m_{Fs}= = 0>$). This is achieved in two-frequency pumping schemes using a forbidden optical $|F_s, m_{Fs}=0>$- $|F'= F_s, m_{F'} = 0>$ transition and simultaneous excitation of atoms from two hyperfine sublevels of the ground state. One laser tuned to the $|F_s> \rightarrow |F' = F_s>$ transition should be π-polarized. The second laser with the arbitrary polarization is tuned to the $|F_s \pm 1> \rightarrow |F'>$ component of the same or other D - line. Two-frequency pumping can give a gain in the inversion of the "0-0" sublevels' populations by 8 times for ^{87}Rb atoms. The achievement of 100 % efficiency of two-frequency pumping is, however, related to realization a few of conditions concerning intensity and width-line of both lasers, interaction time and correct selection of optical transitions. It is shown in [2] that 100 % pumping of a ^{133}Cs beam, for example, can be realized with two stabilized lasers, if they are broadband and uncorrelated. Thus, the excitation intensity should be maintained at a level at which a pumping time, T_{Fs}, for $|F_s >$ state exceeds the correlation time τ_L of each laser ray, $T_{Fs} >> \tau_L$, or accordingly:

$$\Delta\nu_L = 1/\tau_L >> 1/T_{Fs} = (2\pi/h^2)\cdot I_{FsF'}(\nu_0) \cdot |\mathbf{d}_{FsF'}|^2, \quad (4)$$

where τ_L - time of laser wave's coherence; $I_{FsF'}$ - intensity of laser emission hyperfine component.

As regards the ^{87}Rb atom, 24 variants are possible for realizing two-frequency laser pumping using eight pumping transitions. These schemes are distinguished by combination of D - lines' hyperfine components and polarization of those. The calculation results of the two-frequency pumping schemes are presented in Fig.4 in the order of pumping rate lowering. The greatest pumping rate is provided by the scheme denoted as D_2, 1-1' π; D_1 (D_2), 2 -2' σ.

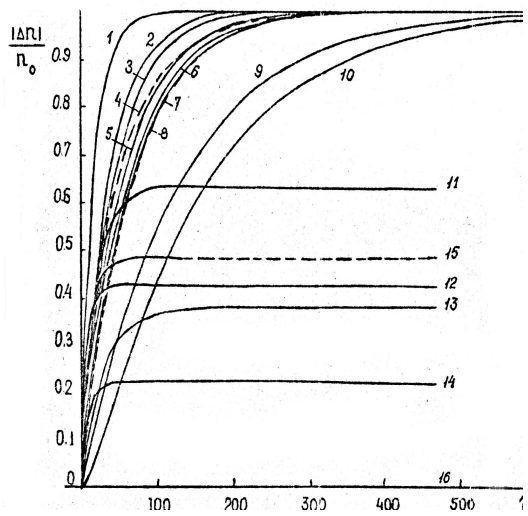


Fig.4. Schemes for two-frequency laser pumping:

1 - D_2, 1-1'π; D_2 (D_1) 2-2'σ
2 - D_2, 2-2'π; D_1, 1-2'π
3 - D_2, 1-1'π; D_2 (D_1) 2-2'π
4 - D_2 (D_1), 2-2'π; D_1, 1-1'σ
5 - D_2 (D_1), 2-2'π; D_2, 1-2'π
6 - D_2, 2-2' π; 1-2' σ
9 - D_1, 1-1' π; D_2, 2-1'σ
10 - D_1, 1-1'π; D_1, 2-1'σ
11- D_2, 1-1'π; D_2, 2-1'π
12 - D_2, 1-1'π; D_1, 2-1'π
14 - D_1, 1-1'π; D_1, 2-1'π
15 - D_2, 1-1'π; $D_1(D_2)$, 2-2'π
16 - D_1, 1-1' π; $D_1(D_2)$, 2-2' π

11÷14 - schemes of two-frequency pumping with accumulation of atoms in $|1,0>$ - state and population of "pockets" $|2, \pm 2>$;

15 - scheme of two-frequency pumping with a concurrence of accumulation in $|1,0>$ and $|2,0>$ states,

(3:1), when exciting by components D_2, 1-1'π and D_1 (D_2), 2-2' π;

16 – scheme of two-frequency pumping by components D_1, 1-1'π; D_1 (D_2), 2-2' π with the equal population of |1, 0> and |2, 0> states.

The calculation results give for each scheme maximum value of pumping efficiency, $|\Delta n|/n_0$, and are true when $\Delta\nu_L << |\nu_{F'=Fs\pm1} - \nu_{F'=Fs}|$ - in the absence of atom leakage effect which may be appeared at the simultaneous excitation of neighbouring hyperfine transitions at the wings of the pumping transition's contour ($\rho \neq 0$). The usage of these schemes in the discriminator requires to create a weak magnetic field within pumping and detection regions as well as to correctly select the radiation direction of both lasers relative to **B** field. This provided the unambiguity of pumping light polarization for beam atoms and correspondence between calculated and realized pumping schemes.

The pumping schemes with one laser are characterized by restricted efficiency and fast pumping. One may observe the related shifts caused by the microwave power. Two-frequency schemes give maximum efficiencies with slow pumping. It increases significantly a contribution of slow atoms to the frequency standard's signal. The detection at the cyclic transition in these schemes results in the distortion of the Maxwell distribution and the appearance of the weight multiplier, 1/v, in the function of atomic velocity distribution.

3.3 Signal-to-Noise Ratio for various Pumping and Detection Schemes

The evaluation of SNR, when detecting a signal of the clock "0-0" transition, based on the analysis of the population distribution in the pumping cycle and model of the noise shot nature, can be made using a relation

$$\frac{S}{N} \cong \frac{\frac{\Delta n(\infty)}{n_0} < P >_y \left[n_0 \tau_p \int_0^\infty f(y)dy \right]^{1/2}}{\left[\frac{\Delta n(\infty)}{n_0} < P >_y + \frac{n^\Sigma(\infty)}{n_0} \right]^{1/2}}, \qquad (5)$$

where $\Delta n(\infty)/n_o$ - equilibrium value of difference in populations of |1,0> and |2,0> operating sublevels at the output from the pumping region; $<P>_y$ – probability of microwave transitions between |1,0 > and |2,0 > states in the discriminator averaged over the distribution of atoms in the beam according to velocities, f(v/α); $n^\Sigma(\infty)$ – sum of populations of the $|F_s>$ state's Zeeman's sublevels proportional to a number of atoms in this state which have not gone to the sublevel not interacting with the light when exciting at the $|F_s> \rightarrow |^2P_J>$ transition. The value of $n^\Sigma(\tau)$, in such a way, characterizes a signal of background atoms; τ_p – time constant of the recording system.

When using two-frequency pumping scheme with the accumulation of all atoms in the |1,0> state (the best D_2, 1-1' π; D_1 (D_2), 2-2' σ scheme), the detecting laser should be tuned to the $|F_s = 2> \rightarrow |F'=3>$ transition.

When pumping with accumulation of all atoms in the |2,0> state (2-2' π, 1-2'π pumping scheme by D_1 or D_2 lines), it would be better to detect at the $|F_s=1> \rightarrow |F'=0>$ transition. In all these cases, the fluorescence of atoms in the detection region, when exciting the microwave resonance in the beam, is being observed on the zero background and has a maximum at the resonance frequency of the clock "0-0" transition.

The results of signal-to-noise evaluation for various pumping and detection schemes are given in Fig. 5-8.

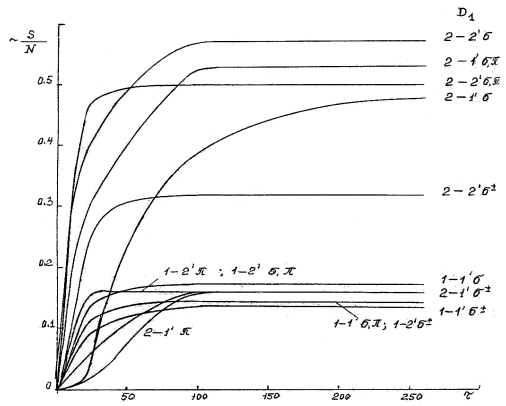


Fig.5. Values of $S/N \cdot (n_o\ \tau_p)^{-1/2}$ for various schemes of laser pumping by D_1 line and fluorescence detection by probing $|F_s = 2>$ state

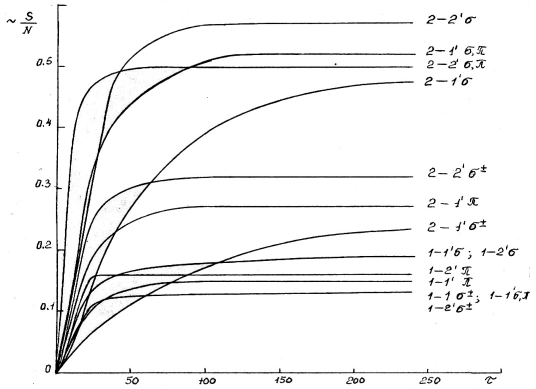


Fig.6. Values of $S/N \cdot (n_o\ \tau_p)^{-1/2}$ for various schemes of laser pumping by D_2 line and fluorescence detection by probing $|F_s = 2>$ state

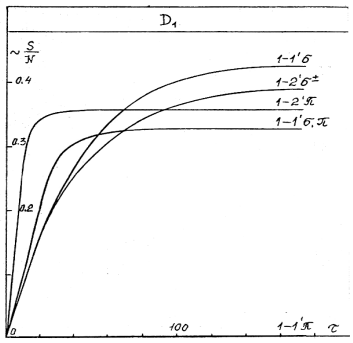


Fig.7. Values of $S/N\cdot(n_o\ \tau_p)^{-1/2}$ for various schemes of laser pumping by D_1 line and fluorescence detection by probing $|F_s = 1\rangle$ state

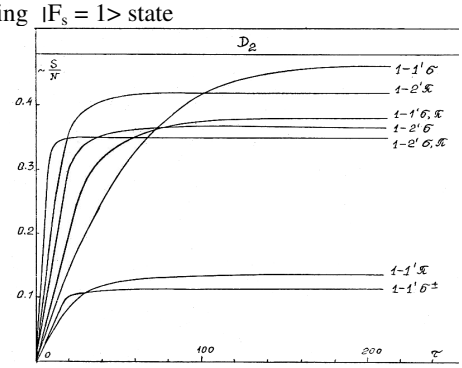


Fig.8. Values of $S/N\cdot(n_o\ \tau_p)^{-1/2}$ for various schemes of laser pumping by D_2 line and fluorescence detection by probing $|F_s = 1\rangle$ state

As is seen from Fig.5 -8, even if pumping through $|2'\rangle$ level (1-2' and 2-2' schemes) gives the same $\Delta n(\infty)/n_0$ values for D_1 and D_2 lines, the SNR values differ significantly. The anomalously low SNR in pumping at the 2-1' π (D_1) transition is explained by an origin of "trap" in $|2, \pm 2\rangle$ states because the atoms from these states are not being excited by the π - light and are being detected at the cyclic 2-3' transition as the noise.

4. OPTICAL DETECTION

The calculations made as regards a number of fluorescent photons emitting by ^{87}Rb atoms when exciting a beam by the laser at pumping and cyclic transitions.

When detecting at the *pumping* transition a value of the resonance fluorescence maximum signal, I_{max}, is not a function of a number of absorbed photons and depends on atomic beam intensity, I_{beam}, and maximum number of photons, β_{max}, which in the mean atom scatters before the transition to the sublevels non-interacting with the light. A maximum number of photons being scattered does not depend on the exciting light intensity and is determined by the selected atomic transition and limited by optical pumping.

In all works on the ^{133}Cs discriminator, in order to find a number of fluorescent photons, β, the populations $n_{m'}$ of the respective excited state, $|F'\rangle$, are being calculated, that is the systems of differential equations are being solved for $n_{m'}(t)$. When calculating a number of fluorescent photons from Rb atoms' beam, we applied the other method which was of greater simplicity and, as in a case of pumping calculation, depending on some universal parameter. As far as a life time of the excited state (τ_{rad} = 28 ns) is much less than the time of atom's transit through the detection region, $2\cdot10^{-5}$ sec, it is possible in principle to realize several "excitation-emis-sion" *cycles* which number is just that universal parameter.

The calculation of β_{max} values is based on the distinction probabilities of dipole transitions between $|F_s\rangle$ and $|F'\rangle$ - states of the hyperfine transition selected for excitation and carried out by analogy with an example in [3]. The calculations are made for all pumping hyperfine transitions of D_1 and D_2 - lines when exciting ^{87}Rb atoms by light of linear popularization (**σ** and **π**). The data is given in Fig. 9, 10 as well as in Table 1 with a number of excitation-emission cycles. $<\beta>$ values are calculated for cases of coherence effects' absence, β_{max} ($B > 0.8$ G), and full trapping, β_{cpt} ($B = 0$).

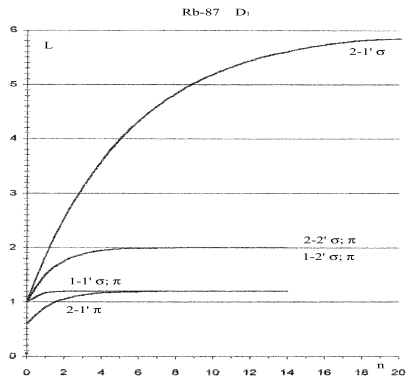


Fig. 9. Fluorescence yield L when exciting ^{87}Rb atom by D_1 line as a function of a number of cycles, n

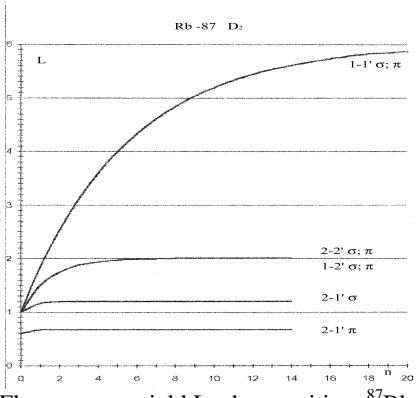


Fig.10. Fluorescence yield L when exciting ^{87}Rb atom by D_2 line as a function of a number of cycles, n

In order to check the *correctness* of the calculation method, the $<\beta_{max}>$ values are also calculated for all hyperfine pumping transitions of D – lines (**σ** and **π**) when exciting ^{133}Cs atoms. The data coincides with L_f (^{133}Cs) - values given in [4] and obtained using solutions of differential equation systems.

When using the *cyclic* detection, a maximum rate of photon re-emission during the atom's transit through the detection region is being determined by the life time of the excited state, τ_{rad}, and is equal to: $\beta_{max}/\tau_v = 1/(2\ \tau_{rad})$. The radiation life time of the $5^2P_{3/2}$ - state of the ^{87}Rb atom is equal to: $\tau_{rad} = 2.65\cdot10^{-8}$ s, $\beta_{max}/\tau_v = 1.9\ 10^7$ ph/s. At $\tau_v \approx 2\cdot10^{-5}$ s, a value of β(Rb) is $\approx$ 380 phot/at. For ^{133}Cs atom being excited at the $|F_s = 4\rangle$ - $|F' = 5\rangle$

transition, $\tau_{rad} = 3.15 \cdot 10^{-8}$ s, $\beta_{max} / \tau_v = 1.6 \cdot 10^7$ ph/s, and $\beta(Cs) \approx 400$ phot/at.

5. EFFICIENCY OF PUMPUNG/DETECTION SCHEMES ($|\Delta n_{o-o} \cdot \beta|$ values)

The Table 1 is composed for $|\Delta n_{o-o} \cdot \beta|$ values characterizing the pumping/detection scheme's efficiency using difference in populations, Δn_{o-o}, and fluorescence yield, β_{max}, and are given for four pumping transitions of both D - lines when exciting ^{87}Rb atoms.

5.1 D_1 line

When using the D_1 - line, the following schemes give the *largest value* of $|\Delta n_{o-o} \cdot \beta|$ (a):

$2 - 1'\ \sigma\ (D_1) \rightarrow |\Delta n_{o-o} \cdot \beta| = \mathbf{206}$ for ^{87}Rb;
$4 - 3'\ \sigma\ (D_1) \rightarrow |\Delta n_{o-o} \cdot \beta| = 66$ for ^{133}Cs.

These schemes, however, provide a lesser signal-to-noise ratio in the discriminator due to the strong noise caused by the long pumping time for realizing Δn_{o-o} (max), and the largest number of fluorescent photons per atom, β, due to strong probabilities of transitions. In addition, these schemes are subjected to the effect of coherent population trapping which vanishes when B > 0.8 G (for ^{87}Rb) and B > 0.3 G (for ^{133}Cs).

Proceeding from $|\Delta n_{o-o} \cdot \beta|$ parameter, the following schemes give a lesser gain :

$1 - 2'\ \pi\ (D_1) \rightarrow |\Delta n_{o-o} \cdot \beta| = 45.4$ for ^{87}Rb
$3 - 4'\ \pi\ (D_1) \rightarrow |\Delta n_{o-o} \cdot \beta| = 34.6$ for ^{133}Cs

However, these schemes have a lesser pumping time (7 times for ^{87}Rb) and are free from CPT effect due to the π – polarization.

5.2 D_2 line

When using the D_2 line, the *most efficient* schemes with respect to the $|\Delta n_{o-o} \cdot \beta|$ parameter are as follows (b):

$1 - 1'\ \sigma\ (D_2) \rightarrow |\Delta n_{o-o} \cdot \beta| = \mathbf{157}$ for ^{87}Rb
$3 - 3'\ \sigma\ (D_2) \rightarrow |\Delta n_{o-o} \cdot \beta| = 48.8$ for ^{133}Cs

For ^{87}Rb, the $2 - 2'\sigma\ (D_2)$ scheme for which $|\Delta n_{o-o} \cdot \beta| = \underline{82.6}$ is more efficient as regards pumping ($\Delta n_{o-o} = 41.3$ %), but lesser efficient with respect to β (3 times lesser). This scheme has such advantages as somewhat lesser pumping time and significantly lesser detection constant. In addition, when the ABT functions according to the two-laser scheme, the stronger cyclic (2-3') transition is being used for the detection. For eliminating the CPT effect, the fields of B > 0.8 G are superimposed within both optical regions.

The respective pumping and detection schemes for ^{133}Cs are as follows:

$4 - 4'\ \sigma\ (D_2)$, pumping $\rightarrow |\Delta n_{o-o} \cdot \beta| = \underline{37.2}$ and
$4 - 5'\ (D_2)$, detection.

The comparison of (a, b) – schemes' efficiency for Rb and Cs atoms shows that, for **both D – lines**, the **$|\Delta n_{o-o} \cdot \beta|$ - parameter is 3 times higher** for **^{87}Rb atoms** than for ^{133}Cs; the population difference, **Δn_{o-o}, is 2 times higher**, and the fluorescence yield, **β , is 1.5 times higher**.

6. CONCLUSION

• Using a "rate equations" formalism, the systems of kinetic equations are composed and numerically solved for populations of eight sublevels of the ^{87}Rb atomic ground state in a cycle of one- and two-frequency pumping.
• Efficiency of pumping "0-0" transition's sublevels for all schemes using D_1 and D_2 – lines of various polarization is analyzed. For each scheme, using a shot-noise limit, a value of maximum SNR ratio is estimated.

The most efficient schemes of one- and two-frequency pumping for ^{87}Rb atom are recognized:
$2\text{-}2'\sigma\ (D_1)\ D_2$ which provides $\Delta n_{o-o} = 41.3$ % ;
$2\text{-}2'\sigma\ (D_1)\ D_2$; $1\text{-}1'\pi$, D_2 (for two-frequency pumping).

It is found that the adopted calculation model and its results are true under conditions when field B > 0.8 G; $\Delta\nu_L \leq 25 \div 40$ MHz when using $F_s = 2 \rightarrow F' = 1; 2; 3$ transitions, $\Delta\nu_L < 3$ MHz when using $F_s = 1 \rightarrow F' = 0;1;2$ transitions ($\rho \leq 0.8\%$).

• Using a method of counting cycles, a number of fluorescence photons of ^{87}Rb and ^{133}Cs atoms is calculated when exciting a beam at various hyperfine transitions of D – lines of σ and π – polarization.
• For all excitation schemes the table of $|\Delta n_{o-o} \cdot \beta|$ values is composed. It is shown that, for both D – lines, the $|\Delta n_{o-o} \cdot \beta|$ parameter is by 3 times higher for ^{87}Rb atoms than for ^{133}Cs atoms; the pumping efficiency, Δn_{o-o}, is 2 times higher and fluorescence yield, β_{max}, is 1.5 times higher.

As a whole, the studies have demonstrated a great promise for using the ^{87}Rb isotope as a working substance as the alternative of ^{133}Cs atoms when developing a high-stable small-size atomic beam frequency standard with laser pumping.

REFERENCES

[1] A.N. Besedina, A.B. Bassevich, A.G. Gevorkyan et al. Atomic beam Rubidium frequency standard with laser pumping and detection, presented at the 5th Russ. Sympos. "Metrology of Time & Space", VNIIFTRI, Moscow, Russia, Oct. 11–13, 1994.

[2] E. de Clercg, M. Tetu, et al. Laser diode optically pumped Cs beam, J. Physique, v. 45, pp.239-247, 1984.

[3] S. Seiikhuis, G. Nienhuis, R. Morgenstern. Inhibition of the Larmor precession during optical pumping in a weak magnetic field, Phys. Rev. A, vol. 33, no. 6, pp. 3977-3982, 1986.

[4] N. Dimarcq, P. Cerez, V. Giordano, G. Theobald. Comparison of pumping a Cs beam tube with D_1 and D_2 lines, J. Appl. Phys, v. 69, no. 3, pp.1158-1162, 1991.

This work was supported by the INTAS-CNES Project no. 03-53-5175.

Session 3B / VCO & synthetisers

A SPICE SYSTEM SIMULATION MODEL FOR PHASE-LOCK LOOPS AND ANTI-JITTER CIRCUITS

Michael J Underhill* and James Brodrick

Toric Limited UK, and * University of Surrey, UK

Abstract

The Anti-Jitter Circuit (AJC), announced at EFTF 1996 has undergone progressive improvement since then (EFTFs 1997 to date and the references [1 to10]). It is a feed-forward jitter cancellation device competing with the Phase-Lock Loop (PLL). A System Simulation Model for AJCs and PLLs and combinations of these has now been created using a SPICE platform (Catena SIMetrix). This model provides the jitter suppression transfer function together with a full representation of all internal circuit noise sources as they impact on the (intrinsic) noise seen at the circuit output. The model, with the working title Toric "SpringBoard", is capable of modelling AJCs and PLLs separately, or as a combination of both. Dividers or phase accumulators, as used in PLL and DDS frequency synthesizers, can be included in the model. The output noise spectrum and the suppression of input noise can be computed very rapidly, typically in less than a second. Time delays from sampling or comparator delay are included. The transient response can also be found, but the wide ratio between maximum and minimum frequency of interest mans that this typically requires a minute or more for each simulation run..

Introduction

The Anti-Jitter Circuit (AJC) is a unique feed-forward circuit for cancelling time and phase jitter. Feedback is used for DC stabilisation and to ensure good cancellation of input jitter. The AJC has been progressively improved since its introduction at 1996 EFTF [1 to 10]. Recent new techniques include:

1. 'AF = AC Feedback' shown with the 'A2S2 = AJC Analytic System Simulator' at Tampa EFTF-IFCS in 2002 [7] – for good spectrum filtering and transient settling.
2. 'EF = Enhanced Feedback' – patented and presented at Montreal UFFC- IFCS 2004 [10] – for very low suppression cut-off frequency with good transient response.
3. 'DF = Digital Feedback' – announced Montreal UFFC - IFCS 2004 [10] – for low intrinsic noise.
4. 'BF = Bridge Feedback' – announced in the presentation of this paper in EFTF Besançon 2005 – for reduced size on-chip filter capacitors and good AJC filter characteristics.
5. 'Step-up-step-down' – for improved suppression and intrinsic noise at all sideband frequencies

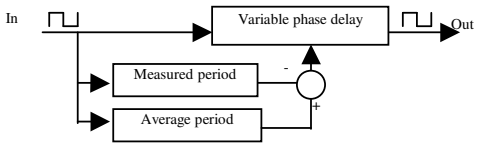


Fig. 1 Basic AJC Feed-forward Noise Cancellation Process.

Fig. 1 shows that the jitter suppression in an AJC is a feed-forward cancellation process. This process can be performed by the very simple circuit shown in Fig. 2.

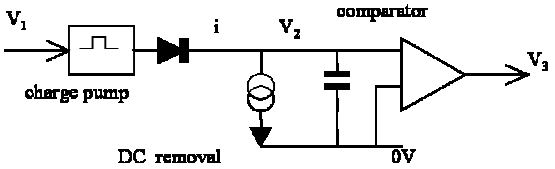


Fig. 2 Simple Circuit of the Basic AJC

Fig.3 shows the key parts of the time waveforms in the AJC of Fig. 2. The input pulse stream V1 is shown with jitter on the middle pulse. The charge pump generates equal area pulses irrespective of the jitter and these are integrated in the integrator capacitor to give a sawtooth waveform V2. The current source creates the discharging down-slope. It has to be controlled with a long time constant so that there is no mean upward or downward drift of the integrator capacitor voltage waveform V2.

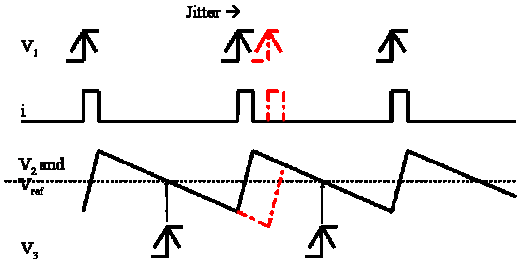


Fig. 3 Time Waveforms of the AJC

The jitter suppression comes from noting that jitter on any pulses does not change the timing of the intersection of the sawtooth down-slope with the Reference voltage level Vref.

Triggering an output monostable from these points gives an output waveform with the jitter removed.

The feedback loop controlling the discharge current sets the lowest frequency of suppression. At the lowest suppression frequency, for loop stability, the loop gain must be approximately unity. On the other hand a DC loop gain of typically 10 is required for the AJC integrator to stay within range as the frequency is changed over a wide range. The way in which this can be achieved with an 'Adiabatic AJC' is shown in Fig. 3.

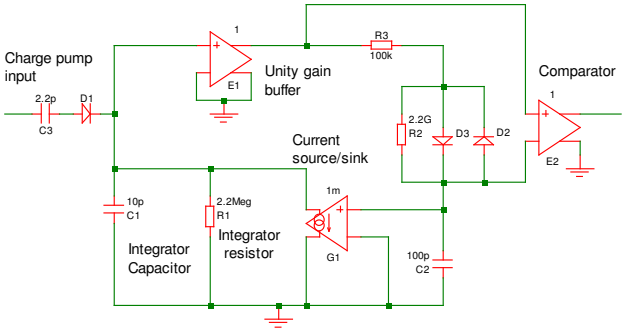


Fig. 4 Buffered Adiabatic AJC with Speed-up Diodes

AJC Design Rules

One purpose of a SPICE System Simulator is to confirm, or create and confirm, the design rules for the system of interest.

The design rules for such an AJC such as the one shown in Fig.4 are as follows:

1. Choose the maximum integrator current as

   $i_{max} = f_{max} \times V_{pp} \times C1$

   where Vpp is the peak-to-peak sawtooth voltage and f_{max} is the maximum frequency of operation

2. Choose the DC loop gain as

   $G_{DC} = gm \times R1 = 1 \text{ to } 10$

   where the higher value is needed to cover a large frequency range. Gm is the transconductance of the current source/sink.

3. Limit the AC loop gain (for critical damping and minimum 'peaking'):

   $g_m \times R1 \times C1R1/C2R2 \leq 1$

Note that the integrator resistor R1 should be as high as possible for the minimum loop bandwidth and suppression frequency, $1/2\pi C1R1$.

But then the loop gain condition (3) typically requires that the loop filter has a long time constant, C2R2. C2 may be made large for a 'discrete component' AJC on a board. On a chip large capacitors are not possible and so high value resistors or an equivalent current source is required.

If diodes are available in the IC process a back-to-back pair may be used to provide a 'speed-up function for start-up and large frequency jumps. If no resistor is connected across the back-to-back diodes the diodes themselves provide an equivalent resistor R1 which becomes very high after the AJC has settled. Fig. 4 shows the back-to-back diodes; and the value shown for R2 is the minimum effective resistance to ensure a well damped settling response. For very fast acquisition the value of R3 can be typically be reduced to about 10k.

The Enhanced Feedback (EF) AJC technique

An additional negative feedback loop may be placed around the AJC to reduce its internal noise and to lower the lowest suppression frequency. It can relieve the requirement that the inner loop should have a low cut-off frequency. It can also increase the suppression obtainable at intermediate frequencies.

The additional blocks required to implement the Enhanced Feedback (EF) technique are shown on the right hand side of Fig.5 (in red).

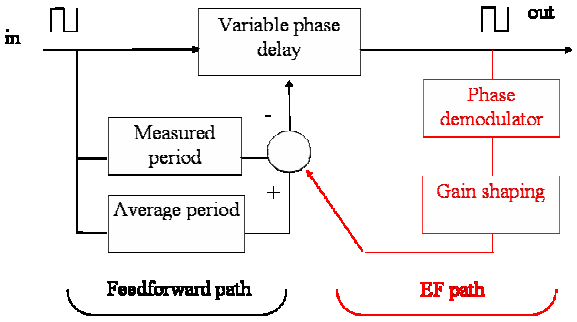


Fig. 5 Block Diagram of EF-AJC Technique

The phase demodulator is a high pass filter (differentiator) with a cut-off frequency that should be at or below the lowest frequency of interest. Its transient time response is shown in Fig. 6.

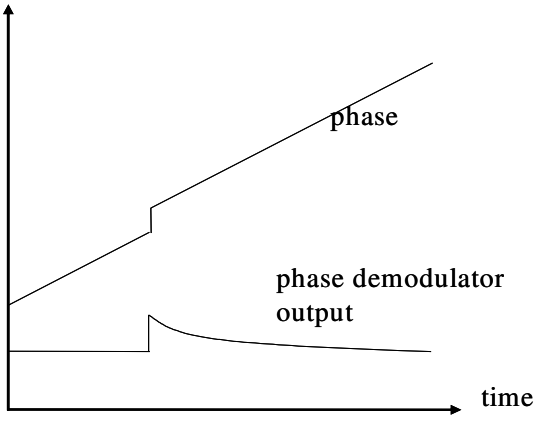


Fig. 6 Phase Demodulator Characteristic

For the EF arrangement shown as a 'SpringBoard' SPICE model in Fig. 8, Fig. 9 shows the improvements that can be obtained in suppression and self-generated 'intrinsic' noise, by using the EF technique. These modelled results are better than for most PLL VCOs. For reference on the noise vertical axis 1µ corresponds to -120dBc/Hz of sideband phase noise, L(f). The improvement of suppression to some 60dB at some frequencies may also be seen.

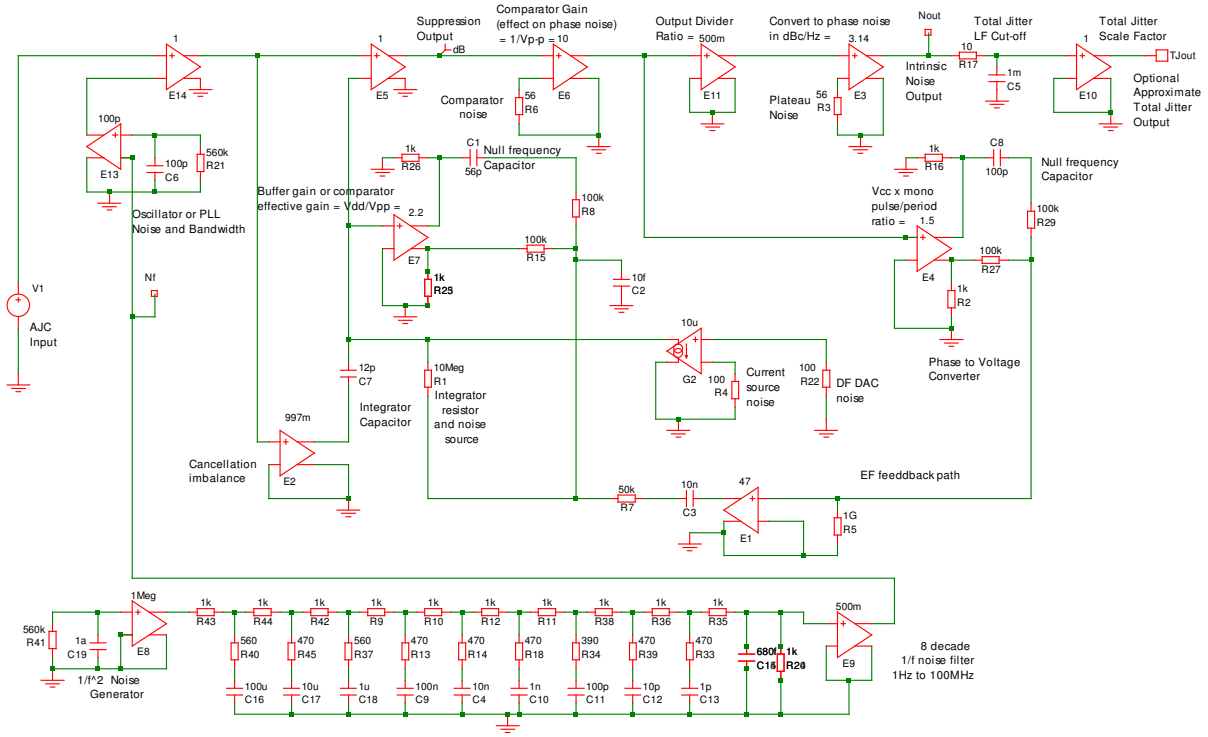


Fig.7 SpringBoard Model of 100 MHz Step-Up-Down EF AJC – with PLL and 1/f noise generators

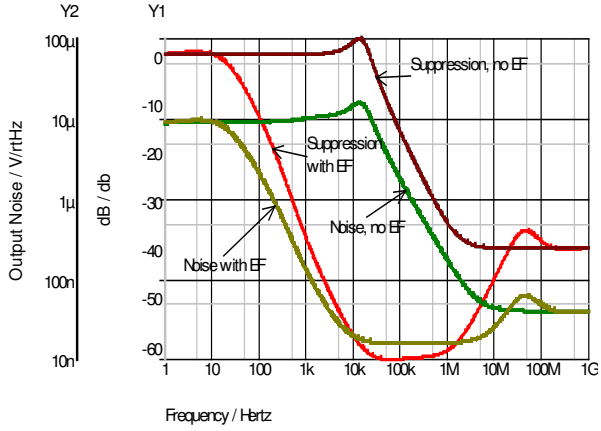


Fig. 8 AJC Suppression and Noise Comparisons with and without the EF loop

The 'SpringBoard' modelling system also provides an approximate total jitter figure of the AJC. This is plotted in Fig.9. The value depends on the lowest frequency included in the integration and this is plotted as the abscissa as shown.

A substantial improvement (about 40db) occurs around the original cut-off frequency of about 10kHz. The asymptotic limit at the lowest frequency showing that the EF technique moves the noise into the ('wander') low frequency region. For period or period-to-period jitter, the high-pass action included in the jitter integral removes the low frequency noise and the benefit of the EF technique remains at these low frequencies.

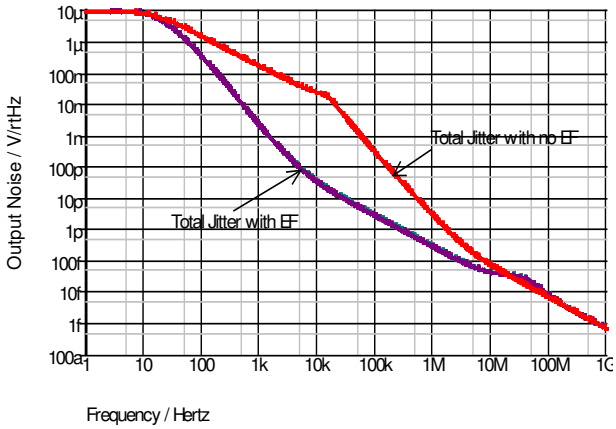


Fig. 9. Approximate total jitter comparison with and without EF

'SpringBoard' SPICE System Simulator

The Toric 'SpringBoard' System Simulator is a 'baseband' model that represents the sideband frequency spectra of suppression and noise directly.

It can be implemented in any version of SPICE from any vendor.

The preferred SPICE version and the version used in development is the "SIMetrix" SPICE demonstration version available as a free download from Catena UK. Typical AJCs and PLLs can just be accommodated in the demonstration version separately, but for these two combined the full version is then recommended. SIMetrix has some generally useful features not in other vendor's SPICE versions and the FFT function is particularly to be recommended.

SpringBoard is a 'High Level' model in using 'ideal' components. For example all AJC functions are translated into resistors, capacitors and summing amplifiers. Although it is a high-level model, real models of components may be substituted in the baseband feedback loop parts of the system. . Time delay elements can also be used to represent sampling delays when appropriate.

Noise sources in general can all be represented by equivalent resistors. The exception is for 1/f noise where either the internal parameters of SIMetrix resistors may be edited or the 1/f noise source shown in Fig. 7, and described below, may be used.

The SPICE main functions used are:

1. AC Analysis for plotting AJC Jitter Suppression against frequency
2. Noise Analysis for plotting the intrinsic noise originating inside the AJC
3. Transient Analysis for plotting phase settling after a frequency step.

The first two processes run very fast with SIMetrix SPICE, typically taking no more than a second or so. The transient analysis process is slower by about 10 times and even longer if time delays are used.

For the SpringBoard AJC circuit in Fig. 7 significant features and ponts to note are:

1. The integrator capacitor C7 is also used a means of coverting input phase into frequency (in this case a voltage representing frequency).
2. The integrator resistor R1 is also assumed to have thermal noise which can be a major contribution to the 'intrinsic' or internal AJC noise that appears at the output. (In fact all real resistors have noise that has to be taken into account.)
3. The AJC cancellation is carried out in differential amplifier E5. The cancellation imbalance is set by gain value assigned to amplifier E2. A value of unity corresponds to perfect cancellation, 0.9 corresponds to a cancellation of 20dB, 0.99 to 40dB and 0.997 to 50dB.
4. A buffer amplifier E7 has been placed in the inner feedback loop. Typically this would be a unity gain buffer amplifying the integrator sawtooth waveform directly. In the simplest AJCs no buffer is used but only a directly connected high value resistor. The loop feedback can be taken from the output of the comparator. This has the advantage of providing auto-biasing of the comparator to its correct switching point. It has the disadvantage that an effective gain of Vdd/Vpp appears at this point in the equivalent feedback loop.

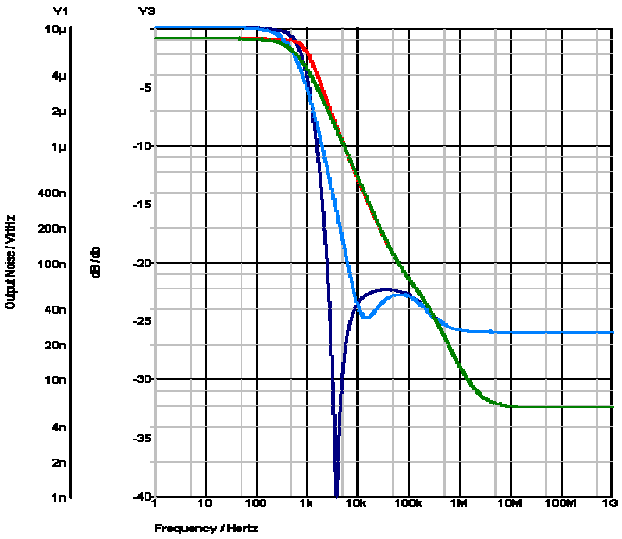


Fig. 10 AJC Suppression and Intrinsic Noise Using Bridge Feedback

5. 'Bridge Feedback' (BF) can be used optionally if an inverting buffer output is available. This is implemented by R8, R15, and C1. (C2 represents unwanted stray capacitance.) The carrier frequency is cancelled but the DC negative feedback is not. C1 allows the 'zero' frequency to be set. Note that in this AJC unlike the one shown in Fig. 4 the required low gain for feedback stability is provided also provided by the integrator resistor R1. Fig 10 shows the effect of BF on the spectra of suppression and intrinsic noise. The intrinsic noise is little affected by the bridge feedback and only around the cut-off frequency. (red and green curves). However the zero position

has a marked effect on the AJC suppression characteristic (light and dark blue curves).

6. Current source G2 is controlled from a low noise DAC and an up-down counter to provide the main feedback control of the integrator discharge ramp of the sawtooth waveform. This technique is known as 'Digital Feedback' (DF) and is found to give major improvements in intrinsic noise (of typically 20dB or more).
7. The AJC input for plotting the AJC suppression is the source V1.
8. The suppression is plotted from the output of the cancellation amplifier E5.
9. The intrinsic noise is measured at the output of E3. The noise voltage at the input of E3 is converted to phase noise in 1Hz bandwidth by giving E3 a gain of $\pi = 3.14$.
10. The effective gain of the comparator for phase noise is 1/Vpp and this is assigned to E6 This shows that a large integrator sawtooth voltage improves the AJC intrinsic noise performance.
11. If a digital divider is placed on the output of an AJC the RMS phase noise is improved by the divider ratio, n. To represent this E11 is given the gain value 1/n.
12. Amplifier E13, C8 and R21 simulate a noise source with a flat top, a cut off frequency and a 6dB/octave roll-off. The top value and the cut off frequency can be set to represent the spectrum of an oscillator on its own or a PLL locked oscillator. In this way SpringBoard can be used to examine the potential for a given AJC to improve the spectra of such sources.
13. All resistors act as noise sources inpractice and SIMetrix correctly assigns thermal noise to each resistor. It is part of the optimisation process of an AJC to choose these to minimise the noise contributions of each and the overall noise of all noise sources in combination.
    It is is also possible to choose a resistor value to represent the noise in active devices.
    There are some examples of this in Fig 7:
14. R4 represents the noise of the current source, and the 100 ohm value has been chosen on the assumption of devices with a 1dB noise figure and a 1k noise match resistance.
15. R6 represents comparator noise and the value of 56 ohms was chosen to be equivalent to a comparator with a -150 dBc/Hz noise performance.
16. R3 represents the plateau noise of theAJC output amplifier, set to -150 dBc/Hz.
17. An approximate estimate of total time jitter is formed by the componentsR17, C5, and E10. A better integration of the noise should be implementable in the future using the 1/f (power and 1/f ½ in amplitude) filter shown at the bottom of Fig 7.
18. The EF (outer) loop is formed by the amplifiers E1 and E4, and the components surrounding them. Note that Bridge Feedback is being used. It is particularly advantageous for EF as the zero has a more pronounced beneficial effect on the stability of this relatively high gain feedback loop. DC feedback can be used with careful attention to biasing. The large value capacitor C3 may be removed.
19. At the bottom of Fig. 7 is an 8 decade 1/f noise filter implementable in any SPICE version. Some of the resistor values have been fine tuned to give an optimum flatness to the frequency response of this circuit shown in Fig. 11. The error is less than ± 0.5dB.
20. The 1/f noise filter may be used with a white noise source to give a 1/f noise source or with a $1/f^2$ noise source to give a $1/f^3$ noise at the filter output.



Fig.11. 1/f Noise Filter, 1Hz to 100MHz – Used with: (a) white noise source for 1/f noise(b) 1/f2 source for 1/f3noise

SpringBoard PLL Simulator

The SpringBoard Simulator may also be applied to optimise the design of any Phase-Lock Loop (PLL). Fig. 12 show a representation of a type 2 third order PLL as a control system

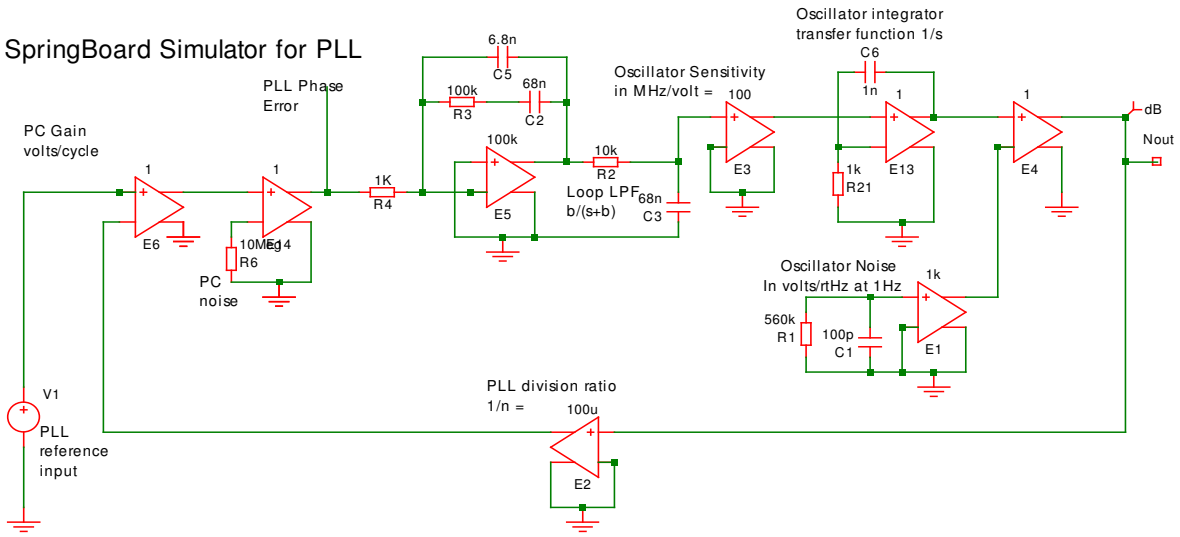


Fig. 12 Springboard (SPICE) simulator for optimisation of Phase-Lock Loop noise and transient response.

As an example, Fig. 13 shows how the output noise of a PLL is optimum if the loop bandwidth is chosen to be the intersection of the phase comparator noise spectrum referred to the VCO output with the VCO noise spectrum. The middle curve and top leftmost curves show optimum designs for low and high phase comparator (PC) noises respectively. The bottom curve shows the case for too low a PLL bandwidth for a low PC noise. The top rightmost curve shows the case of the PLL bandwidth being too low for the case of a high PC noise.

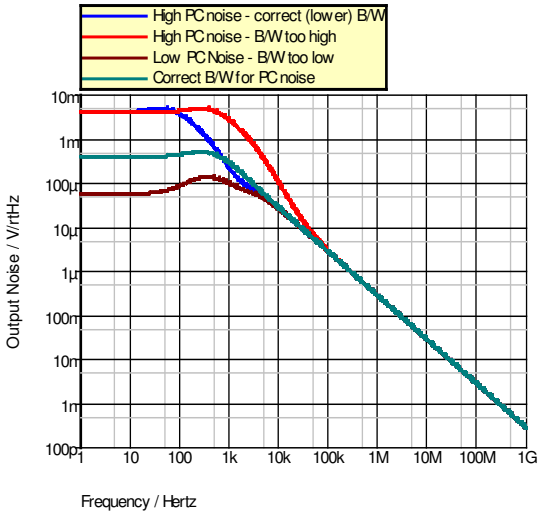


Fig. 13 PLL Noise Spectrum Optimisation with SpringBoard (Spice) Simulator

This shows that SpringBoard is useful both for PLL and AJC optimisation.

Conclusions

Any version of the standard circuit simulator SPICE can act as the platform for a High Level Systems Simulation of Noise, Transfer Function, and Transient Response of Phase-Lock Loops (PLLs) and Anti-jitter Circuits (AJCs) and combinations of these.

References

[1] M. J. Underhill and M. J. Blewett, "Spectral Improvement of Direct Digital Frequency Synthesisers and Other Frequency Sources", *Proc. 10th European Frequency and Time Forum*, Brighton, pp. 452-460, 5-7 March 1996.

[2] M.J. Underhill and M.J. Blewett, "Performance Assessment of a Delay Compensation Phase Noise and Time Jitter Reduction Method", *Proc. 11th EFTF*, Neuchatel, pp. 364-368, 4-6 March 1997.

[3] M.J. Underhill, S. Stavrou, M.J. Blewett and N. Downie, "The Anti Jitter Circuit for Low Spurious DDS Square Waves and Low Cost Fractional N Synthesis", *Proc. 12th EFTF*, Warsaw, pp. 292-297, 10-12 March 1998.

[4] M.J. Underhill, "The Adiabatic Anti-Jitter Circuit", *Joint IFCS and 13th European Frequency and Time Forum*, Besancon, pp. 611-614, 4-6 April 1999.

[5] M.J. Underhill, "The Adiabatic Anti-Jitter Circuit", *IEEE Trans. UFFC*, Vol. 48, No. 3, pp. 666-674, May 2001.

[6] M.J.Underhill "Phase noise limits of the anti-jitter circuit and on-chip RC oscillators", *Proc. 14th EFTF*, Torino, Italy, March 2000.

[7] M.J.Underhill, "The Noise and Suppression Transfer Functions of the Anti-Jitter Circuit", *Joint IFCS and 17th EFTF*, Tampa, USA, 5-8 May 2003.

[8] M.J. Underhill, "Anti-Jitter Circuits, EF-AJC", UK Patent Application #0416627.8, 26th July 2004

[9] M.Underhill and P.Brown, "Estimation of Total Jitter and Jitter Probability Density Function from the Signal Spectrum", *Proc. 18th EFTF*, Guildford ,UK, April 2004.

[10] M. J. Underhill and J. Brodick, "The Performance of the Anti-Jitter Circuit with Enhanced Feedback ('EF-AJC')", Proceedings of the 2004 IEEE International Frequency Control Symposium and Exposition, pp158 – 164. (2004 IEEE International Ultrasonics, Ferroelectrics, and Frequency Control Joint 50th Anniversary Conference *IFCS*, Montreal, Canada, 23-27 August 2004).

AN INTEGRATED DIFFERENTIAL SINGLE-CHIP VCO FOR S-BAND

Hannu Salminen[1], Pekka Eskelinen[2] and Jan Holmberg[1]

[1]VTT Information Technology, Finland
[2]Helsinki University of Technology, Finland
email: pekka.eskelinen@hut.fi

ABSTRACT

Phase modulation in its various forms is used to give systems a reasonable compromise of frequency usage and technical performance. Local oscillators that are readily capable of giving properly phased outputs are favoured and frequency agility precludes narrow-band phase shifters. A differential circuit that is extendable to a multiple local oscillator source has been designed in the form of a single chip integrated voltage controlled oscillator. The prototype unit is manufactured with a 0.35 µm SiGe Bi-CMOS process. The true circuit area of the prototype is 0.9 mm^2 including on-chip inductors. The prototype chips run on 2 V supply and draw about 3.1 mA current, have a tuning range from 2.1 to 3.6 GHz and an output power of -20 dBm. The tuning sensitivity is 2 GHz/V, frequency pushing 47 MHz/V and harmonic level below -18 dBc in all cases. One of the greatest difficulties was the low Q of substrate-based coils. At 1 MHz from the center frequency the phase noise level was -102 dBc/Hz. Some challenges in the use of modern simulation tools for phase noise analysis and oscillator frequency tuning are finally highlighted.

1. INTRODUCTION

Many modern communication and radar systems utilize modulation and detection schemes, which are based on so-called I and Q signals having a phase shift of 90 degrees. In some other designs, opposite phase signals are encountered to give a bi-phase spectrum. An example of such circuitry, although just of block diagram level, is highlighted in Figure 1. These efforts with the modulation and demodulation circuits are targeted towards achieving an optimum balance between band usage and transmission capacity [1].

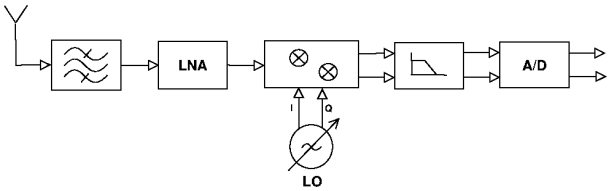


Figure 1. An example of modern communication receiver circuit employing multiple-output local oscillator.

Of course, in radar work the aim is to improve detection possibilities [2]. Regardless of application, these modulator devices require local oscillators giving suitably phased outputs, for example in quadrature form [3]. Frequency agility is typically desired [4], whereby circuit topologies relying on inherently narrow-band external phase shifting networks are less favourable. This paper describes a single chip voltage controlled oscillator concept that later enables an alternative and easy approach with multiple outputs having appropriate phase shifts.

2. OSCILLATOR CIRCUIT

A completely integrated design having no external components was selected as a basis for the VCO project. A similar approach is commercially used e.g. in the ubiquitous POS-modules manufactured by MiniCircuits. Our prototype unit is manufactured with a 0.35 µm SiGe Bi-CMOS process, which gives a reasonable compromise between a high cut-off frequency of 60 GHz and low manufacturing costs. These principal characteristics are summarized into Table 1. In fact, at the time of design, two alternatives - a high speed and a high voltage process were offered - but the clear benefits of the much higher cut-off frequency are apparent. It further supports relatively easy integration with common CMOS-type logic modules, as indicated e.g. in [5] and [6]. Figure 2 shows a concept view of a HBT (Heterojunction Bipolar Transistor) that has been processed from SiGe.

Table 1. Some selected parameters of the selected High Speed SiGe BiCMOS process.

Parameter	Value
Maximum supply voltage	3.3 V
Metal layers	4
Poly layers	4
High resistivity poly	1.2 kΩ/sq
Poly1/Poly2 capacitance	0.9 fF/µm^2
Metal2/Metal3 capacitance	1.25 fF/µm^2
BV_{CEO}	2.7 V
f_T/f_{MAX}	60/70 GHz

The initial oscillator circuit concept is illustrated in Figure 3. It is based on two transistors in differential

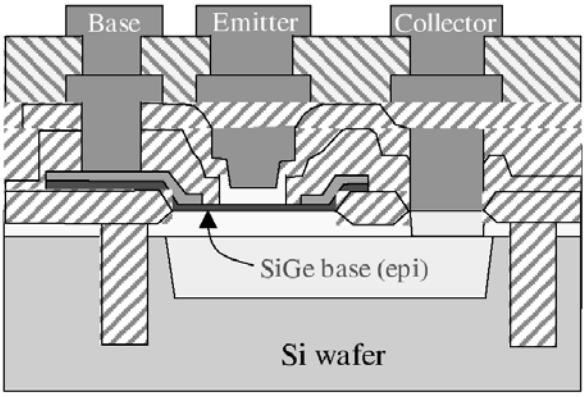


Figure 2. The principal structure of a SiGe Heterojunction Bipolar Transistor (HBT).

configuration. This negative resistance topology is known to be very reliable and robust. Bias networks and mandatory output buffer stages have been omitted from the picture to enhance clarity. Two different oscillators were actually designed in a very short time, partly to test the feasibility of the new SiGe process. The first oscillator type has no tuning input and hence no varactor diode. However, it was clear that very valuable process-related data can be obtained from that as well. The second design, processed in the same run, is the actual VCO. Figure 4 depicts the final physical lay-out of the VCO design on SiGe substrate. One of the greatest challenges was the low Q of substrate-based coils [7], clearly visible in the photo that was taken with the probe station camera. The true circuit area of the prototype is 0.9 mm^2 including on-chip inductors.

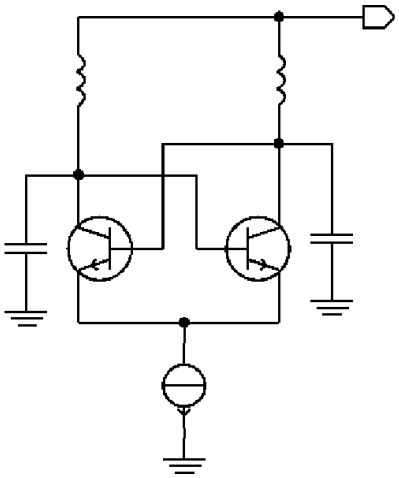


Figure 3. A differential, two-transistor oscillator used in this design project.

Because there was very little time available for the entire design of the test circuits, very strict specifications were avoided. However, the following target values, shown in Table 2, were set.

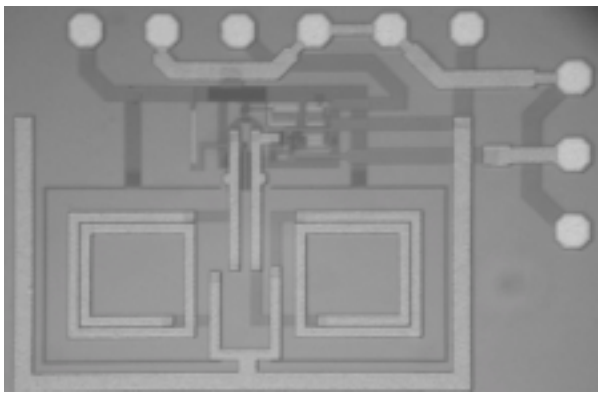


Figure 4. The final VCO implementation on the substrate as seen through the probe station microscope.

Table 2. Target values for the prototype SiGe oscillators

Parameter	Value
Center frequency	2.4 GHz
Supply voltage	< 3 V
Supply current	< 10 mA
Output power	> -20 dBm
Output SWR	< 1.5
Realization	no external coils

3. SIMULATED PERFORMANCE

Basically, three computer simulation tools were available for this design task. One of the most versatile circuit simulators is APLAC (Analysis Program for Linear Integrated Circuits) because it is inherently open and the user is free to modify all features and parameters as required by the current problem at hand. However, it has two shortcomings: first, versatility brings unavoidably some complexity to the usage and secondly, not all chip manufacturers support the APLAC component models. A common simulation tool is, of course, Agilent's ADS (Advanced Design System), which has many useful characteristics but is more closed in nature, e.g. there is no programming possibility. A local limitation was the lack of certain ADS modules necessary for design rule checking (DRC) and integrity test (LVS). Mentor software was also theoretically suitable for this design, but unfortunately it was most of the time unavailable. However, mandatory crosschecks and final output formatting (into GDS-II) were done with it.

Some relevant examples of simulated circuit performance are collected into Figs. 5 to 7. First, a lot of simulation runs were done to the fixed-frequency version. Some concern was caused by the combination of relatively low supply power and desired final output RF power as that might yield to distortion problems. Figure 5 shows the time-domain plot of the final emitter follower stage as simulated in ADS. Despite the swing is not so high, actually 90 mV_{pp} corresponding to -17 dBm into 50 Ω, the characteristic shape is rather far from

sine. If the bias current of the final stage were increased, we would of course get rid of the lower clipping but at the expense of increased DC power consumption. With this time-domain performance, we can rather easily predict the spectrum characteristics that are illustrated in Figure 6. The level of the second harmonic in this simulation is just -10 dBc. Finally, Figure 7 shows the tuning transfer function of the VCO prototype.

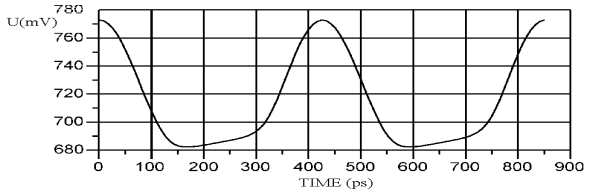


Figure 5. Simulated output voltage waveform of the fixed-frequency version into 50 Ω load. Test point is at the emitter follower.

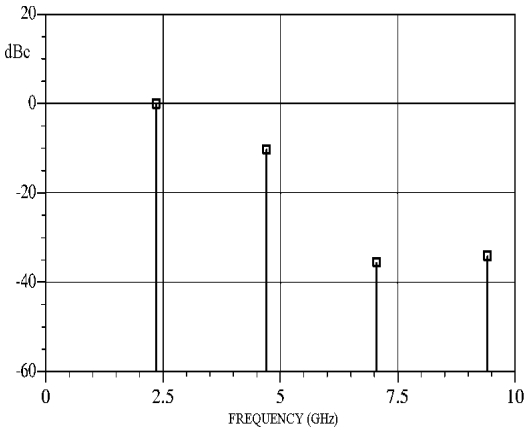


Figure 6. Simulated wide-band spectrum for the fixed-frequency version.

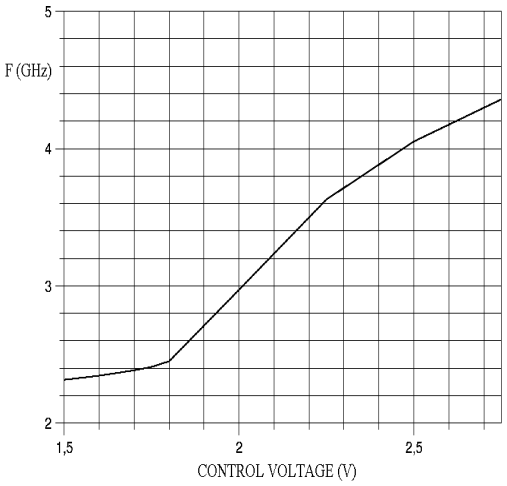


Figure 7. Simulated tuning characteristics of the VCO circuit. From about 1.8 V the response is 2.6 GHz/V.

This simulation suggests a high sensitivity of about 2600 MHz/V, which is perfect for some frequency-agile military systems, e.g. radar intermediate frequency oscillators, but less suitable without "reduction" capacitors (placed parallel to the actual varactor) to typical communication applications

4. MEASUREMENTS

The manufactured prototype chips run on 2 V supply and draw about 2.1 mA current and have a wide measured tuning range from 2.1 to 3.6 GHz without any component changes or re-dimensioning. Critical RF parameters were measured with a Cascade Microtech probing system [8] connected to a Rohde&Schwartz FSEM microwave spectrum analyser. The measured frequency tuning sensitivity is 2 GHz/V at nominal supply level as can be seen in Figure 8. Frequency pushing was observed to be 47 MHz/V. This is illustrated in Figure 9. A fixed frequency comparison unit on same wafer shows only 2 MHz/V. After some consideration the reason is clear as the supply voltage has a direct connection through the choke coil to the other end of the varactor diode.

The measured oscillator output power level is -20 dBm. Highest harmonic levels were below -18 dBc. A close-up of the carrier is highlighted in Figure 10 with 300 kHz resolution bandwidth in the spectrum analyzer. At 1 MHz from the center frequency, the phase noise level was -102 dBc/Hz [9]. The entire phase noise recording is plotted in Figure 11. Here simulated values agree from 100 kHz but show too optimistic behaviour closer to carrier. Extended simulations indicate that much of the relatively high phase noise comes from the constant current source formed by a transistor pair. Obviously the very low Q of the spiral inductors impairs the noise characteristics further.

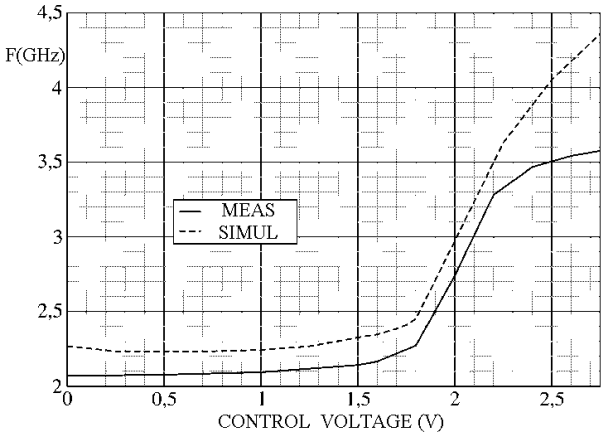


Figure 8. Measured tuning performance of the first VCO prototype. Without apparent bias error the characteristic is very similar to the simulated one up to 2.2 V.

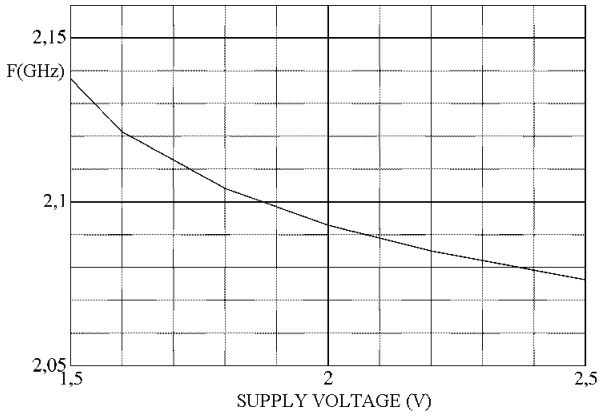


Figure 9. Measured frequency pushing in the VCO prototype. This result is apparent from the selected circuit topology.

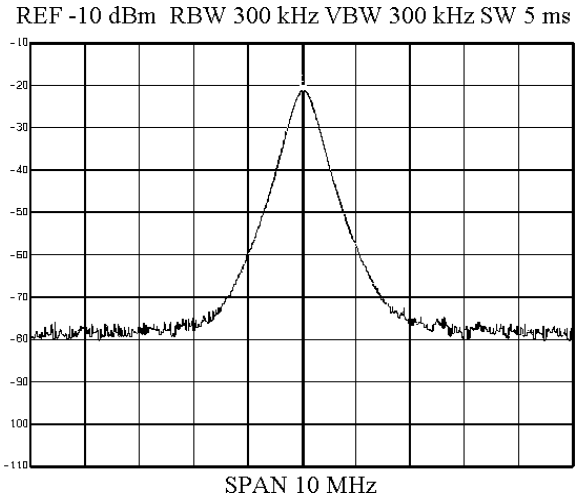


Figure 10. A close-up of the measured VCO spectrum.

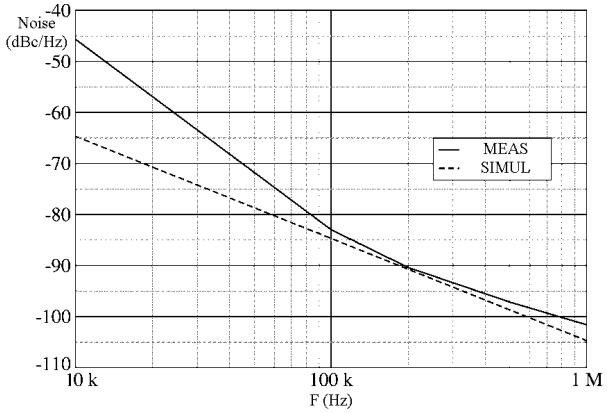


Figure 11. A comparison of measured and simulated VCO phase noise.

5. CONCLUDING REMARKS

Basically the designed and constructed first SiGe BiCMOS VCOs were very successful. Main design targets were met and useful information was obtained for coming complete quadrature output oscillator implementation. One of the greatest challenges is the low Q of substrate-based spiral coils that also require considerable area on the chip. The size can be considerably reduced by using stacked coils. However, at higher frequencies, e.g. at 5 GHz, their size is not such an issue. VCO tuning sensitivity was very high but this was partly expected as no parallel capacitors to the varactor were used. In coming designs the varactor operating point should also be optimised for highest Q. Prototype oscillators had center frequencies some 10 per cent below simulated values. Means to improve correlation should be found as no additional tuning components are desired. Phase noise simulation was found to be challenging, as in this case the simulated values agree from 100 kHz but show much too optimistic behaviour closer to carrier. Also, some discrepancies were occasionally found between results obtained with different simulation packages. These were mainly traced back to questionable parametric input and component data, not in the hands of the typical circuit designer.

REFERENCES

[1] R. Ludwig and P. Bretchko, RF Circuit Design, Upper Saddle River, Prentice Hall, 2000.

[2] P. Lacomme et al, Air and Spaceborne Radar Systems, Norwich, SciTech Publishing, 2001, pp. 190-192.

[3] S. Hackl et al, A 28 GHz Monolithic Integrated Quadrature Oscillator in SiGe Bipolar Technology, @ IEEE Journal of Solid State Circuits, Vol. 38, No. 1, pp. 135-137, January 2003.

[4] M. Kihara, S. Ono and P. Eskelinen, Digital Clocks for Synchronization and Communications, Norwood, Artech House, 2003, pp. 222-229 .

[5] D. Harame et al, Current Status and Future Trends of SiGe BiCMOS Technology, @ IEEE Trans. on Electron Devices, Vol. 48, No. 11, pp. 2575-2594, November 2001.

[6] J. Cressler, SiGe HBT Technology: A New Contender for Si-Based RF and Microwave Circuit Applications, @ IEEE Trans. on Microwave Theory and Techniques, Vol. 46, No. 5, pp. 572-589, May 1998.

[7] J. Burghartz et al, RF Circuit Design Aspects of Spiral Inductors on Silicon, @ IEEE Journal of Solid State Circuits, Vol. 33, No.12, pp. 2028-2034, December 1998.

[8] K. Ang, M. Underhill and D. Robertson, Transmission Line Stabilised Monolithic Oscillators, in the Proc. of the 14th EFTF, Turin, Italy, 2000, pp. 88-92.

[9] B. Neubig and W. Briese, Das Grosse Quarzkochbuch, Feldkirchen, Franzis' Verlag, 1997, pp. 315-339.

Phase noise modelling and optimisation of microwave SiGe amplifiers for sapphire oscillators applications

S. Gribaldo[1], R. Boudot[2], G. Cibiel[3], V. Giordano[2], O. Llopis[1]

[1] LAAS-CNRS, 7 av. du Colonel Roche, 31077 Toulouse, FRANCE
[2] Institut FEMTO-ST, dept. LPMO, 32 av. de l'Observatoire, Besançon, FRANCE
[3] CNES, 18 av. Edouard Belin, 31401 Toulouse, FRANCE

Abstract

The nonlinear modelling of different microwave SiGe bipolar transistors has been performed. Using these models, the phase noise of an amplifier is computed, taking into account two different types of noise, the microwave additive noise floor and the up-converted 1/f noise. The simulation technique combines different approaches available in a commercial CAD software. Theoretical results are then compared to the experiment, both for a single stage amplifier and a two stage amplifier. Finally, the phase noise of an optimised sapphire oscillator is calculated.

I. Introduction

Optimizing a low phase noise oscillator is still a very difficult task. One of the possible approaches is to focus on the phase noise of the amplifier included in the feedback loop, which is the main cause of phase noise in microwave oscillators. Inside the cavity bandwidth, the amplifier phase noise is simply converted into oscillator frequency noise, as described in Leeson's paper [1]. Thus, any improvement of the amplifier phase noise results directly in an improvement of the oscillator frequency noise. The advantages of studying amplifier phase noise instead of oscillator phase noise are numerous. Firstly, it is easier to simulate (and optimize) an open loop circuit (an amplifier) than a closed loop circuit (an oscillator). The frequency is indeed externally imposed in an open loop circuit, and this facilitates the convergence of the algorithm. Secondly, the phase noise measurement of an amplifier is possible from the linear behaviour (low input power) up to strong compression, whereas the oscillator can only be measured in compression [2]. Studying phase noise on such a large input power range leads to very interesting information on the device, and is the best way to check the validity of a modelling approach [3]. Moreover, there is no phase loop effect in an amplifier (this parameter is always difficult to control in an oscillator experiment).

This paper focus on the phase noise modelling of different commercially available microwave bipolar transistors (hereafter referenced as TB1 and TB2), all of silicon-germanium (SiGe) type. The transistors have first been modelled using a conventional large signal model extraction technique. Then the noise sources have been added to the device model and the phase noise has been simulated using dedicated techniques and a commercial harmonic balance software. Finally, the phase noise of various amplifiers is simulated, and the interest of some two stages circuits to get simultaneously low phase noise and high gain performance is pointed out.

II. Devices nonlinear modelling

Our models are based on the classical SPICE Gummel-Poon model [4], to which extrinsic passive elements are added in order to take into account the device fixture. The model is extracted from DC or pulsed I(V) characterization and multi-bias S parameters measurements. It is validated in linear and nonlinear regime, using respectively S parameter data and output power versus input power data at different harmonic frequencies (Figures 1 and 2). There is in both cases a good agreement between the measured and the simulated values, which allows us to go further in complexity by adding the noise parameters to the model.

III. Transistor noise modelling

Phase noise in a transistor, or an amplifier, is the results of two different processes [5]. The first one is the conversion close to the carrier of the device low frequency (LF) noise. The second one is the addition of the high frequency (HF) noise. The two processes acts very differently on the phase noise. The LF noise conversion is a multiplicative process, which means that the noise level follows the signal level. The HF noise is additive, and has thus a minimum impact on phase noise at high signal level.

Both noise processes have to be taken into account in order to be able to accurately simulate the amplifier phase noise (and also, in a second step, the oscillator phase noise).

The HF noise addition is probably easier to model than the LF noise conversion. At low input power, it can be calculated using equation (1).

$$S_\phi(f) = \frac{FkT}{P_{in}} \qquad (1)$$

where F is the amplifier noise figure, k the Boltzman constant, T the absolute temperature and P_{in} the amplifier input power. However, some problems arise at high input power. The device noise figure is increased because of device compression [5] (a nonlinear noise figure can be defined). Two approaches can then be used to model the amplifier additive phase noise : either the nonlinear noise figure has been measured in some way [5], and is used to calculate, with equation (1), the phase noise, or a nonlinear noise simulation is performed with a dedicated software (see next paragraph) and using the device natural HF noise sources (resistors thermal noise and junctions schottky noise).

Identifying the LF noise sources in a transistor is a more difficult task. Moreover, active device electrical models are often far from device physics, and finding the right location in the model of a given LF noise source is almost impossible. Locating the noise sources is a prerequisite to phase noise simulation because, in nonlinear operation, the classical equivalent noise source approach is not valid. As an example, a noise source which affects a nonlinear element will be converted, and another noise source which is at the transistor output can be very weakly converted [3]. Moreover, the noise source itself can be affected by the RF large signal, even if the time constants involved in the LF noise generation are very slow compared to the signal period. In other words, the noise does not depend only on the transistor DC conditions (even if these conditions include the rectification of the microwave large signal) [6]. The way to take into account these effects in an equivalent circuit model is still a disputed subject. The noise source can be associated to a nonlinear element of the equivalent circuit [3], or considered itself as nonlinear [7]. However, an equivalent model is by no means a rigorous representation of devices physics, and the only accurate solution to this problem could be in physical approaches [6] or microscopic models. But these models cannot be used directly to compute the noise in a complex system like an oscillator.

Our approach of the problem is a little different. It uses transistor extrinsic LF noise sources, but the effect of the large signal on these noise sources is taken into account : the device LF noise is measured under large signal conditions. This approach is not completely rigorous, because the LF noise measurement is rarely performed in the same conditions that are experienced by the device in oscillator or amplifier operation. However, it has already proven its efficiency [8], and it is simple enough to authorize a circuit optimisation with the device model.

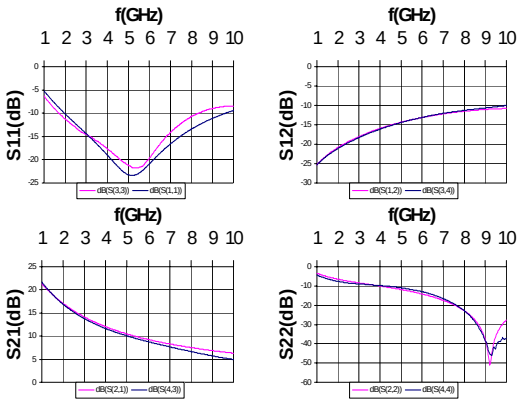


Figure 1 : Comparison between model and measurement [S] parameters of TB1 (Vce = 2V and Ic= 10 mA)

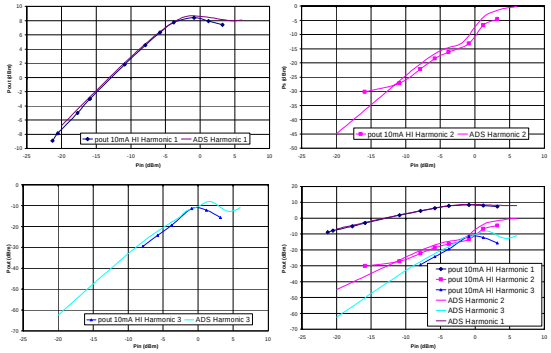


Figure 2 : Comparison between model and measurement TB1 output power P_{out} vs. input power P_{in} (Vce = 2V and Ic= 10 mA)

Two extrinsic noise sources are generally considered in a classical (linear) bipolar transistor model: the base voltage noise source and the base-emitter current noise source. The base-emitter current noise has a physical meaning. It is related to the current flowing in the base-emitter junction, and is composed of a 1/f component, proportional to this current, and a schottky noise floor. This is, in Spice models, the unique 1/f noise source in the transistor. The effect of the base current noise on phase noise is very strong. Actually, in many oscillators, the 1/f frequency fluctuations can be described entirely as a nonlinear conversion of this unique LF noise source. However, it is possible to minimize its effect using a bypass with a high value capacitance [9] or a low impedance bias network [8]. In this case, all the current noise flows into this bypass, and the nonlinearity control voltage is stabilized. What is remaining is the base voltage noise source, which takes into account the voltage fluctuations taking place at the base or the emitter electrodes, but also many other intrinsic fluctuations in the device (it is an equivalent noise source). It is this source that has been characterized, in our modelling approach, versus the microwave input power. The device, used in common emitter configuration and loaded onto 50 Ω, is fed by a microwave signal at 3.5 GHz and its low frequency

behaviour is analyzed through bias Tees in a conventional way (i.e. using a low noise amplifier on the collector access).

Figures 3 shows this type of measured data for a SiGe bipolar device. The equivalent input voltage noise is measured from linear (or even static) condition up to more than 10 dB compression. A sudden increase of the noise is observed when the device enters into compression. This effect has to be taken into account in the device model (if not, the simulated phase noise would be lower of 10 dB to 15 dB compared to the measured phase noise).

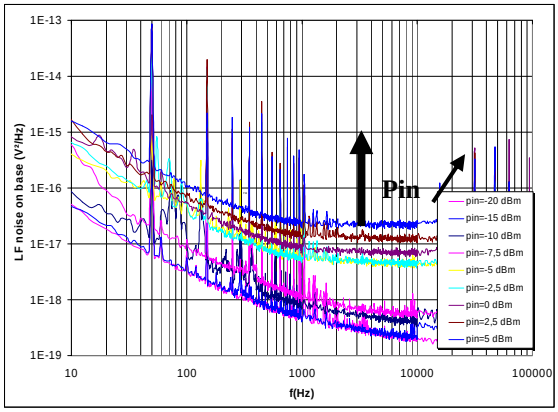


Figure 3 : LF equivalent input voltage noise spectral density ; device TB1, with input microwave power varying from 0 (quiescent device) up to +5 dBm (Ic = 10mA, Vce = 2 V)

In order to model the LF noise increase with the microwave power, two empirical functions of the microwave power P_{in} have been used : $k_1(P_{in})$ and $k_2(P_{in})$. The voltage noise dependence on P_{in} is described using the following formula :

$$S_V(P_{in}) = (S_{V\,1/f} + k_1(P_{in}) * S_{V\,floor}) * (1 + k_2(P_{in})) \quad (2)$$

$Sv_{1/f}$ and Sv_{floor} being the spectral power densities (1/f and noise floor) measured on the quiescent device.

Fitting the noise spectra (Figure 4) leads to the following expressions for these two functions :

$$k_1 = a_1 * \left(e^{\left(\frac{P_{in}}{b_1}\right)} - \frac{1}{c_1}\right) \qquad k_2 = a_2 * \left(e^{\left(\frac{P_{in}}{b_2}\right)} - \frac{1}{c_2}\right) \quad (3)$$

which are depicted in Figure 5.

Adding this “nonlinear” noise source to the transistor nonlinear electrical model allows us to simulate both amplifier and oscillator phase noise.

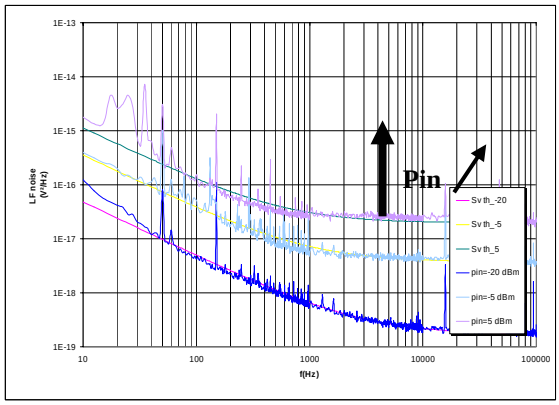


Figure 4 : LF equivalent input voltage noise spectral density ; device TB1 ; measured and modeled

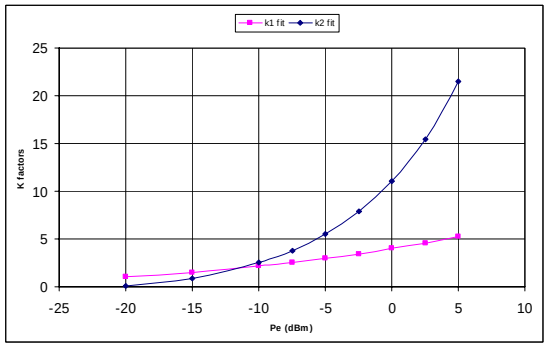


Figure 5 : k_1 and k_2 vs. 3.5 GHz input power variations, (Ic = 10mA, Vce = 2 V) – device TB1

IV. Phase noise simulation

The above described model is implemented on a commercial software : Agilent ADS. Various approaches may be used on ADS to simulate phase noise. However, many of these tools are restricted to oscillator simulation (“pnmx” and “pnfm”) and special techniques must be implemented to simulate an amplifier phase noise. The first one is the quasi-static perturbation technique, which consist in introducing a small static voltage (or current) shift to evaluate the effect of a LF voltage (or current) noise on the phase of the microwave signal through the amplifier (this is the equivalent of the ADS “pnfm” technique in oscillator simulation). The second one takes advantage of the ADS “envelope simulator”, which allows to simulate the effect of a low frequency modulation on a nonlinear amplifier. The third one uses directly the nonlinear noise modules available in ADS. However, the last technique seems to be restricted to the simulation of the amplifier noise floor, i.e. the additive phase noise, and the 1/f phase noise must be simulated by one of the two other techniques.

Providing such a two step process is realized i.e. simulation of the conversion noise using a quasi-static phase perturbation (or,

equivalently, the envelope technique) and simulation of the additive noise using the ADS nonlinear noise tools or, as an alternative, the measured nonlinear noise figure and equation (1), it is possible to reach a very good agreement between measured and simulated phase noise. Figure 6a and 6b represents such a comparison, for a SiGe bipolar device maintained between two 50 Ω loads and feed by a 3.5 GHz signal.

The device residual phase noise is measured using previously described techniques [5,8,10], which minimize the effect of the source noise (and particularly, of its AM noise). The relative increase of the phase noise floor at low carrier level is typical of an additive noise (equation (1)). The 1/f phase noise is almost constant at high power level, and increases at very low input power (-20 dBm). The phase noise is simulated from a device nonlinear model similar to the one described in paragraph III and using ADS and the above described two step simulation (a different technique for the additive and the multiplicative noise contributions). The simulated data fit well the measured data on the whole input power range.

With such a modelling approach, it is possible not only to predict an oscillator phase noise, but also to optimize the final result. The optimization of the amplifier is easier than the one of the oscillator, for already mentioned reasons (convergence). Moreover, some simple rules may help in designing a low phase noise oscillator with this technique. There is actually a trade-off to find between the gain requirements and the amplifier phase noise. It has been found in previous studies that matching a transistor for high gain performance results in a poor phase noise performance [9]. On the contrary, a resonator requires a sufficient decoupling to reach a high enough loaded Q factor (Q_L). Finally, a 6 dB losses coupling ($Q_L = Q_0/2$) can be considered as an optimum coupling both for additive phase noise [11] or conversion phase noise [8] contributions, in a single stage amplifier oscillator. This means that the amplifier gain should be higher than 7 dB or, more realistically to take into account the circuit losses, 9 dB. This gain performance is easy to reach at moderate frequencies (up to 5 GHz, as an example), giving many possibilities to optimize the transistor phase noise versus RF loading, but is more difficult at higher frequencies without matching the device for maximum small signal gain (which again, is far of being an optimized load versus phase noise performance). Another important case is the one of sapphire oscillators, and particularly of cryogenic oscillators. In this case, the resonator is reached through long cables, in order to prevent it from being heated. These cables induce extra losses, and the use of an amplifier featuring a small signal gain of at least 14 dB becomes mandatory.

We found that, both for 10 GHz applications and for cryogenic 5 GHz to 7 GHz applications, it was impossible to get optimized results by keeping a single stage amplifier. Therefore, a low phase noise two stages amplifier has been studied.

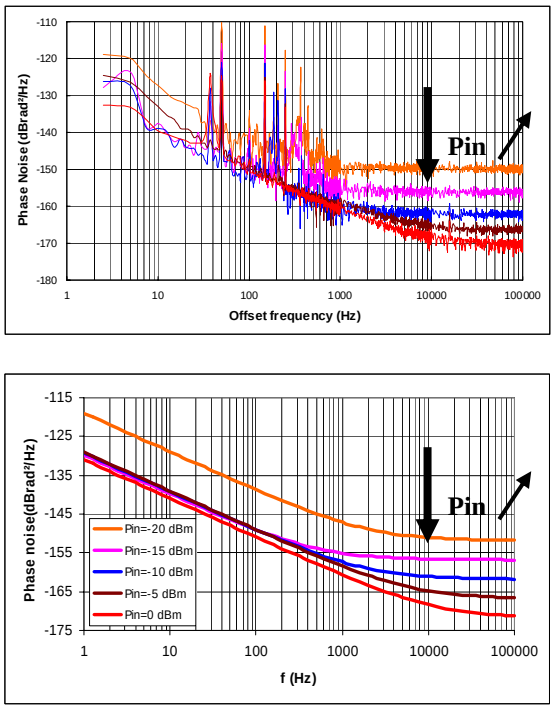


Figure 6 (6 a and 6b) : Measured (above, 6a) and simulated (below, 6b) 3.5 GHz phase noise of a bipolar transistor (TB2) at different microwave input power levels ranging from -20 dBm up to 0 dBm.

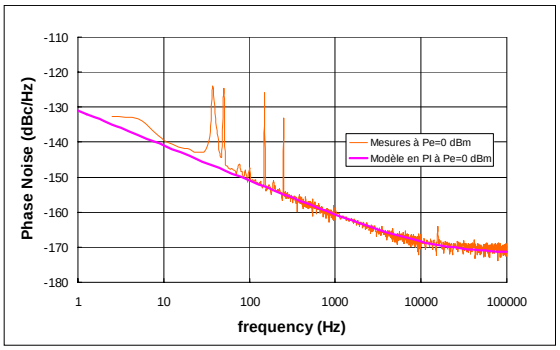


Figure 7 : Phase noise simulation – measurement comparison at high power (0 dBm) ; TB2 at 3.5 GHz.

V. Two stages low phase noise amplifier and cryogenic oscillator simulation

Different combinations of the modeled devices have been studied in order to design a two stages amplifier. If two identical stages are considered, the overall phase noise should be increased of about 3 dB, if the phase noise of each stage is not too sensitive to the input power level.

This is probably true for the 1/f phase noise, as shown, as an example, in Figure 6. But this is not true for the additive noise floor, which improves relatively to the signal at high level. Finally, the phase noise of the two stages amplifier should be 3 dB higher than the one of the single stage amplifier in the 1/f region, and should be imposed by its first stage in the noise floor region.

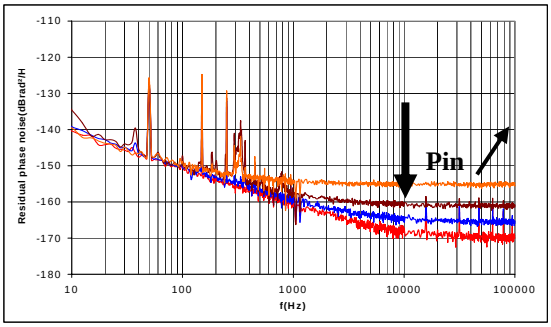


Figure 8 :3.5 GHz measured residual phase noise of two cascaded TB2, maintained onto 50 Ω load (Pin =-15 dBm,-10 dBm, -5dBm and 0 dBm).

The measured phase noise of such an amplifier, realized with two TB2 devices, is depicted in Figure 8, for different input power levels. The measured phase noise is very close to the one of the single stage device (Figure 6), and is mainly imposed in the case of this transistor by the first stage (both 1/f and additive phase noises are improved at high power).

Therefore, we have chosen this approach to design a low phase noise two stages amplifier for cryogenic applications. The gain-phase noise trade-off is analysed and optimized on each stage, using a single stub approach such as in previous one stage amplifier design [8]. The transistors are biased for low 1/f noise i.e. using a low impedance bias network on the transistors bases [8]. The complete circuit is still not yet available, but the oscillator simulated results are very promising, as shown in Figure 9. This final simulation has been performed on ADS, using the oscillator phase noise dedicated tools of this software.

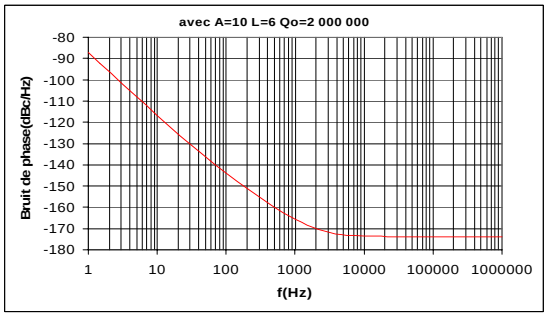


Figure 9 : Simulated phase noise of an oscillator with a two stages TB1 amplifier and cryogenic sapphire resonator @ 5 GHz (Ic = 10mA, Vce =2V)

Conclusion

A modelling technique, dedicated to microwave amplifier phase noise calculation, has been presented. This technique takes into account the two main noise contributors to phase noise in silicon bipolar transistor amplifiers: the 1/f converted low frequency noise and the additive high frequency noise. CAD calculation of each noise type is presented, taking into account the nonlinear effects which change both the device low frequency noise and the device high frequency noise figure when it is pushed into compression. This model compares well to the experiment on various single stage microwave amplifiers. It is also used to optimise a two stages amplifier dedicated to cryogenic sapphire oscillator applications. Preliminary simulated results for this oscillator are given.

References

[1] *D.B. Leeson, "A simple model of feedback oscillator noise spectrum," Proc. Letters of IEEE, Vol. 54, N° 2, pp. 329-330, 1966.*

[2] *O. Llopis, et al, "Low level and reflection phase noise measurements," IEE Electron. Lett., Vol. 37, N° 2, pp. 126-129, 2001.*

[3] *O. Llopis, et al, "Nonlinear noise modelling of a PHEMT device through residual phase noise and low frequency noise measurements," Proc. of the IEEE MTT Symposium, pp. 831-834, May 2001*

[4] *G. Massobrio, P. Antognetti, "Semiconductor Device Modelling with SPICE" McGraw-Hill Ed.*

[5] *G. Cibiel, L. Escotte, O. Llopis, "A study of the correlation between HF noise and phase noise in low noise silicon based transistors", IEEE Transactions on Microwave Theory and Tech., vol 52, n° 1, Janvier 2004, pp. 183-190.*

[6] *F. Bonani et al. "Noise source modelling for cyclostationary noise analysis in large signal device operation", IEEE Trans. on Electron. Dev., n° 9, sept 2002, pp. 1640-1647.*

[7] *H.J. Siweris, B. Schiek, "A GaAs FET oscillator noise model with a periodically driven noise source", 1986 EuMC Proceedings, pp. 681-686.*

[8] *G. Cibiel, et al, "Optimization of an ultra low phase noise sapphire - SiGe HBT oscillator using nonlinear CAD", IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 51, No. 1, pp. 33-41, January 2004.*

[9] *M. Regis, O. Llopis, J. Graffeuil, "Nonlinear modeling and design of bipolar transistors ultra low phase noise dielectric resonator oscillators" IEEE Trans. on MTT, Vol. 46, N° 10, pp. 1589-1593, October 1998.*

[10] *G. Cibiel, et al, "AM noise impact in low level phase noise measurements," IEEE Trans. on UFFC, Vol. 49, N° 6, pp. 784-788, 2002.*

[11] *J. K. A. Everard, "Low noise oscillators" Proc. of the IEEE MTT-Symposium, pp. 1077-1080, 1992.*

Design and Characterization of New Microwave Synthesizers for Metrological Applications.

D. Chambon, S. Bize, F. Chapelet, F. Narbonneau, M. Lours, A. Clairon, and G. Santarelli
BNM-SYRTE, Observatoire de Paris, 61 Avenue de l'Observatoire, 75014 Paris, France

M. Tobar
The University of Western Australia, School of Physics, 35 Stirling Hwy, Crawley, Western Australia.

Keywords - **Sapphire cryogenic oscillator, atomic fountain, frequency synthesis chain, phase noise, frequency stability**

Abstract - **This paper describes the recent improvements on the local oscillator, used as reference for the various experiments at SYRTE. The new system, delivering an high stability and low phase noise signal at 1 GHz shows better performances than the previous one, operating at 100 MHz. In order to explore the new facilities offered by this local oscillator, frequency synthesizer used as probe signal for the atomic fountain, have been redesigned. Accuracy impact evaluation of the synthesizers driving atomic fountain is also presented.**

I. INTRODUCTION

At SYRTE, the reference signal is obtained by the combination of a Sapphire Cryogenic Oscillator (SCO), an H-Maser through a set of ultra low noise frequency synthesizers: we commonly call this system "Local Oscillator". The SCO operates continuously for 25 days between each LHe refill, and delivers an ultra low phase noise microwave signal at 11.932 GHz. The frequency stability (Fig. 1) of this oscillator is estimated to be $\sigma_y(\tau) \leq 1-2\times10^{-15}$ for $1\ s < \tau < 1000\ s$ with a drift of about 1.5×10^{-13} per day. The H-Maser reaches the 10^{-15} level stability from about 1000 s.

This local oscillator signal is distributed to all of SYRTE metrology and physics applications. Figure 2 shows the dissemination of this reference local oscillator: three atomics fountains: FO1 [1], FOM (the transportable fountain) both based on ^{133}Cs and a dual fountain (FO2)[2, 3] operating with both ^{133}Cs and ^{87}Rb. The reference signal is also distributed to inertial sensor experiments (gyroscope and gravimeter) which need an ultra low noise signal to reach high sensitivity. Low phase noise and ultra stable microwave sources are also used to drive mode-locked femtosecond lasers in high resolution frequency comparisons between microwave and optical clocks [4]. The dissemination via a fibre optic link of 44 km has also allowed the comparison between an optical frequency standard operating at 30 THz (CO2 laser stabilized on OSO4 absorption) with a microwave frequency standard (SCO of SYRTE) [5].

A first synthesizer converts the microwave signal (11.932 GHz) delivered by the SCO to 11.98 GHz and 100 MHz. These signals are phase coherent and slowly phase-locked on a H-Maser which guarantees the long term stability. The proximity of the Rb/Cs fountain FO2 allows generating directly the Cs hyperfine clock transition frequency at 9.192 GHz from the 11.98GHz without significant degradation [6]. For all the distant applications (> 20 m from the SCO) demanding metrological reference (FOM, FO1, optical frequency standard with cold Strontium atoms and OsO_4 and SF_6 at LPL), the 100 MHz signal degraded by 15 dB comparing to the 11.98 GHz signal, was distributed via RF cables. From this metrological signal at 100 MHz, we designed frequency synthesis chains to generate the 9.192 GHz signal. With this scheme, we obtained on the FO1 fountain a stability of 2.2×10^{-14} at 1 s. The power spectral density of frequency fluctuations Sy(f) of this local oscillator was a significant limitation for the stability of the FO1 fountain. The work described in this paper has been done with the aim of overcoming this limitation: we increase the distribution signal from 100 MHz to 1 GHz. This frequency can still be distributed with RF cables over 100-200 meters with reasonable insertion losses (35dB/200m).

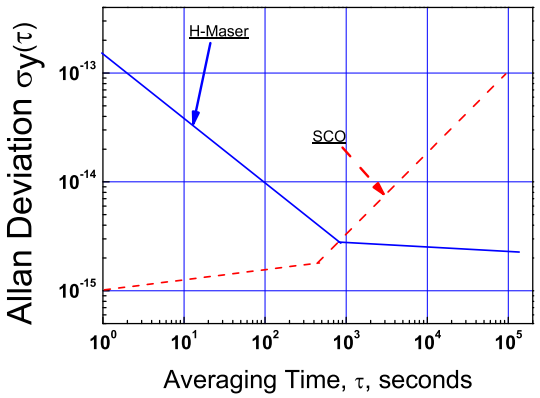


FIG. 1: Allan deviation of the SCO and the H-Maser

II. DESCRIPTION OF THE DOWN CONVERTER SYNTHESIS CHAIN DELIVERING THE REFERENCE SIGNAL AT 1 GHZ

Our goal is to convert the microwave signal of the SCO at 11.932 GHz to a 1 GHz signal without any

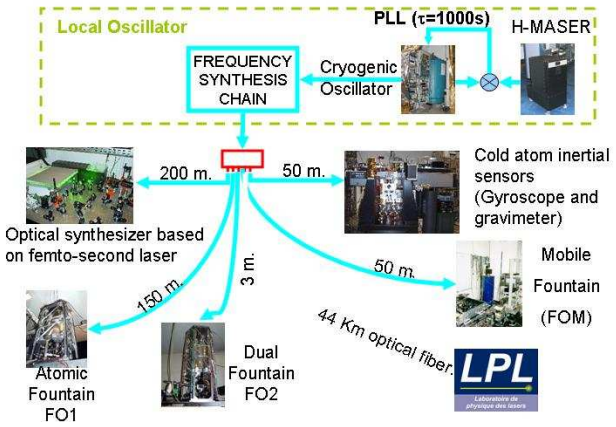


FIG. 2: Distribution schematic of the metrological signal

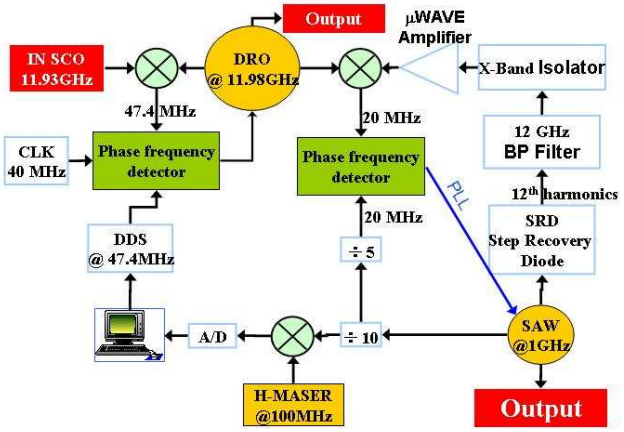


FIG. 3: Schematic of the down converter synthesis chain at 1GHz

significant degradation of the SCO's phase noise. Additionally, this new reference signal must be phase coherent with the H-Maser for the long term stability. From the free-running SCO output signal, we produce a microwave signal at 11.98GHz (Fig. 3) by offset phase locking a Dielectric Resonator Oscillator (DRO) via a Direct Digital Synthesizer (DDS). This signal is slightly tunable and nearly reproduces the SCO phase noise performances. In order to transfer the low phase noise and high short term frequency stability of the SCO, we need to bridge the gap between 11.98 GHz down to 1 GHz. Fig. 3 describes the synthesizer used for this conversion. A Surface Acoustic Wave (SAW) oscillator signal at 1 GHz is frequency multiplied by using a Step Recovery Diode (SRD). This SRD generates all the harmonics of the 1 GHz signal, we select the 12th harmonics using a band pass filter at 12 GHz. These two microwave signals at 11.98 GHz and 12 GHz drive the LO and RF ports of a mixer. The IF signal at 20 MHz is compared, via a phase/frequency detector, to a 20 MHz signal obtained by frequency dividing the SAW output by 50. The zero beat signal is used to phase lock the 1 GHz oscillator on the 11.98 GHz signal with a bandwidth of about 4 kHz. Finally, both signal the 1 GHz SAW and the 11.98 GHz DRO needs to be phase coherent with the H-Maser; we use the baseband beat signal of the SAW divided by 10 and the H-Maser output to perform, via a computer, a slow digital phase lock loop ($\tau = 1000s$) controlling the DDS frequency. We use a signal at 40 MHz, which is phase coherent with the 1 GHz oscillator, as a clock signal for the DDS.

A. Experimental results

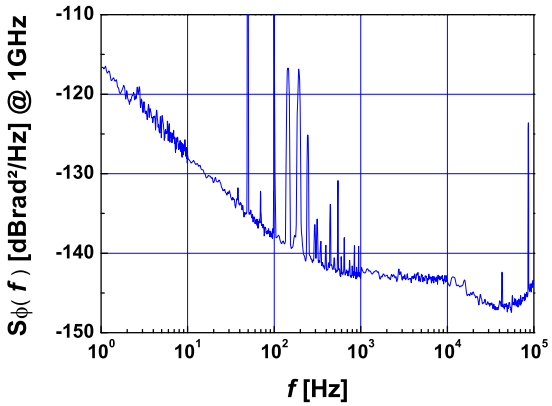


FIG. 4: Phase noise power spectral density of the down converter synthesis chain at 1GHz

To measure the various performances of this new 1 GHz reference, we build two identical synthesizers and we beat their output signals at 1 GHz. The resulting beat-note signal is used to characterize the phase noise and the phase stability. Fig. 4 shows the phase noise spectral density of one device measured at 1 GHz. We observe a flicker noise at 1 Hz from the carrier which is around $S_\phi(f) = -117$ dBrad^2/Hz, with a white phase noise floor of -145 dBrad^2/Hz. From this measurement, we deduce that there is no significant degradation of the SCO in term of phase noise.

The phase stability is represented in term of Modified Allan deviation in Fig. 5. We obtain a $Mod\sigma_y(\tau) = 1.2 \times 10^{-15}$ at 1 s $\tau = 1$ s and a level at 10^{-18} for 20000 s of averaging time.

According to this new reference signal, we redesign new high performances synthesizers which generate the 9.192 GHz signal to drive our atomic fountains.

III. A NEW LOW NOISE SYNTHESIZER GENERATING THE 9.192 GHZ FOR THE CESIUM FOUNTAINS.

A. Hardware description

Fig. 6 represents the block diagram of the synthesizer which generates the clock transition signal at 9.192 GHz from the reference signal at 1 GHz. This

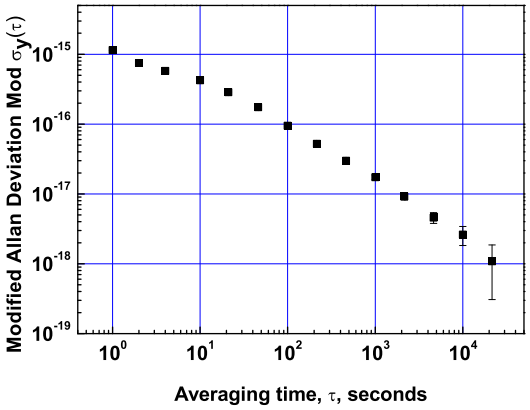


FIG. 5: Fractional Phase Stability of the down converter synthesis chain measured at 1GHz

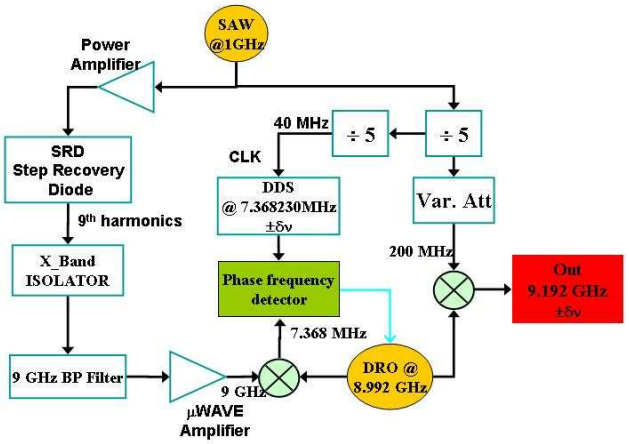


FIG. 6: Synthesis chain generating a 9.192 GHz signal from the new reference at 1 GHz

1 GHz output is fed to a step recovery diode which generates a frequency comb. We select the 9th harmonics with a band pass filter at 9 GHz. To phase lock a Dielectric Resonator Oscillator at 8.992GHz, we mix the two microwave signals at 9 GHz and 8.992 GHz. The resulting beat-note signal at 7.368 MHz is phase compared to a DDS signal to phase lock the DRO via an appropriate loop filter. The frequency clock of the DDS is obtained by dividing the 1 GHz signal by 25. Finally, we mix the DRO's output signal (+7 dBm) at 8.992 GHz with a 200 MHz signal (about -40 dBm), obtained by dividing the 1 GHz signal. The resulting low power (-60 dBm) 9.192631770 GHz signal is well frequency shifted from the DRO output in order to strongly reduce the microwave leakage to the fountain clock. This signal can be frequency tuned with a resolution of 1 μHz via the DDS and the output level can be adjusted using a voltage control attenuator at 200 MHz.

B. Experimental results

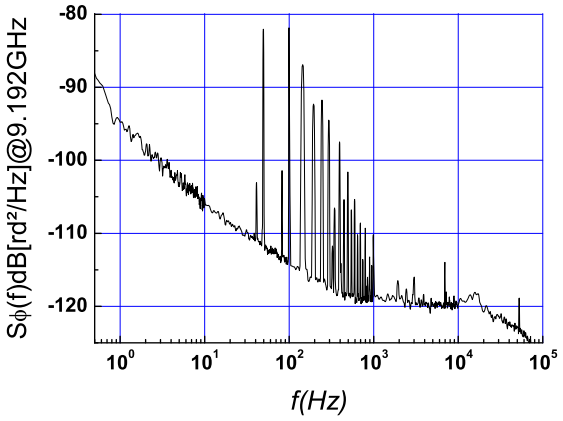


FIG. 7: Phase noise power spectral density of the synthesis chain at 9.192 GHz

To evaluate the phase noise and the phase stability of the previously described synthesizer, we build two identical devices. Fig. 7 shows the phase noise spectral density of one device measured at 9.2 GHz. At 1 Hz from the carrier, we have $S_\phi(f) = -96$ dBrad^2/Hz, with a white phase floor of -120 dBrad^2/Hz which is compatible with the goal to do not significantly degrade the phase noise of the SCO.

A Modified Allan deviation of the phase stability is plotted on the Fig. 8 We obtain $Mod\sigma_y(\tau) = 1.2 \times 10^{-15}$ for $\tau = 1$ s of integration time and a level $\leq 10^{-18}$ for 3 hours of averaging time which is compatible with the accuracy target of 10^{-16} for atomic fountains.

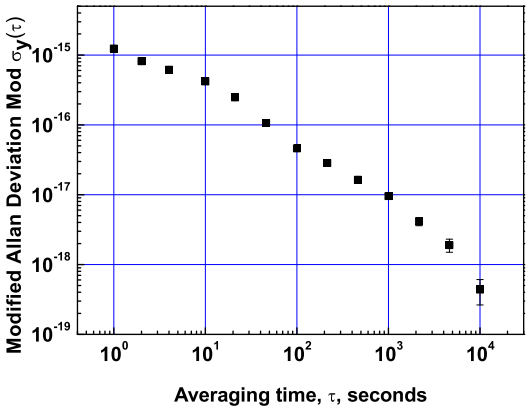


FIG. 8: Phase stability of the synthesis chain at 9.192 GHz

The previous measurement do not fully evaluate the possible frequency shifts of the atomic resonance, induced by the synthesizers imperfections. For instance microwave leakage, asymmetric spurious sidebands and the phase perturbations synchronous with the fountain

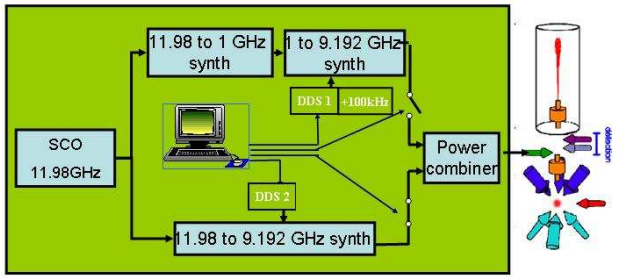


FIG. 9: Differential measurements set-up

cycle needs to be evaluated "in situ" comparing two synthesizers with a very different design. This measurement is done by driving the atomic fountain alternatively with the two synthesizers and the frequency shift evaluating by a differential measurement method described in the following.

IV. FREQUENCY SHIFT EVALUATION WITH A DIFFERENTIAL MEASUREMENT.

The differential measurement method is commonly used to evaluate frequency shift between two fountain configuration which can be sequentially applied with a measurement frames shorter than 100 s. In this way, the long term local oscillator frequency instability is rejected and the measurement frequency stability is optimized. This differential method has already been used in [7, 8]. In our case, frequency shifts not related to the synthesizers, are cancelled (black body, collision shift, quadratic Zeeman effect...). Fig. 9 shows the measurement set-up. The atomic interrogation signal is generated in two different ways: using the previous described devices (1 GHz down converter and the 9.2 GHz signal synthesizer) and a direct conversion 11.98 GHz to 9.2 GHz synthesizer described in [6]. This two synthesizers, with a completely different design, are likely not to have the same spurious sidebands content, phase perturbation and microwave leakage. Thus the differential measurement lead to a realistic estimation of the total synthesizer related frequency offsets.

The measurement sequence is done by 50 fountain cycles (about 65 s) for each synthesizers, interlaced by 5 s of blanking time. When one synthesizer is used to probe the atoms, the other one is switch off (100 dB output attenuation) and frequency offset of 100 kHz from the atomic resonance.

The frequency stability obtained with both synthesizers is 1.6×10^{-14} (see Fig. 10). This outstanding performance, which is continuously achieved by the FO2 fountain, allows a stringent measurement in a very short time. In this way, the evaluation of the frequency difference reach a resolution slightly below 10^{-16} in about one day measurement. The measured frequency offset between the two configurations is 6.5×10^{-17} fully consistent with 0 within the error bars. This result proves the synthesizers performance in term of frequency offset.

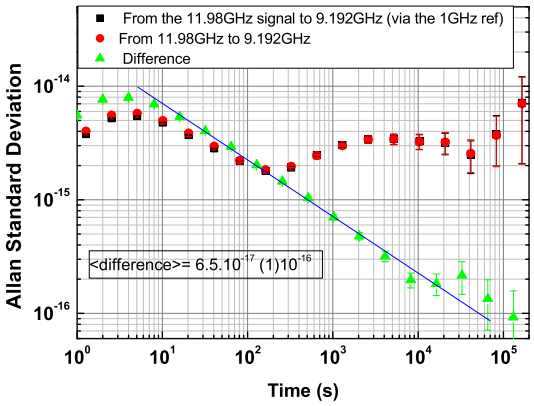


FIG. 10: Synthesizers differential measurements

V. CONCLUSION AND PROSPECTS.

The new reference signal at 1 GHz is fully operational with high performances. The synthesis chain, which generate this 1 GHz signal from the SCO's signal at 11.932 GHz , has been fully characterized in term of phase noise and phase stability. We can now disseminate, without any significant degradation of the SCO's performances (1.2×10^{-15} for $\tau = 1\ s$), this 1 GHz signal to all the experiments of the laboratory and to distant experiments such as the Laboratoire des Physiques des Lasers (LPL), using an optical fiber [9].

With this improved reference signal, we developed new synthesizers to generate the Cs clock transition frequency at a very low phase noise and phase stability level. With this synthesizer, the frequency stability of the atomic fountain is limited by the atomic projection noise [10] to about 1.6×10^{-14}.

A differential measurement test has been carried out, using the FO2 fountain, with two different microwave synthesizers. The observed frequency difference between the two synthesizers, is consistent with 0 with a resolution measurement of 1×10^{-16}. We can conclude that there is no accuracy limitation from the synthesizers at a level of 1×10^{-16}. Despite this good results, we will further investigate for possible frequency offset by driving the fountain with higher power levels. The putative daily perturbation which degrades the frequency stability from $50 \times 10^3\ s$ to $100 \times 10^3\ s$ needs to be further investigated.

[1] P. Rosenbusch *et al.*, Proceeding EFTF 2004.
[2] S. Bize *et al.*, in *Proc. of the 6th Symposium on Frequency Standards and Metrology* (World Scientific, Singapore, 2001), p 53.

[3] H. Marion *et al.*, Proceeding EFTF 2004.
[4] I. Courtillot *et al.*, Opt. Letters. **28** 468 (2003).
[5] F. Narbonneau *et al.*, Proceeding EFTF 2004.
[6] D. Chambon *et al.*, Proceeding 18th European Frequency and Time Forum. University of Surrey, Guildford, UK (2004).
[7] F. Pereira *et al.*, Physical Review Letters **89** Number 23 (2002).
[8] Y. Sortais *et al.*, Physical Review Letters **85** p. 3117 (2000).
[9] A.Amy-Klein *et al.*, "Long-distance frequency dissemination with a resolution of 10^{-17}" in Physical Review Letters. Accepted to be publish.
[10] G.Santarelli *et al.*, "Quantum projection noise in an Atomic Foutain: A high stability Cesium Frequency Standard" in Physical Review Letters June 1999

19th European Frequency and Time Forum, Besançon, France

Session 5D / Measuring methods and instrumentation

CALIBRATION OF FOUR EUROPEAN TWSTFT EARTH STATIONS WITH A PORTABLE STATION THROUGH INTELSAT 903

D. Piester[1], R. Hlavać[2], J. Achkar[3], G. de Jong[4], B. Blanzano[5], H. Ressler[5], J. Becker[1], P. Merck[3], O. Koudelka[5]

[1]Physikalisch-Technische Bundesanstalt (PTB), Bundesallee 100, 38116 Braunschweig, Germany (dirk.piester@ptb.de)
[2]National Physical Laboratory (NPL), Teddington, United Kingdom (roman.hlavac@npl.co.uk)
[3]BNM-SYRTE, Observatoire de Paris (OP), Paris, France (joseph.achkar@obspm.fr)
[4]NMi Van Swinden Laboratorium B.V. (VSL), Delft, Netherlands (gdejong@nmi.nl)
[5]Joanneum Research GmbH and Technical University Graz (TUG), Graz, Austria (otto.koudelka@joanneum.at)

Keywords: TWSTFT, time transfer, calibration, uncertainty evaluation, Galileo.

ABSTRACT

The differential delays of four two-way satellite time and frequency transfer (TWSTFT) earth stations were determined by using a portable TWSTFT station. This station was assembled by TUG and visited the time laboratories of PTB, OP, NPL, and VSL from 5th to 16th July 2004. A number of calibration measurements was performed during a four hours slot at each location. These measurements were supplemented by differential measurements between the portable and the co-located local stations. Differential delays between the portable and co-located earth stations show a statistical uncertainty below 0.6 ns for a standard TWSTFT measurement. The final closure measurement at PTB allows a stability analysis of the differential delay between the portable and the local station. The deviation between the two co-locations is only 0.4 ns. We achieved total estimated uncertainties down to 0.9 ns. As a further test the results were checked for closing errors and also against previous calibration results.

1. INTRODUCTION

Two-way satellite time and frequency transfer (TWSTFT) is one of the leading techniques for remote frequency standards and time scales comparison [1, 2]. Accuracies in the sub-nanosecond regime are achievable. Many systematic errors which are present in receive-only methods, such as GPS common-view time transfer, are negligible small. While unknown delay differences between the earth stations' transmitting and receiving path do not hamper comparisons of remote frequency standards it is observed that time scale comparisons even on the level of hundreds of nanoseconds are impossible without a proper calibration of the participating ground stations.

TWSTFT is being used by the Bureau International des Poids et Mesures (BIPM) to relate the Coordinated Universal Time UTC(i) scales realized in the institutes i. Up to now, most of the TWSTFT time links have been calibrated using GPS common-view measurement results which limit the attainable uncertainty to several nanoseconds. This can be clearly seen in the uncertainty values of UTC - UTC(i) published for the first time in BIPM Circular T 205 [3]. Consequently, the CCTF Working Group on TWSTFT stimulated to perform calibrations of TWSTFT links using an adequate technique. The most convenient way is the use of a portable TWSTFT station.

Beside the extensive calibration activities of the U.S. Naval Observatory (USNO) [4], in 1998 and 2003 European calibration trips were carried out by two persons visiting three or two European stations, respectively [5, 6]. We report on the calibration of four TWSTFT earth stations during a two-week schedule. The portable station needed to be conducted by only one person as the invited stations were well prepared and supported the activity. We report on the evaluation concept of the calibration and describe the execution and results of the exercise in detail, including a discussion of the uncertainty budget estimation. Finally, the possible influence of so-called closing errors is discussed.

2. THE CALIBRATION TRIP

Four European institutes agreed to have the differential earth station delays of their TWSTFT systems in Ku-band determined in a way which was successfully demonstrated one year ago for the IEN – PTB link [6]. Pseudo-random noise (PRN) phase-modulated signals were exchanged via the geostationary satellite INTELSAT 903 at 325.5°E with uplink and downlink frequencies of 14170.5975 MHz and 11120.5975 MHz, respectively. Each station transmits a pre-determined characteristic PRN signal at 2.5 MChip with Mitrex compatible codes. A standard session consists of nominal 120 time difference values (one measurement each second). The midpoint of a quadratic fit function is calculated at each station i, named TW(i) further down, and exchanged among the stations, as recommended in the ITU-R documents [7]. The travelling portable reference station (TS), designated TUG01 and first used in 1998 [5], was provided and operated by Joanneum Research on contract basis [8]. The campaign is illustrated in Fig. 1. The TS was sequentially operated at

four different European time laboratories (the Physikalisch-Technische Bundesanstalt – PTB in Germany, the Bureau National de Métrologie, Systèmes de Référence Temps Espace - BNM-SYRTE (in the following named Observatoire de Paris - OP) in France, the National Physical Laboratory – NPL in the United Kingdom, and the National Metrology Institute Van Swinden Laboratorium B. V. – VSL in the Netherlands) during a two-week schedule. The initial measurements at PTB were verified by a second series concluding the campaign.

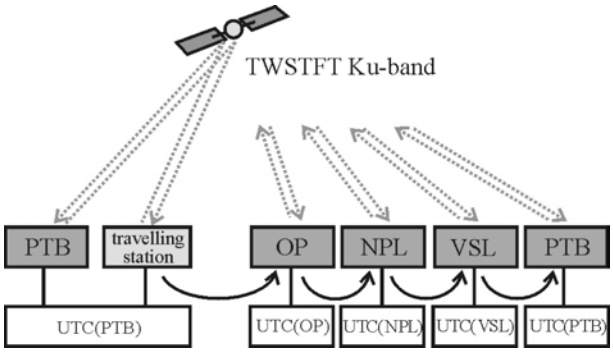


Figure 1: Schematic set-up representation of the travelling TWSTFT station (TUG01) sequentially operated at PTB, OP, NPL, VSL, and back at PTB.

The campaign was carried out from 5th to 16th July 2004. The TS was supplied with a new 4-segment antenna (diameter 1.2 m) to facilitate transport and deployment by one engineer. Most of the outdoor and indoor equipment including cables were the same as in previous calibration exercises [5, 6]. The TS modem was SATRE 036. Since the last campaign some microwave components (LNB, filter) have been replaced due to a change of the geostationary satellite and its transponder frequencies. Thus, comparisons with previous results are hampered by the unknown delay change due to the exchange of these components.

The measurements started on 6th July (MJD 53192) at PTB, followed by measurements at OP (9th July, MJD 53195), NPL (12th July, MJD 53198), VSL (14th July, MJD 53200), and again at PTB (16th July, MJD 53202). In Fig. 2 photographs of the TS beside the local outdoor equipment are depicted (a to d). There were normal weather conditions (cloudy and sometimes rainy, temperature ranging from 10° C to 22 °C) during the whole trip which had no noticable impact on the results of the campaign.

A half hour time-table was arranged between all sites and the TS to enable TWSTFT measurements in all possible combinations. This time-table was repeated seven times to a total schedule lasting four hours from 10:00 UTC to 14:00 UTC. Additional measurements before and after the schedule were recorded at all sites between the TS and the local station. At VSL the schedule was enhanced by using otherwise idle time slots of the schedule.

In Fig. 2 (e to h) the indoor setup at each site is shown. At OP, NPL, and VSL the TS indoor equipment was installed inside the respective time and frequency laboratory. At PTB, the TWSTFT antenna and transceiver of the local station are mounted on a roof top in a distance of about 200 m from the time and frequency laboratory where the local modem and automation systems are located. The TS antenna and transceiver were mounted side by side to the local outdoor equipment and the indoor set-up was located in the same building (in a office two floors below). Because no reference frequency and 1PPS were available at the TS set-up, 1 PPS and 10 MHz frequency were supplied by a caesium clock (model HP5071A, ID C9) from the PTB clock ensemble, which was brought to the TS set-up for the duration of the experiment.

3. CALIBRATION TECHNIQUE

To determine the permanent stations' individual differential delays at each side, the TWSTFT equipment was measured versus the TS, both connected to the local realization of UTC(i). This relative delay difference is called the common clock difference between the local station of laboratory i and the TS and given by

$$\begin{aligned} \mathrm{CCD}(i,\mathrm{TS}) = {} & \tfrac{1}{2}\mathrm{TW}(i) - \tfrac{1}{2}\mathrm{TW}(\mathrm{TS}) \\ & + \mathrm{REFDLY}(i) - \mathrm{REFDLY}(\mathrm{TS}) \end{aligned}, \qquad (1)$$

REFDLY(TS) is the connection of the TS's modem TX 1PPS (PPSTX(TS)) to the local UTC(i) and determined using

$$\begin{aligned} \mathrm{REFDLY}(\mathrm{TS}) = {} & [\mathrm{UTC}(i) - \mathrm{CLK}(i)] \\ & + [\mathrm{CLK}(i) - \mathrm{REF}(\mathrm{TS})] \\ & + \tfrac{1}{2}[\mathrm{REF}(\mathrm{TS}) - \mathrm{PPSTX}(\mathrm{TS})] \end{aligned}, \qquad (2)$$

where CLK(i) is the input reference clock for the TS's modem. The difference to UTC(i) has to be determined. REF(TS) is a 1PPS generated by the indoor unit of the TS and related to the modem reference frequency. Thus, REF(TS) is phase coherent to the PPSTX which is the transmission output of the modem. In the TS measurement configuration the start input of the time interval counter is not triggered by the PPSTX but by the REF(TS). This configuration introduces the factor ½ in eq. (2) and differs from the normal practise in the laboratories. In a second step, link calibration constants were computed from the differential earth station delays. Generally, delay differences between the transmission and reception paths of both, the local station and the TS, are leading to a non-zero result of the CCD.
After determination of the CCD(i,TS) at at least two sites k and l, a calibration constant for a time comparison between them can be computed using

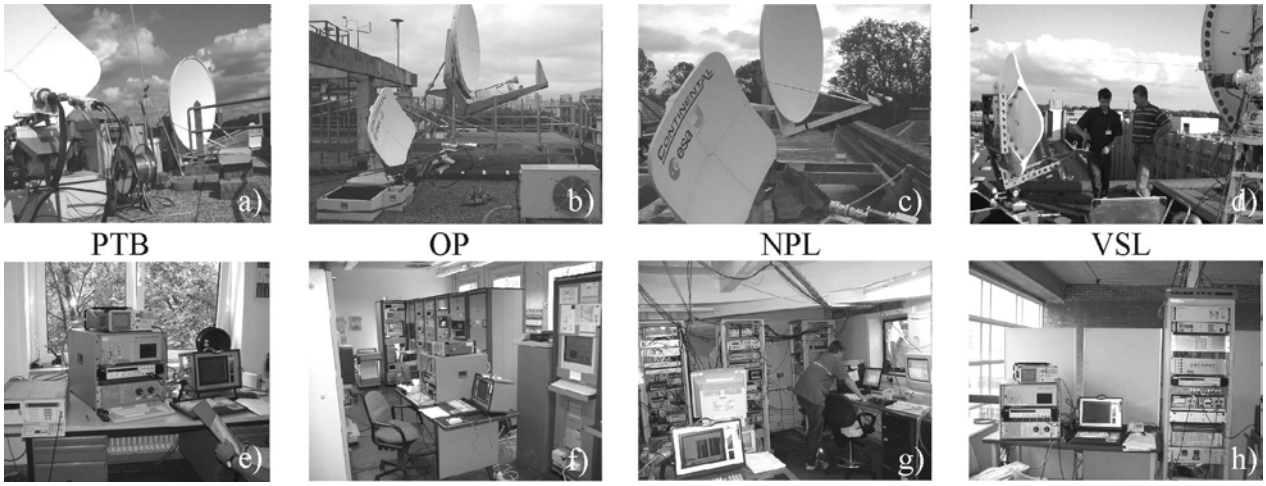


Figure 2: The TUG portable station collocated at four European TWSTFT earth stations. The upper photographs show the outdoor the lower photographs the indoor setup at PTB (a, e), OP (b, f), NPL (c, g), and VSL (d, h), respectively.

$$\begin{aligned} \mathrm{CALR}(k,l) = {} & \mathrm{CCD}(k,TS) - \mathrm{CCD}(l,\mathrm{TS}) \\ & + \mathrm{TCD}(l) - \mathrm{TCD}(k) \end{aligned} , \qquad (3)$$

where TCD(i) is the earth rotation correction (Sagnac effect) for the one-way path from the satellite to station i, calculated following Ref. [7]. Having completed this exercise, the difference of the time scales UTC(k) and UTC(l) can be further on determined in routine operations according to

$$\begin{aligned} \mathrm{UTC}(k) - \mathrm{UTC}(l) = {} & \tfrac{1}{2}\left[\mathrm{TW}(k) + \mathrm{ESDVAR}(k)\right] \\ & - \tfrac{1}{2}\left[\mathrm{TW}(l) + \mathrm{ESDVAR}(l)\right] \\ & + \mathrm{REFDLY}(k) \\ & - \mathrm{REFDLY}(l) \\ & + \mathrm{CALR}(k,l) \end{aligned} . \qquad (4)$$

ESDVAR(i) is the monitored differential earth station delay variation due to changes in cabling, etc. This value is set to zero at the moment when a new calibration value is applied.

4. RESULTS

The results of single calibration measurements between the TS and the collocated station are shown in Fig. 3. In the upper part the REFDLY(TS) measurements (label **I** to **IV**) are shown in detail:

$$\begin{aligned} \mathrm{I}: \quad & \mathrm{UTC}(i) - \mathrm{CLK}(i) + C_{\mathrm{I}} \\ \mathrm{II}: \quad & \mathrm{CLK}(i) - \mathrm{REF}(\mathrm{TS}) + C_{\mathrm{II}} \\ \mathrm{III}: \quad & \mathrm{REF}(\mathrm{TS}) - \mathrm{PPSTX}(\mathrm{TS}) + C_{\mathrm{III}} \\ \mathrm{IV}: \quad & \mathrm{REFDLY}(\mathrm{TS}) + C_{\mathrm{IV}} \end{aligned} . \qquad (5)$$

Constants C_{I} - C_{IV} have been applied for a better visualization of the data and are summarized in Table 1. **I** represents the relation between UTC(i) and the modem reference CLK(i). At OP, NPL, and VSL both signals were phase coherent, while at PTB the use of the clock C9 introduced additional noise. The corresponding time differences between UTC(PTB) and C9 during the TW sessions were calculated by using a linear regression for the first co-location (MJD 53192) and the mean value for the closure (MJD 53202).

Table 1: Constants applied in eq. (5)

	PTB_1	OP	NPL	VSL	PTB_2
C_{I} (ns)	-48660	-360	-8361	-26	-48773
C_{II} (ns)	-41	-46	-95	-92	-41
C_{III} (ns)	-731	-724	-723	-769	-731
C_{IV} (ns)	-49074	-775	-8824	-510	-49187

Measurements **II** and **III** were ideally phase coherent. Especially the data of the first visit at PTB reveal a significant slope, which we attribute to warm up processes in the modem. One may recognize the same but very slight slope in the second PTB data. **IV** is the REFDLY(i) (see eq. (4)) taken to calculate the CCD(i) after eq. (1). CCD(i) is shown in the lower part of Fig. 3. The black dots represent the results of the individual measurements, the grey dots are the mean over all data of one day. The error bars represent the standard deviation. PTB_1, OP, and VSL data show slight drifts over the total measurement time, which do not exceed one nanosecond at least for PTB_1 and VSL. The data of OP and NPL show a little bit more noise. The PTB_2 data reveal a local minimum which may be due to clock drifts of the TS reference. The calibration results shown in Fig. 3 are summarized in Table 2. The standard deviation (sigma) of the CCD(i) is always below 0.6 ns. In the last column the Sagnac correction TCD(i) is given.

Table 2: Calibration campaign results

Laboratory	CCD (ns)	sigma (ns)	TCD (ns)
NPL	-824.08	0.57	76.298
OP	6998.16	0.43	86.103
PTB	+7.20	0.31	94.295
VSL	-29.12	0.23	84.324

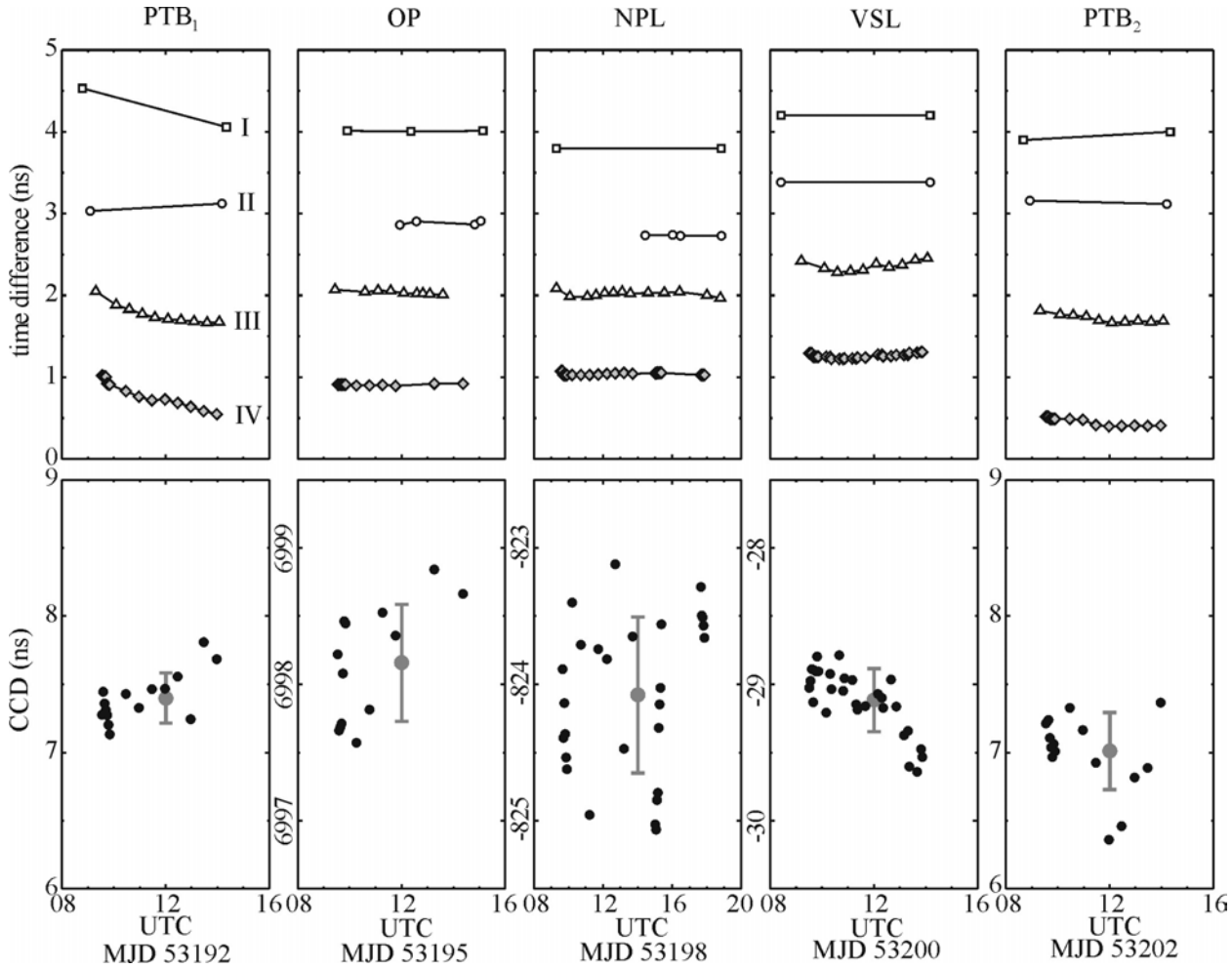


Figure 3: REFDLY measurements of the TS (upper part, for details see the text) during the calibration trip from the 1st visit at PTB (left) to the 2nd (right). In between the measurements recorded at OP, NPL, and VSL. The lower part shows the single measurements (black dots) as well as the means and standard deviation (grey symbols) of the common clock difference CCD between the TS and the local station.

The resulting calibration constants are shown in Table 3. The overall uncertainty of the calibration constants can be calculated using the following equation:

$$U = \sqrt{u_{A,k}^2 + u_{A,l}^2 + u_{B,1}^2 + u_{B,2}^2 + u_{B,3}^2} \,. \qquad (6)$$

$u_{A,i}$ is the standard deviation of CCD(i) determined at laboratory i. Ideally the determination of the two CCD(i) for one link calibration should be performed simultaneously. In practice this is not possible. To account for potential delay variations of the three stations, which are involved for one link calibration, we estimate them from the two experiments TS at PTB, the initial and the closure. The difference between both measurements is 0.39 ns. If this difference is shared equally between both stations every single station shows the same instability, and thus contributes with 0.28 ns. Because three stations are involved for one link calibration the total amount is $u_{B,1}$ = 0.48 ns. The TS has to be related to the local UTC(i) which requires a measurement of the UTC(i) reference with the TS's time interval counter for REFDLY(i) determination. We have to account for this by applying $u_{B,2}$ = 0.5 ns according to the time interval counter specifications. $u_{B,3}$ reflects all other systematic errors, e.g. the stability of the connection to the local UTC (0.1 ns), influence of code changes, Tx and Rx power, C/N_0 (overall 0.2 ns). PTB used a portable clock C9 to connect the TS to UTC(PTB). Thus additional 0.3 ns should be assumed for links where PTB is involved. The total estimated uncertainty is just at the border to the sub-nanosecond region (see Table 3).

5. COMPARISONS WITH PREVIOUS CALIBRATIONS

Except the links including OP, which TWSTFT equipment was recently established and thus had not been calibrated, all others had been initially calibrated with Circular T (i.e. relaying on GPS measurement and calibration) [9]. To compare the present CALR values (column 2 in Table 3) with old ones, one has to take into account the values of the monitored differential earth

Table 3: Applied calibration constants and uncertainty budget evaluation. For details see the text.

Link	CALR(k,l) (ns)	$u_{A,k}$ (ns)	$u_{A,l}$ (ns)	$u_{B,1}$ (ns)	$u_{B,2}$ (ns)	$u_{B,3}$ (ns)	U (ns)
NPL - OP	7832.05	0.57	0.43	0.48	0.50	0.22	1.02
NPL - PTB	849.28	0.57	0.31	0.48	0.50	0.37	1.02
NPL - VSL	802.99	0.57	0.23	0.48	0.50	0.22	0.95
OP- PTB	-6982.77	0.43	0.31	0.48	0.50	0.37	0.95
OP - VSL	-7029.06	0.43	0.23	0.48	0.50	0.22	0.88
PTB - VSL	-46.29	0.31	0.23	0.48	0.50	0.37	0.88

station delay variations ESDVAR (see Ref. [7]) which may introduce additional uncertainties. We neglect these uncertainties here.

The NPL-PTB link had been calibrated with Circular T on MJD 52434 with an estimated uncertainty of 5.0 ns. The expectation calibration value is 853.36 ns. The offset of 4.1 ns to the new calibration is within the estimated uncertainty of the previous calibration. The link NPL-VSL had been calibrated on MJD 51434 with Circular T with an estimated uncertainty of 5.0 ns. We expected a calibration result of 818.93 ns which is 15.9 ns off in comparison with the actual result. The link PTB-VSL had been calibrated with Circular T on MJD 51017 (July 1998) with an estimated uncertainty of 5.0 ns. The new calibration deviates significantly from the old constant (-34.59 ns). In summer 2002 (MJD 52416), an Agilent 5071 Opt001 clock was transported from PTB to VSL and back [10]. The result of the clock transportation was that the TWSTFT calibration constant of the PTB should be reduced by 7.8 ns. But this shift was never applied until now. Considering this observation, fictively the expected calibration value is –34.59 ns – 7.8 ns = -42.39 ns. The difference to the actual calibration is reduced to only 3.9 ns. This is still a little bit larger than the reported total estimated uncertainty of the clock transport of 1.5 ns.

6. CLOSING ERRORS

In the data evaluation of this report we assumed that the delays of the receiving paths of all participating stations are constant, i.e. not dependent on the station which transmits a dedicated signal. Generally, the receive path delay depends on the received spectrum. Because different earth stations transmit different spectra, the delay difference in the local earth station k determined relatively with respect to a TS is generally different from the delay difference during time transfer between station k and l. Those delay differences may cause so-called closing errors [11], which can be observed if two time scales are compared via different TWSTFT measurement paths, i.e. direct measuring UTC(k) – UTC(l), and using a third "relay" station m measuring [UTC(k) – UTC(m)] + [UTC(m) – UTC(l)] [12, 13]. A non-zero difference is called the closing error, whose origin is not well understood. In previous studies, no significant relations between delay variations and operating parameters were detected [13]. Furthermore, instabilities in the closing error, e.g. jumps of the same magnitude as the effect itself, were observed [12]. In a theoretical analysis [11] J. Davis showed that closing errors may deteriorate the accuracy of a calibration experiment as presented in this report. In the ITU-R documents [7] this technique is characterized by the key words: Portable earth station used in a relative mode (REL). To circumvent possible closing errors Davis proposed to use a portable earth station as an independent time-transfer system (IND). Both techniques are illustrated in Fig. 4. However, additional stations can be easily calibrated only by using the same TS in the REL mode. If the CCD of an additional station is determined all links to the previously visited stations would be calibrated assuming the TS has not changed its internal delays.

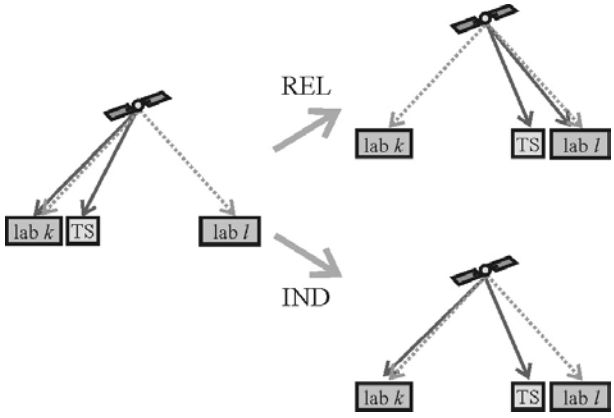


Figure 4: Different calibration techniques employing a portable travelling earth station (TS): TS used in a relative mode (REL) and as an independent time-transfer system (IND). The dotted arrows indicate the link to be calibrated, the red ones indicate the TWSTFT measurements to be evaluated for calibration.

The technique REL is usually applied by TUG campaigns, while IND is usually used by the USNO in their calibration campaigns e.g. for the link USNO-PTB [14]. We recorded data of all possible measurement combinations between the participating stations to enable a comparison between the REL- and IND-technique for the current campaign. In Fig. 5, results of both calibration techniques are compared. Each link between the four sites was investigated by analyzing the measurements of one day when the TS was located at

one site of the link. The error bars reflect the combined standard deviation of the two techniques. The standard deviation of $[UTC(k)-UTC(l)]_{REL}$ consists of the statistical uncertainty of both common clock experiments at station k and l which are necessary for calibration, as well as the TWSTFT measurement of the calibrated link. Only two measurements contribute to the statistical uncertainty of $[UTC(k)-UTC(l)]_{IND}$, the common clock measurement at k and the remote measurement between TS located at site l, and k.

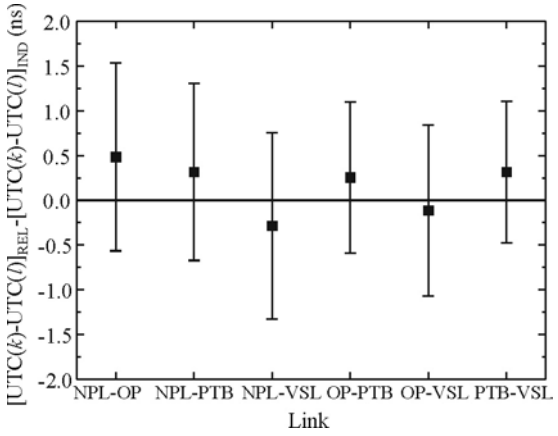


Figure 5: Analysis of possible common clock errors, by comparing the two different calibration techniques REL and IND.

No result shows a significant deviation from zero. Thus, we expect that there is no appreciable deterioration of the REL results. Nevertheless, a future detailed analysis may reveal systematic phenomena which probably lead forward to a better understanding of TWSTFT.

7. CONCLUSION

The differential delays of four European TWSTFT earth stations were determined by using a portable TWSTFT station. Calibration constants with a total estimated uncertainty down to 0.9 ns were achieved. A test for closing errors showed no significant deterioration of the accuracy of the calibration. The exercise demonstrates that a two-week schedule is sufficient for the calibration of four TWSTFT earth stations and only one person is needed to conduct the portable station during travelling, provided that the invited stations are well prepared and support the activity. This is important because in the future low cost high accuracy procedures are necessary to calibrate TWSTFT equipment of time laboratories contributing to the Galileo System Time.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support of A. Bauch T. Polewka, C. Richter (PTB), D. Valat (SYRTE) during assembling and operating the TS at the respective home laboratories and for helpful discussions. The support of INTELSAT for providing satellite transponder time for scientific purposes is gratefully acknowledged.

REFERENCES

[1] D. Kirchner: Two-Way Time Transfer Via Communication Satellites; *Proc. of the IEEE*, vol. 79 (7), pp. 983-989, 1991.

[2] D. Kirchner: Two-Way Satellite Time and Frequency Transfer (TWSTFT): Principle, Implementation, and Current Performance; *Rev. of Radio Science*, Oxford University Press, pp. 27-44, 1999.

[3] Circular T 205, BIPM, February 2005, see http://www.bipm.fr/en/scientific/tai/time_ftp.html.

[4] L. Breakiron et al.: The Accuracy of Two-Way Satellite Time Transfer Calibrations; *to be published in Proc. of the 36th PTTI*, Washington, DC, December 2004.

[5] D. Kirchner, H. Ressler, P. Hetzel, A. Söring: The calibration of three European TWSTFT stations using a portable station and comparison of TWSTFT and GPS common-view measurements; *Proc. of the 27th PTTI*, Reston, December 1998, pp. 365-376, 1999.

[6] F. Cordara et al.: Calibration of the IEN-PTB TWSTFT Link with a Portable Reference Station; *Proc. of the 18th EFTF 2004*, Guildford, April 5-7, paper 137, 2005.

[7] Recommendation ITU-R TF.1153-2. The operational use of two-way satellite time and frequency transfer employing PN time codes; Geneva, last update 2003.

[8] O. Koudelka, H. Ressler, B. Blanzano: Two-Way-Satellite-Time-Transfer Calibration Campaign, *Report: IAS.2003.AF.022-01*, Joanneum Research, Graz, 2004.

[9] Circular T 203, BIPM, December 2004, see http://www.bipm.fr/en/scientific/tai/time_ftp.html.

[10] D. Piester, P. Hetzel, A. Bauch: Recent Time and Frequency Activities at PTB, *Proc. of the 34th PTTI*, Reston, December 2002, pp. 365-376, 2003.

[11] J. A. Davis: Theoretical Treatment of Closing Errors Occuring in Two-Way Satellite Time Transfer; *Proc. of the 9th EFTF 1995*, Besancon, pp. 537-540.

[12] J. A. Davis et al.: European Two-Way Satellite Time Transfer Experiments Using the INTELSAT (VA-F13) Satellite at 307°E; *IEEE Trans. Instrum. Meas.*, vol. 44 (2), pp. 90-93, 1995.

[13] J. A. Davis et al.: Closing Errors Observed During Two-Way Time Transfer Experiments; *Proc. of 9th EFTF 1995*, Besancon, pp. 81-86.

[14] D. Piester et al.: Time Transfer between USNO and PTB: Operation and Calibration Results; *Proc. of the 35th PTTI*, San Diego, December 2003, pp. 93-101, 2004.

Dual frequency absolute calibration of GPS receiver for time transfer

A. de Latour, G. Cibiel, J. Dantepal, J.-F. Dutrey, M. Brunet, L. Ries, J.-L. Issler

CNES 18, avenue Edouard Belin, 31401 Toulouse Cedex, France - 33 (0)5 61 27 31 31
antoine.delatour@cnes.fr, gilles.cibiel@cnes.fr

ABSTRACT

This paper suggests a receiver time delay calibration method using a GPS simulator. The method has been used to get the receiver time delay of an Ashtech Z12T receiver. The method enables to compute an absolute time delay in the L1- and L2-frequency-bands within few nanoseconds accuracy. The measured absolute time delay depends mainly on the GPS simulator calibration. This method has several advantages : the radiofrequency environment is controlled and the calibration does not depend on the UTC-GPS time offset.

1. INTRODUCTION

Future Galileo receivers (Rx) and GPS Rx dedicated to the time transfer involve to be able to make accurate and repeatable calibrations of timing Rx.
The most widely used approach to calibration of Rx is co-locating the Rx, which will be calibrated with a reference Rx. This calibration approach, differential calibration, has been established as a method for measuring relative electrical delays between two Rx [PET 00]. To accomplish this, one Rx is designated as the reference Rx and is in constant circulation among timing laboratories. A relative calibration is performed between this Rx and the laboratory Rx. Therefore, this method is cumbersome and the stability of the calibration cannot be determined.
A better solution to this calibration problem would be a repeatable method to measure the electrical delays for each Rx individually. Work at the Naval Research Laboratory (NRL) has shown that the delays of an Ashtech Z-12T Rx can be measured relative to a GPS signal simulator [WHI 01], [LAR 03]. Our method is based on this one. The difference between the two methods comes from the different ways used to calibrate the GPS simulator:

- the NRL method used the zero-crossing method
- the CNES method used correlation functions

The CNES method improves the simulator calibration, the time delay is indeed estimated over more than one transition contrary to the zero-crossing method.
This paper explains the method used to calibrate GPS Rx for time transfer. In addition, the results issued from the Ashtech Z12T Rx calibration are shown. The calibration is performed in two main steps:

- *1st step*: the GPS signal simulator calibration in the L1-and L2-frequency bands
- *2nd step*: the GPS Rx calibration

The Ashtech Z-12T Rx has been chosen for this experiment because it has a repeatable internal reference. This internal reference is determined by the difference in the time of arrival at the Rx of the 20 MHz clock input signal (the 20 MHz is generated a 10 MHz frequency reference multiply by two) and the one pulse per second (1 PPS) input signal. Because the 20 MHz clock and 1 PPS input to the Rx are both coherent and externally generated, the calibration delay remains the same when the receiver power is cycled.

During the presented experiments, Rx and GPS signal simulator are located in a thermally controlled area, where the temperature is equal to 20.5 degrees Celsius with a maximum deviation of 1 degree.

2. GPS SIMULATOR CALIBRATION

2.1 General strategy

The internal delay of the simulator corresponds to the delay between the beginning of the PRN C/A or P ranging code and the internal one pulse per second synchronized with the GPS time. The simulator calibration is performed using only the PRN C/A ranging code for the following reasons:

- the chips in the C/A code and the P code have the same rise- or fall-time, because they share the same simulator bandwidth

- the generation of the reference P code is complex, because the P code period is one week contrary to the C/A code, whose period is equal to one millisecond.

The simulator delay is estimated using the correlation function, for each GPS frequency bands, between the following signals:

- the reference C/A code, synchronised with the simulator PPS and sampled at 10 GHz
- the C/A code generated by the GPS simulator at the radio frequency output and sampled at 10 GHz

The 10 MHz reference frequency is generated by an active Hydrogen MASER clock (EFOS 16 from Observatory of Neuchâtel).

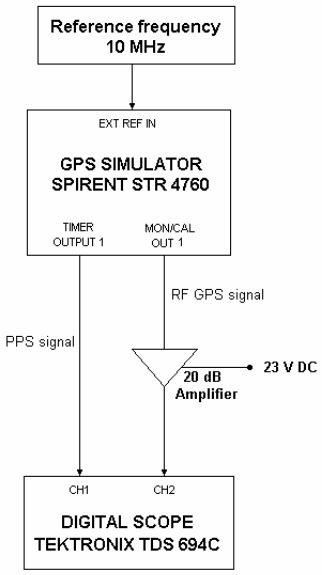


Figure 1 : Synopsis of the GPS simulator calibration

The GPS simulator is a Spirent STR 4760 (serial number 2041), which is calibrated by Spirent every year. To get the RF signal, we use the MON/CAL OUT 1 port, which allows the GPS signal to be monitored at a level approximately 60 dB higher than the one on the front panel. This output port enables to have a GPS signal level, which is well above the thermal noise floor. That's one reason why this port has been used for the simulator calibration. The second reason is that this port is used by Spirent for the annual calibration. The PPS signal has a 1s period with a high time of 100 ms. The PPS signal is synchronised with 1s epochs in simulated GPS time.

The PPS signal is transmitted with a 50-Ohms RF cable, which induced a delay of 32.131 ns with a standard deviation of 0.007 ns.

The RF GPS signal goes through different elements:

- a 50 Ohms RF cable with an isolator :
  - mean group delay in the L1 frequency band: 31.578 ns
  - standard deviation group delay in the L1 frequency band: 0.105 ns
  - mean group delay in the L2 frequency band: 31.633 ns
  - standard deviation group delay in the L2 frequency band: 0.105 ns
- an amplifier: Mini-Circuits ZFL – 2500 VH with the following properties :
  - frequency range: 10 MHz to 2.5 GHz
  - gain: 20 +/- 1.5 dB
  - maximum input power: 10 dBm
  - mean group delay : 0.55 ns
  - standard deviation group delay : 0.01 ns

The RF GPS signal has a delay respectively of 32.128 ns and 32.183 ns with 0.105 ns of standard deviation, at the input of the digital scope, in the L1- and L2-frequency-bands.

2.2 Dual frequency calibration method

The calibration method enables to evaluate the GPS internal delay in the L1- and L2-frequency bands. The calibration is performed in four steps:

- *1st step*: running the internal calibration of the GPS simulator
- *2nd step*: creating a scenario for the GPS simulator
- *3rd step*: storing the PPS and the RF signals with the digital scope
- *4th step*: processing the stored data to get the estimated delay

Internal calibration
The internal calibration is performed to have a better stability concerning the estimated simulator delay. After running the internal calibration, the GPS simulator has its optimum system performance.
The calibration utility aligns the code and carrier phases of each channel within the simulator unit.

GPS simulator scenario
To be able to calibrate the GPS simulator, the Single Channel Utility has been used. This utility allows one channel of the satellite simulator to be used as a

simple GPS signal source. The GPS signal characteristics may be interactively controlled. These main simulator characteristics are the following: the PRN code, the navigation data message, the signal level, the Doppler.

Storing the PPS and GPS RF signals

The PPS signal and the RF signal have been stored using the digital scope Tektronix TDS 694 C. Each signal is stored in a file with a Mathcad format. In each file the sampling frequency is equal to 10 GHz and there are 120000 samples.

The Figures 2 and 3 give a first approximation of the simulator delay, which seems to be equal to 100 ns (C/A code in the L1 frequency band). Looking at the Figure 3, we can say that the zero-crossing method cannot be used without pre-processing the GPS signal because of the noise level.

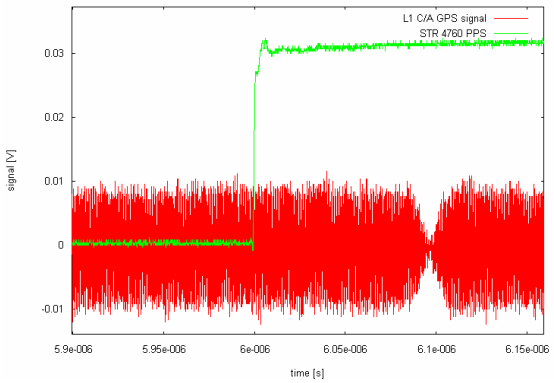


Figure 2 : GPS signal (C/A code in the L1 frequency band) with the PPS signal generated by the simulator

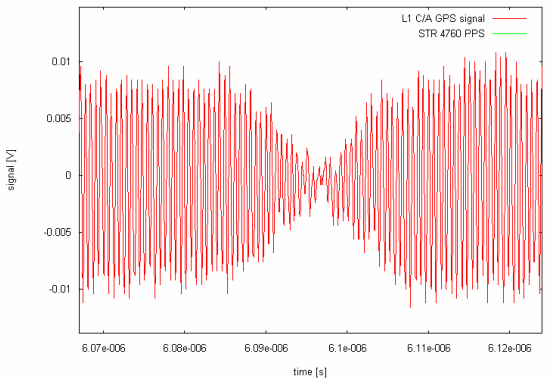


Figure 3 : zoom on the GPS signal (C/A code in the L1 frequency band)

Processing the stored data

The method used to estimate the simulator delay looks like a standard GPS signal acquisition. Two parameters have to be found:

- the code offset versus the PPS generated by the simulator
- the residual carrier frequency offset

These two parameters are evaluated simultaneously. The number of residual carrier frequency bins that must be observed is equal to 21 with a step of 4 kHz. The number of code offset bins that are observed is equal to 120000 with a step of 100 ps.

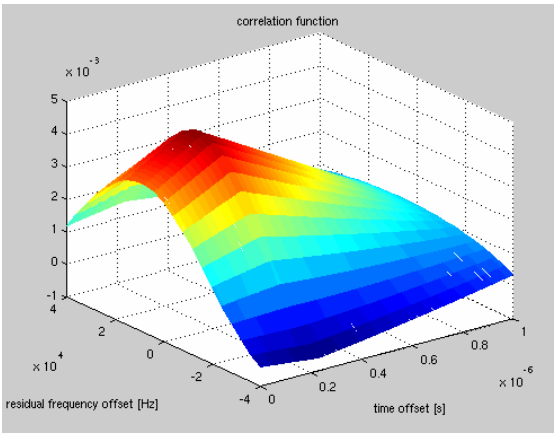


Figure 4: zoom on the correlation function near the maximum value (L1)

For each frequency offset, the following operations are performed:

- the RF GPS signal is extended from 120000 to 131072 samples using the zero-padding method ($131072 = 2.10^{17}$). The zero-padding improves the correlation function.
- the RF GPS signal is multiplied by an oscillator, whose frequency is equal to the L1-central frequency (1575.42 MHz) plus the residual carrier frequency offset
- the base band GPS signal is filtered with a Nyquist filter, whose parameters are the following:
  - roll-off factor: 0.5
  - Nyquist frequency: 40 MHz
- the base band GPS signal is correlated with the reference signal, which is synchronised with the PPS signal generated by the simulator
- if the absolute maximum value of the correlation function is not bigger than 90% of the roughly estimated maximum amplitude value of the input signal, then the whole process has to be made one more time adding 30 degrees to the local oscillator phase.

The correlation function enables to know the following values:

- the estimated time offset
  - 231.5 ns (C/A code in the L1 frequency band)
  - 234.5 ns (C/A code in the L2 frequency band)
- the estimated residual frequency offset:
  - 8 kHz (C/A code in the L1 frequency band)
  - 16 kHz (C/A code in the L2 frequency band)

These results have been reached adding a phase offset to the local oscillator phase.

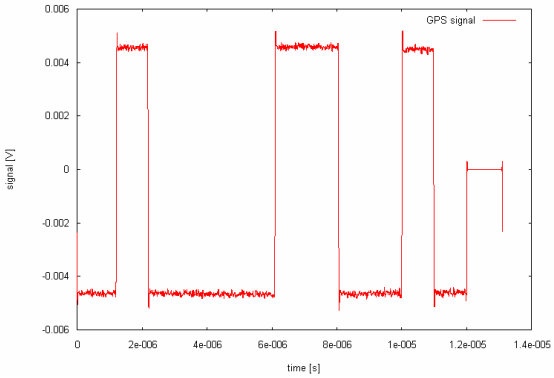


Figure 5 : base-band signal using the optimal residual frequency offset (1575.44 MHz) in the L1 frequency band

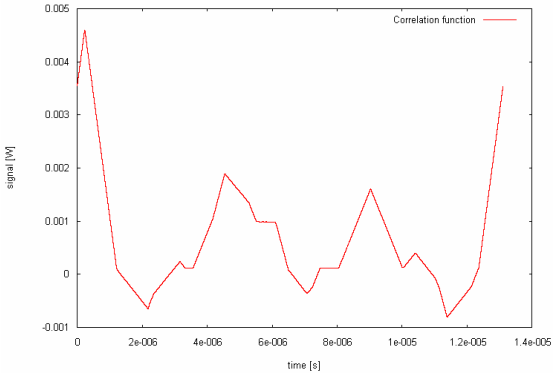


Figure 6 : correlation function using the optimal residual frequency offset (1575.44 MHz) in the L1 frequency band

The stored signal length is roughly equal to 12 code chips and the pulse of the PPS signal is located in the middle. To improve the correlation function, the reference signal is shifted by 6 code-chips. The GPS second is the time, where the signal exceeds the 1 Volt level. This level has been used in the experiment (4). Considering these two remarks, we can respectively estimate the GPS simulator delay in the L1- and L2-frequency bands: 105 ns and 108.47 ns.

2.3 Results

Before storing the signals, an internal calibration is done each time, to get a more reliable delay estimation of the simulator calibration.

Date	L1 correlation delay [ns]	L1 delay [ns]	L2 correlation delay [ns]	L2 delay [ns]
17/02/2005	239,9	105,1	243,3	108,45
18/02/2005	239,9	105,1	243,5	108,65
08/03/2005	239,8	105	243,4	108,55
09/03/2005	239,6	104,8	243,1	108,25
Average	-	105	-	108,47

Table 1: GPS simulator calibration results

	L1 standard deviation [ns]	L2 standard deviation [ns]
PPS cable	0,1 ns	0,1 ns
RF cable	0,10 ns	0,10 ns
GPS simulator*	0,12 ns	0,15 ns
Interchannel biais**	0,33 ns	0,33 ns
Tick-to-phase	0,1 ns	0,1 ns
Correlation accuracy	0,1 ns	0,1 ns
Results	0,41	0,42

Table 2: Measurement errors for the simulator calibration

* the measured value is below the SPIRENT measured value from 1.668 ns
** this value is given by SPIRENT.

3. GPS RECEIVER FOR TIME TRANSFER CALIBRATION

3.1 Calibration method

Once the simulator internal delay has been computed, the calibration of the GNSS receiver can start. The calibration of the GPS receiver is based on the pseudoranges-measurements. The GPS internal delay is calculated thanks to the following equation:

$$\frac{R_X R}{c} - \frac{SR}{c} - SD = R_X D \qquad (1)$$

with,

- $\mathbf{R_XR}$: Pseudorange computed by the GPS receiver
- **SR**: Pseudorange simulated by the GPS simulator
- **SD**: GPS simulator internal delay
- $\mathbf{R_XD}$: GPS receiver internal delay
- **c**: light celerity (299,792,458 ms^{-1})

The GPS simulator internal delay has been calculated during the simulator calibration. The pseudoranges simulated by the GPS simulator are stored in a file with an ASCII format. The pseudoranges computed by the GPS receiver are stored in a file with a Rinex format. The comparison between the two sets of measurement enables the estimation of the GPS receiver internal delay.
As mentioned already, the Ashtech Z12T Rx, involved in the experiment, are not the standard Z12 Rx. Indeed, these specific receivers were designed to

perform repeatable measurements. These measurements depend on the so called “Tick to Phase” (TtP) value (time difference between the 1 PPS rising tick and the next rising zero of the 20 MHz). As a consequence the estimated GPS receiver internal delay is related to this TtP value. The whole experiment has been performed with a unique TtP value.

3.2 Calibration configuration

The calibration of the GPS receiver is achieved with the following configuration.
To calibrate the GPS receiver, the output level must be the standard GPS level. For this reason, attenuators are added in the RF link to compensate the 60dB-amplification of the simulator (*see 2.1 paragraph*). The attenuation value used was of approximately 60 dB. The generated PPS signals have a 1s period with a high time of 100 ms. The PPS signals are synchronised with 1s epochs in simulated GPS time.

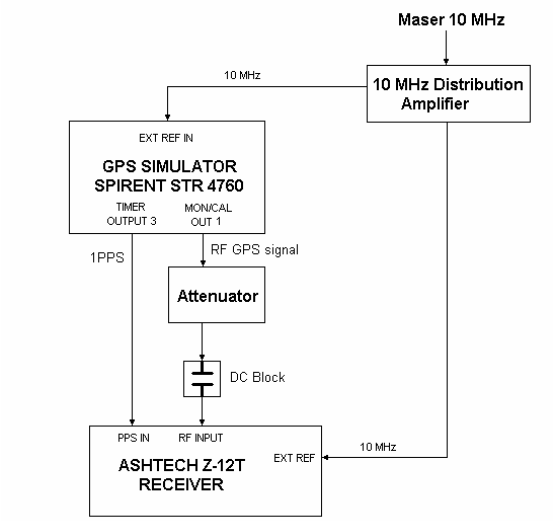


Figure 7 : Synopsis of the GPS receiver Z-12T calibration

The RF link between the simulator and the receiver has got the following elements:

- a DC block
- two attenuators : 60 dB
- a 50 Ohms RF cable

The effect of the RF link is as follow :

- L1 frequency band
  - mean group delay: 5.99 ns
  - standard deviation group delay: 0.55 ns
- L2 frequency band
  - mean group delay: 5.42 ns
  - standard deviation group delay: 0.5 ns

3.3 GPS Simulator scenario

The GPS simulator can generate simultaneously a radiofrequency signal issued from several channels. Each channel is calibrated separately by the simulator, using the tool called Channel Alignment. The simulated scenario is a specific scenario with:

- a receiver is located at a fixed-position (0° N, 0°E)
- no ionosphere and no troposphere
- two transmitted pseudorandom codes in the L1 frequency band: the C/A code and the P code
- one transmitted pseudorandom code in the L2 frequency band: the P code
- the GPS constellation is the default GPS constellation from Spirent, with 24 healthy space vehicles

3.4 Calibration results

The calibration has been performed on two Ashtech Z12 T Rx, one from the BIPM (LP02944) and the other from the CNES (5555-1122A0).

Pseudorange difference
The Table 3 is focused on the pseudorange difference, which corresponds to the difference between the simulated and the received pseudorange (see equation 1).

	Mean at L1 [m]	Std Dev at L1 [m]	Mean at L2 [m]	Std Dev at L2 [m]
CNES	96,763	0,165	96,557	0,682
BIPM	92,353	0,205	97,443	0,671

Table 3 : Pseudorange difference

The Table 4 corresponds to the delay of this pseudorange difference. Note the impressive difference between the deviation at L2 frequency band and the deviation at L1 frequency band. This point must be cleared for the next experiments.

	Mean at L1 [ns]	Std Dev at L1 [ns]	Mean at L2 [ns]	Std Dev at L2 [ns]
CNES	322,767	0,55	322,079	2,275
BIPM	308,056	0,684	325,035	2,238

Table 4 : Delay difference

Tick-to-phase calculation

The important aspect of the Ashtech Z-12T Rx is the repeatability of the internal reference. The 20 MHz zero crossing that immediately follows the rising tick of the 1 PPS signal marks the internal reference (zero) time for the Rx. Whether this internal reference can be determined by opening the receiver case and inspecting the jumper on the circuit board : its position is clearly marked as rising or falling. The 1 PPS input does not have a zero rise time. Therefore, the rising 1 PPS signal is measured at the point it achieved 1 Volt. A TtP of 10.1 ns has been measured. The accuracy of determining the TtP is assumed to be 0.1 ns.

Rx calibration results

The receiver delays are computed using the equation (1).

	(RxR-SR)/c at L1 [ns]	- SD at L1 [ns]	- RF cable delay at L1 [ns]	= RxD at L1 [ns]
CNES	322,767	-105,005	-5,99	211,772
BIPM	308,056	-105,005	-5,99	197,061

Table 5 : Receiver delay at L1

	(RxR-SR)/c at L2 [ns]	- SD at L2 [ns]	- RF cable delay at L2 [ns]	= RxD at L2 [ns]
CNES	322,079	-108,075	-5,42	208,584
BIPM	325,035	-108,075	-5,42	211,54

Table 6 : Receiver delay at L2

	L1 standard deviation BIPM-receiver [ns]	L2 standard deviation BIPM-receiver [ns]	L1 standard deviation CNES-receiver [ns]	L2 standard deviation CNES-receiver [ns]
RF cable	0,55	0,5	0,55	0,5
Tick-to-phase	0,1	0,1	0,1	0,1
SD	0,41	0,42	0,41	0,42
(RxR-SR)/c	0,68	2,24	0,55	2,27
result	0,97	2,33	0,88	2,37

Table 7: Measurement errors for the receiver calibration

These results can not be directly compared with the NRL's result [WHI 00] [LAR 03], because the measures have not been performed at the same TtP.

4. CONCLUSIONS

This experiment has succeeded in performing an absolute GPS receiver calibration. The GPS simulator has been calibrated thanks to a new method proposed by CNES. The new method called "acquisition like process" has improved the simulator calibration accuracy as expected.

The measured receiver delays must be computed with several TtP value in order to compare our calibration with the NRL's calibration. However, the isolator and DC-block must be replaced in the next experiments. Moreover it would be interesting to measure the interchannel biais of the GPS simulator in order to reduce the error budget.

ACKNOWLEDGEMENT

We would like to thank Joe White and Ron Beard, from NRL, and Gérard Petit, from BIPM, for helpful discussions.

REFERENCES

[ASC 98] Ascarrunz F.G., Parker T.E., Jeffers S.R., "Pseudo-random code correlator timing errors due to multiple reflections in transmission lines", *in Proceedings 30th Annual Precise Time & Time Interval (PTTI) Meeting*, pp. 433-438, 1998

[LAR 03] Larson K. M., "Development of a Carrier Phase Time and Frequency Transfer System", *PhD dissertation,* University of Colorado, 2003

[PET 00] Petit G., Jiang Z., Ulrich P., and Taris F., "Differential calibration of Ashtech Z-12T receiver for accurate time comparisons", *in Proceedings of 14th EFTF*, pp. 40-44, 2000

[WHI 01] White J., Beard R., Petit G., and Powers E., "Dual frequency absolute calibration of a geodetic GPS receiver for time transfer", *in Proceedings of 15th EFTF*, pp. 167-172, 2001

ON TWOFOLD USE OF DUAL-MIXER TIME-DIFFERENCE MULTIPLICATION

R. Barillet[*], J. Čermák[**], L. Šojdr[**], A. Kuna[**]

* SYRTE, Observatoire de Paris, 61 avenue de l'Observatoire, 75014 – Paris, France
Roland.Barillet@obspm.fr, fax: + 33 (0)1 4325 5542
* *Institute of Radio Engineering and Electronics, Czech Academy of Sciences, Chaberská 57, 182 51 Praha 8, Czech Republic, cermak@ure.cas.cz, sojdr@ure.cas.cz, kuna@ure.cas.cz, fax: +420 284 680 222

ABSTRACT

The paper describes progress in development of two related applications of the dual-mixer time-difference multiplication (DMTDM): for measurement of frequency or time stability, that is currently the main application, and for measurement of short time delays with a very high resolution (0.1 ps) that has not yet been investigated in the low radio frequencies (5 – 120 MHz) because of its inaccuracy.

In the application for stability measurement, the apparatus developed at IREE already reaches an outstanding performance (ADEV = 8×10^{-16} at τ= 10 s, 5 MHz). An ultimate improvement of long-term stability requires the reduction of the sensitivity to the environment.

In the application for time-delay measurement, a different apparatus was developed on the same general basis, but with reduced mismatches and non linearity in order to improve the uncertainty of the measurement. The dependence of the measurement error on frequency can be modeled and allows us to fit the theoretical and experimental data, thus providing an estimate of the time delay of interest. We have developed three different models. The first and the second models give a satisfactory agreement (4 ps) in the estimated time delay, but do not fully take advantage of the available resolution. The third model implementation currently is in progress. It should reduce the accuracy under 1 ps and ideally to about 0.1 ps with an appropriate processing of the data.

1. INTRODUCTION

The dual-mixer time-difference multiplication has originally been devised to improve the sensitivity of short-term frequency-stability measurement in the time domain [1]. So far the DMTDM has almost exclusively been used for this purpose. The excellent noise performance of DMTDM can be demonstrated on the system developed at IREE in collaboration with SYRTE. At 5 MHz, this system provides the background time stability of 3.5 fs in terms of time deviation (TDEV) at the basic sampling interval $\tau_0 = 0.2$ s, with the flicker PM floor ≈ 2 fs. The corresponding background frequency stability is $7\text{x}10^{-15}$ at 1 s and $8\text{x}10^{-16}$ at 10 s in terms of Allan deviation (ADEV) [2].

A question arises as to whether and how this excellent noise performance can be employed in other applications besides the stability measurement. An obvious use is for measurement of the time delay in different elements (e.g. cables, connectors etc.) that can be inserted into one of the input signal paths. In [3] it has been hinted that the measurement uncertainty is degraded by several effects such as signal reflections, nonlinearities, voltage offsets, and temperature drifts, which otherwise do not manifest themselves in the measurement of short-term stability.

In this paper we will discuss these two uses of DMTDM, i.e. what they have in common and what is different about them. We will demonstrate the case on the IREE DMTDM systems. Concerning the stability application we will mention the problem of the determination of the noise limit, and the ways of further improvement of the IREE system.

The main attention, however, will be paid to the time-delay measurement. Since a large portion of the measurement error is dependent on frequency, a method has been proposed in [3] which makes use of measurement at a whole range of frequencies while keeping the multiplication coefficient constant. The frequency dependence of the error can be modeled and by fitting the model to the measured data, the measurement uncertainty can be substantially improved. We will describe the progress in application of this method which now can be verified on the DMTDM unit designed specially for time-delay measurement.

2. MEASUREMENT BASICS

Assume a model where the time-dependent time difference x(t) is defined by two sine-wave signals at nominal frequency ν

$$v_1(t) = A_1 \sin 2\pi\nu t, \qquad (1)$$

$$v_2(t) = A_2 \sin 2\pi\nu[t + x(t)] \qquad (2)$$

where A_1 and A_2 are the amplitudes whose variations can be neglected. The reference time, t, is defined by the signal v_1 in this case but generally there is no preference for either of the signals.

The time difference x(t) carries the information about the angular (time, phase or frequency) stability between the two signals and also about the amount of their desynchronization (for ideally synchronized signals x(t) = 0). In the real measurement, the difference x(t) is represented by the zero-crossings of $v_1(t)$ and $v_2(t)$, respectively, which are detected by the DMTDM system.

A time-interval measurement setup with the DMTDM is shown in a simplified form in Fig. 1. There, ν_B is the beat-note frequency and $m = \nu/\nu_B$ is the coefficient by which the time difference x is multiplied. The counter thus provides samples of the multiplied time difference, $x_k^* \approx mx_k$, at instants $t_k = t_0 + k\tau_0$, where k is an integer and $\tau_0 = 1/\nu_B$ is the basic sampling interval.

It should be noted that if the counter were sensitive enough to measure x(t) directly, there would be no need for the use of DMTDM, but the noise related to the very large input bandwidth of the counters does not allow for high enough sensitivity.

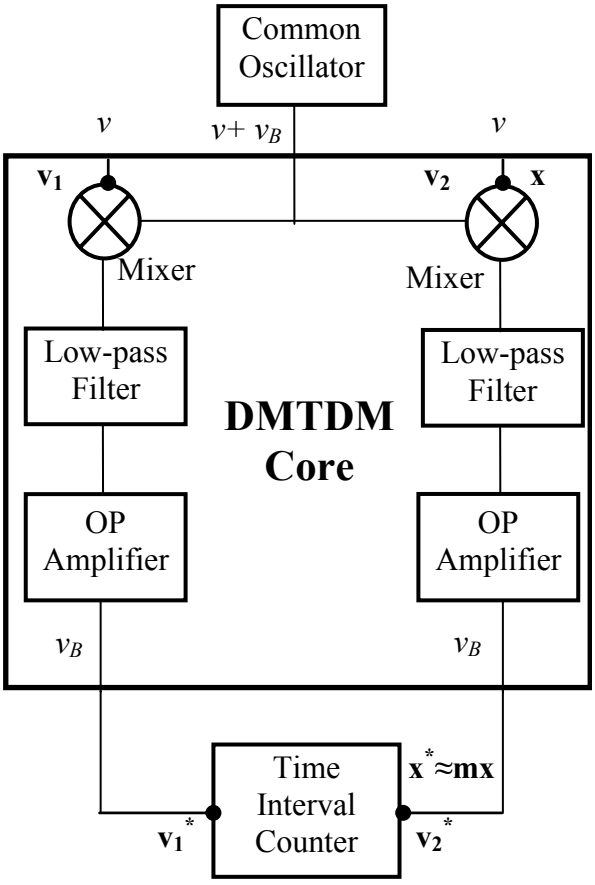


Fig. 1. Basic measurement setup with DMTDM.

At time t_k the counter provides the time difference which can be described as

$$x^*(t_k) = m[x(t_k) + \varepsilon(t_k) + E(\nu, p_{R,j})] + \varepsilon^*(t_k) \quad (3)$$

where x is the input time difference, ε is the error introduced by the measurement system before multiplication, and ε^* is the error introduced after multiplication (primarily by the counter). Error ε represents the input bias and its instability due to voltage offsets, flicker PM in the mixers, differential phase noise of the common oscillator, temperature variations, etc. Error E is mainly due to signal reflections which are specific to a measurement configuration R. This error is dependent on frequency and can be expressed in terms of j=1, … N_R parameters ($p_{R,j}$) characteristic of the configuration R = 1,2… .

The time, t, in (3) is defined by signal v_1*. In order to simplify the notation, we will denote the quantity at time t_k by index k, e.g. $x(t_k) = x_k$ etc. For simplicity of the model, we do not discern between x_k and ε_k measured at ν, and the filtered x_k and ε_k measured at ν_B.

General equation (3) applies to both stability and time-delay measurements. Fortunately, in neither of the two applications we need to measure x but the changes in x.

3. MEASUREMENT OF FREQUENCY STABILITY

In this application, the measurement setup is that shown in Fig. 1 where the signals v_1 and v_2 originate from two uncorrelated oscillators whose mutual stability is to be measured. Thus with the common oscillator, which is a high-quality quartz oscillator, there are three frequency sources involved in the measurement.

When measuring highly stable oscillators, which is typically the case, the signals v_1 and v_2 can be kept quasi-synchronous, i.e. within $x << 1/\nu$, for a considerable period of time which is the condition for rejecting the noise contribution from the common oscillator.

The frequency stability is commonly characterized by ADEV(τ), $\tau = n\tau_0$, defined from the second difference of x_k. For the first difference of x^*_k, we have

$$x^*_{k+1} - x^*_k = m\,[x_{k+1} - x_k + \varepsilon_{k+1} - \varepsilon_k + E_{k+1}(\nu, p_{R,j}) - E_k(\nu, p_{R,j})] + \varepsilon^*_{k+1} - \varepsilon^*_k \quad (4)$$

Obviously, all biases are canceled by the first difference which thus characterizes solely the variations. Since the measurement is performed at a fixed frequency and the configuration remains the same during the measurement, we have $E_{k+1} - E_k = 0$, excluding very small thermal drifts that can only be observed in long term.

Thanks to large multiplication coefficient (in the IREE system, $m = 10^6$), the difference $\varepsilon^*_{k+1} - \varepsilon^*_k$ can be neglected with respect to the multiplied differences in the bracket and (4) can be rearranged in the form

$$x_{k+1} - x_k \approx \frac{x^*_{k+1} - x^*_k}{m} - (\varepsilon_{k+1} - \varepsilon_k) \qquad (5)$$

Without the need of writing the second difference, it is obvious that the measured stability of x(t) is the stability of $x^*(t)/m$ degraded by the fluctuations of the error ε_k.

Supposing that the IREE version-2 represents the state-of-the-art noise performance of the classical DMTDM structure at 5 MHz, a question arises about to what extent the flicker PM floor of 2 fs (see Fig. 2 reproduced from [2]) reaches the DMTDM noise limit.

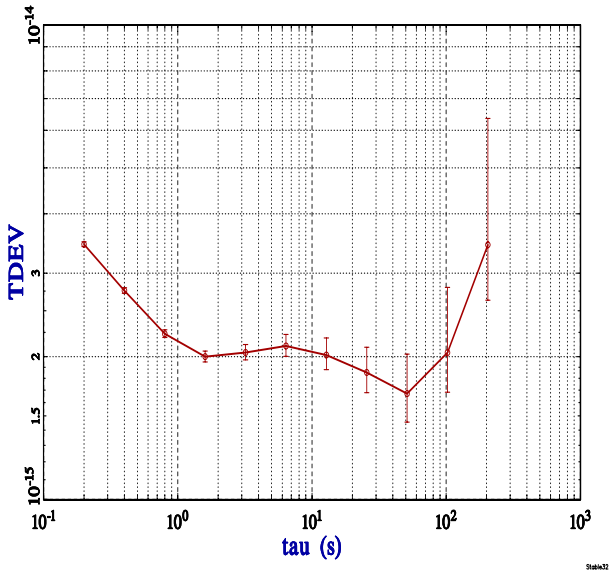


Fig. 2. Time stability of current IREE DMTDM system (TDEV in seconds).

This noise limit is given by a sum of flicker PM contributions from the input components, i.e. primarily from the mixers and slightly from the low noise preamplifiers, and possibly from some environmental effects. Since we have made a careful choice of low-noise input components and optimized the design, there is likely no margin for the improvement by making further modifications in the input circuitry. In version-2, however, we have observed the presence of non-stationary noise that may disturb the long-term performance (positive slope of the plot in Fig. 2), and even the flicker floor itself. These spurious effects are due to inappropriate construction, mainly to input cabling and connectors, that increases the sensitivity for the pickups due to environment including the motion of people in the vicinity of the experiment. Thus a reconstruction of version-2 is a necessary condition for any further decrease in the flicker PM floor (semi-rigid cables and SMA connectors should be used, and better grounding and more effective shielding should be ensured).

The drawback of the current IREE system is also its relatively low bound of the flicker PM floor (≈ 100 s, see Fig. 2). The effects (most likely due to temperature) that manifest themselves on the long run and which impair the system long-term performance have not yet been appropriately investigated.

Improvement can also be made by increasing the maximum desynchronization for which the ultra-low-noise performance is ensured, i.e. where the impact of the common oscillator can be neglected (currently we have $x < 1$ ns). Namely, the common-signal frequency $\nu + \nu_B = 5$ MHz + 5 Hz (i.e. $m = 10^6$) is obtained by dividing 10 MHz + 10 Hz by two by means of a relatively noisy digital divider. In the new version the digital divider will be replaced with a less noisy regenerative divider.

4. MEASUREMENT OF TIME DELAY

The basic measurement setup is shown in Fig. 3.

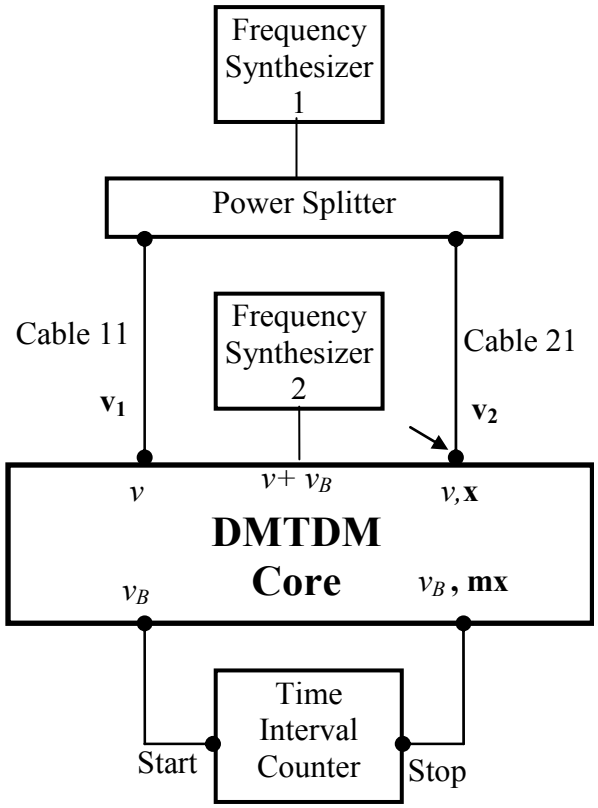


Fig. 3. Experimental setup for time-delay measurement.

The first measurement is performed without the delay dx in the configuration R = 1 as shown in Fig. 3. The second measurement is carried out with dx inserted

into the path of signal v_2 at the end of cable 21 (marked by the arrow in Fig. 3). The detail of the insertion of dx (represented by a short cable) is shown in Fig. 4.

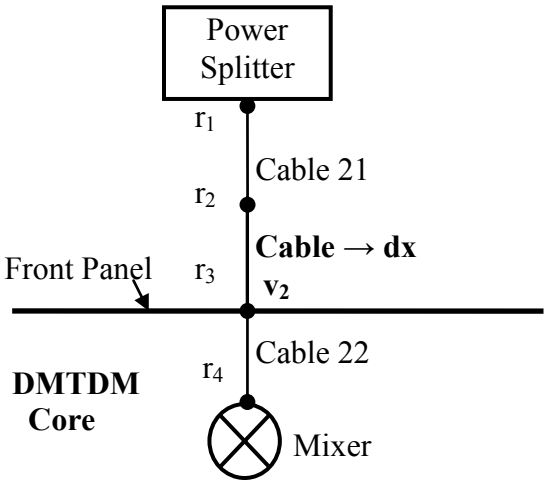


Fig. 4. Configuration with inserted time delay.

Since in measurement 2 the configuration is different, we will have R = 2. Note that we have three reflection points in measurement 1, and four in measurement 2.

Assume simply that we insert the delay dx at an instant θ, $t_k < \theta < t_{k+1}$ Referring to (3), we can write for measurement 1 taken up to t_k

$$x^*_k = m[x_k + \varepsilon_k + E(\nu, p_{1,j})] + \varepsilon^*_k \qquad (6)$$

and for measurement 2 taken from t_{k+1}

$$x^*_{k+1} = m[x_{k+1} + dx + \varepsilon_{k+1} + E(\nu, p_{2,j})] + \varepsilon^*_{k+1} \qquad (7)$$

In changing the configuration, we also affect the error ε (e.g. due to larger desynchonization) but now assume that the change in ε is negligible in comparison with the change in E.

Denoting the averages of quantities up to t_k by index 1 and from t_{k+1} by index 2, and neglecting again $\varepsilon^*_2 - \varepsilon^*_1$, we can write the difference

$$x^*_2 - x^*_1 = m\,[x_2 - x_1 + dx + \varepsilon_2 - \varepsilon_1 + E(\nu, p_{2,j}) - E(\nu, p_{1,j})] \qquad (8)$$

In (8), the biases of x and ε are canceled but the difference $E(\nu, p_{2,j}) - E(\nu, p_{1,j})$ may remain considerably large. This is the essential problem of the time delay measurement.

Fortunately the dependence of E on frequency allows us to devise a model that can be fit to the data measured at a frequency variable within a certain frequency range. In order to avoid the impact of other frequency-dependent effects, the range must not be too large. To perform this measurement, we use two frequency synthesizers instead of two oscillators (see Fig. 3). The Synthesizer 1 generates the probe signals v_1 and v_2 at variable frequency ν while the synthesizer 2 provides the common signal at variable offset frequency $\nu + \nu_B$ in such a way that the multiplication coefficient $m = \nu/\nu_B = 10^6$ is kept constant.

4.1 Measurement reproducibility

The measurement reproducibility provides the essential information on the potential in measurement uncertainty. Fig. 5 shows the standard deviation of 25 measurements of the time delay of 5.3 ns which we measured repeatedly in the range of 3 to 149 MHz in 1 MHz steps over about 6 hours. Sigma decreases as 3 ps/f (MHz), with values slightly larger than this dependence at frequencies > 40 MHz. The peaks are at n × 5 MHz and are probably related to interferences between several oscillators at 5 and 10 MHz. This graph shows that with an appropriate weighting of the data, one can expect an accuracy of about 0.1ps with such a set of data, if the other contributions are properly modeled.

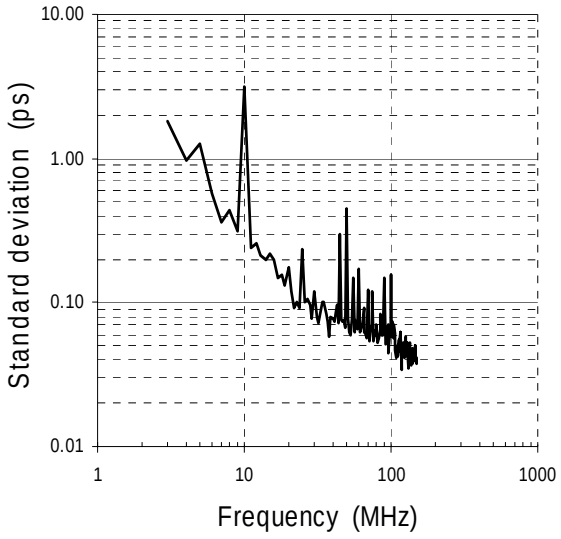


Fig. 5. Reproducibility of time-delay measurement.

4.2 Error models and first experimental results

We start with a model based on equation (8) modified in the form

$$\Delta x^* = m\,[\Delta x + \Delta\varepsilon + \Delta E(\nu, p_j)] \qquad (9)$$

where we have introduced the differential error ΔE characterized by a set of N parameters $(p_j) = p_1, \ldots, p_N$. The time delay dx is included into ΔE and as such it becomes one of the model parameters, say $p_1 = dx$.

In the first approximation we can neglect the errors Δx and $\Delta\varepsilon$ due to time variations in x and ε, respectively, so that we obtain

$$\Delta x^*(\nu) \approx m\, \Delta E(\nu, p_j) \qquad (10)$$

By fitting the model error $\Delta E(\nu, dx, p_2, .. p_N)$ to measured $\Delta x^*(\nu)/m$ we can find dx and the rest of the parameters. The adequacy of the model can be assessed through standard deviation (σ) of the residuals.

In what follows we will mention three different models of ΔE:

1) The simplest model is based on one reflected roundtrip between mixer and power splitter (PS). In this model, we consider 13 parameters to be adjusted (see details in Fig. 4): 3 delays in cables (dx, x_{21}, and x_{22}) and the reflection coefficients of the mixer and of the PS (r_4 and r_1) with 4 parameters each – modulus and phase are linearly variable with frequency. The modulus of the reflection of the third harmonic generated at the input of the mixer is taken into account with 2 parameters. The fit is shown in Fig. 6. We have $\sigma = 19.7$ ps, due to large fluctuations ≈ 65 ps p-p, probably related to reflections in the connecting points which are not taken into account in this model. The third harmonics does not give any significant improvement. The p-p fluctuation of the residuals versus frequency is only 6 times smaller than that of the basic measurement, which is not a sufficient improvement. The delay of interest provided by the fit is dx = 3226.8 ps.

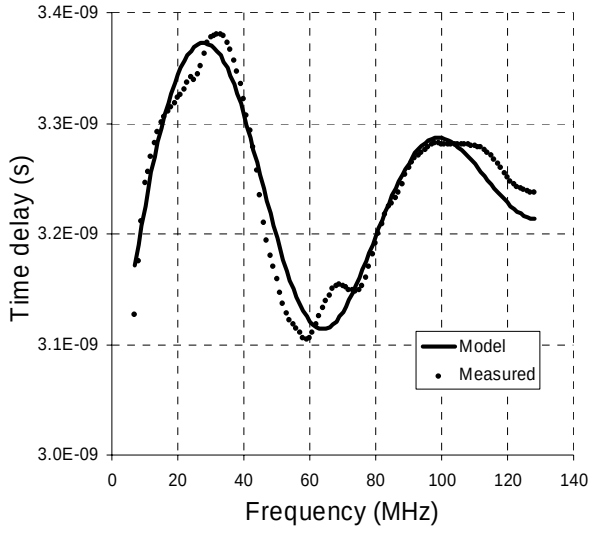


Fig. 6. Fit obtained with the simplest model.

2) Parametric model with reflections in all connecting points r_1 through r_4 (see Fig.4). Its implementation requires 25 parameters including the effect of the 3rd harmonic. The fit is shown in Fig. 7.

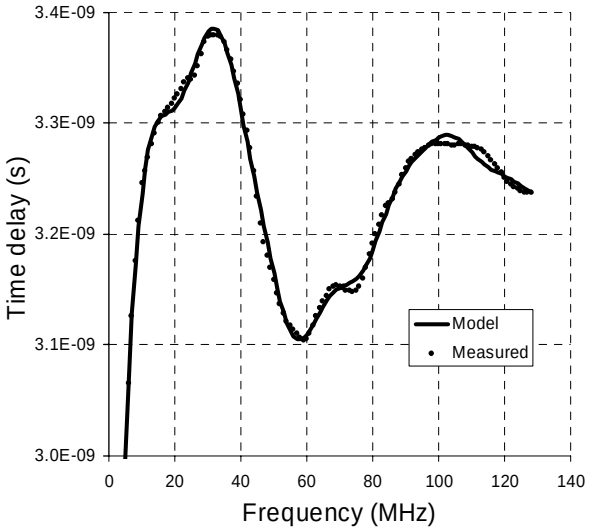


Fig. 7. Fit obtained with the parametric model.

This model gives a better result, with residuals $\sigma = 5.9$ ps but with 25 parameters! The introduction of the 3rd harmonic is efficient since it reduces the residuals from 19.7 ps down to 5.9 ps, though there is no phase adjustment of the 3rd harmonic (assumed in phase with the reflected fundamental harmonic). The residuals are shown in Fig. 8. The delay provided by the fit is dx = 3231.1 ps. The cable delays x_{21} and x_{22} are in rough agreement with the expected values.

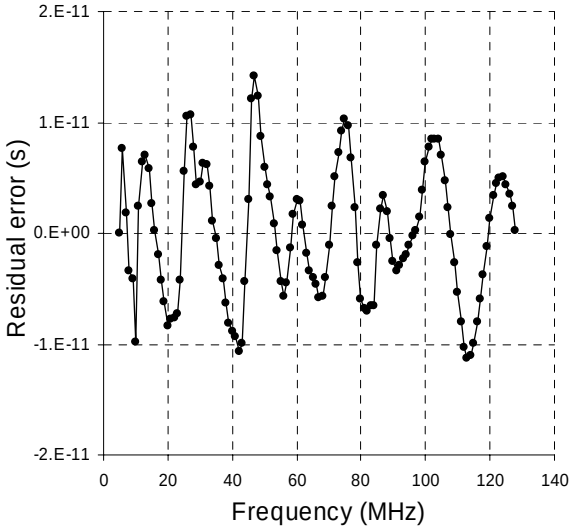


Fig. 8. Residual error with the parametric model.

This approach and its results concerning the reflection coefficients are not consistent with the theory of transmission lines, though the estimate of dx corresponds to that of the previous model within 4 ps. Both estimates are in agreement with the measurement with a fast scope (the latter carried out with about 100 ps uncertainty).

3) The model with an improved consistency is based on two unknown reflection coefficients, i.e. r_4 at the mixer and r_1 at PS (4 parameters each). The other reflection coefficients, r_2 and r_3 (different when seen from the left and from the right) are calculated. We have a total of 15 parameters that are completely independent: eight for r_1 and r_4, three for the cable delays (dx, x_{21}, x_{21}), two for the ratios of characteristic impedances of the three lines, and finally two for the third-harmonic modulus.

Currently this model has been implemented only for the basic configuration (Fig. 3) without dx. The fit is shown in Fig. 9.

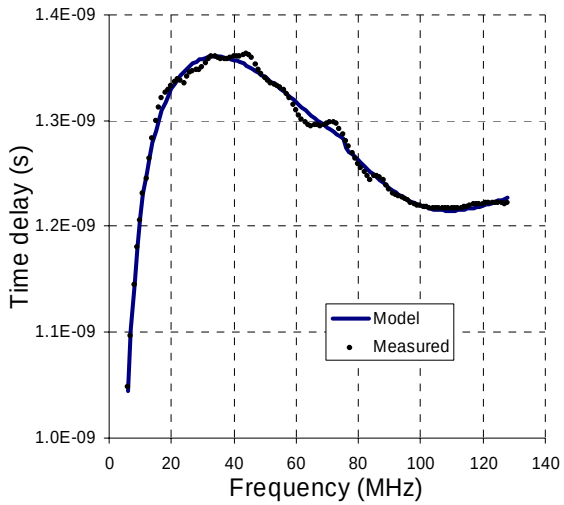


Fig. 9. Fit obtained with the improved-consistency model.

The residuals give $\sigma = 5.10$ ps (the result at 5 MHz removed). The values of x_{11} and x_{22} are not in a good agreement with the values obtained with the second model. Because of its inconsistency, this is not surprising. The values found for reflections at the mixer and PS, $r_4 = 0.16$ and $r_1 = 0.17$, respectively, are possible but larger than expected. The inclusion of the reflection of the third harmonic has a small but not negligible effect (σ =5.96 ps without the third harmonic).

5. CONCLUSIONS

Concerning the use of DMTDM in the stability measurement, the noise floor of TDEV = 2 fs can possibly be slightly improved by reducing the non stationary fluctuations related to technical aspects that are not yet optimized. There is a good margin for improvement in long-term performance (increase of the upper bound of flicker PM) and also for the increase of permitted desynchronisation.

Concerning the use of DMTDM in the time-delay measurement, a lot of further investigation has yet to be done. That specially applies to the devising of a more accurate error model. Summing up the results with the error models devised so far, we can draw the following conclusion: in spite of large residuals in the first model and of the lack of self-consistency in the parametric model, the estimated values of the delay of 3.2 ns measured on an inserted coaxial cable agree within 4.3 ps (the uncertainty is not known). The parametric model shows an interesting effect of the third harmonic component (H3), even without phase adjustment.

In principle, the third model is fully self-consistent and it should suppress the drawbacks of the two previous. Its first results in the basic configuration of the DMTD are consistent with the physical reality. The H3 component gives a small but not negligible improvement in the basic configuration without dx. We expect that in the configuration with dx, the improvement will be much larger. The phase adjustment of H3 could bring an additional improvement. The use of the fifth, seventh, … harmonics can also be considered in order to enable an operation with a higher RF level which would provide a lower noise.

It should also be noted that basically we do not have a standard of the delay dx to directly verify the expected measurement uncertainty on the order of 10^{-13} s. In a first step, the DMTDM will have to be used in order to accurately measure the standard cables that are compared in the framework of the EUROMET TF.TI-K1 ”Comparison of time interval (cable delay) measurement”.

ACKNOWLEDGMENTS

The work at IREE has been supported by the Institutional Research Plan (#AV0Z 20670512). The collaboration between IREE and SYRTE has taken place under the EUROMET Project #651.

REFERENCES

[1] D. W. Allan and H. Daams: “Picosecond time difference measurement system”, *Proc. 29th Annual Symp. Freq. Contr.*, Atlantic City, NJ, USA (1975), pp 404-411

[2] L. Šojdr, J. Čermák and R. Barillet: "Optimization of Dual-Mixer Time-Difference Multiplier", *Proc. 18th European Frequency and Time Forum*, Guildford, UK, (2004), CD-Session 6B/130.pdf

[3] R. Barillet, J. Čermák, J.Y. Richard, L. Šojdr: "Application of dual-mixer time-difference multiplication in accurate time-delay measurement”, *Proc. IEEE International Ultrasonics, Ferroelectrics, and Frequency Control 50th Anniversary Joint Conference*, Montreal, Canada, (2004), in press

Time and frequency transfer across Gigabit Ethernet networks

Kenji Hisadome and Masami Kihara*,

NTT Microsystem Integration Laboratories, Nippon Telegraph and Telephone Corporation,
3-1 Morinosato-Wakamiya, Atsugi, 243-0198 Japan
phone: +81 46 240 2276 fax: +81 46 240 3256 E-mail: hisadome@aecl.ntt.co.jp
*Polytechnic University,
4-1-1 Hashimotodai, Sagamihara, 229-1196 Japan
phone: +81 42 763 9226 fax: +81 42 763 9211 E-mail: kihara@uitec.ac.jp

ABSTRACT

Long distance transmission systems are currently based on SDH (Synchronous Digital Hierarchy) or SONET (Synchronous Optical Network) frame configurations, but IP (Internet Protocol) traffic is rapidly increasing, and telecommunication companies are installing Ethernet-based systems that enable them to provide IP-based services directly. To support the reference frequency based functions running on SONET/SDH systems, this paper describes a new frequency and time technique for GbE (Gigabit Ethernet); it provides the reference clock regeneration function offered the current SDH/SONET framework, and the time transfer function used in reference time based services.

1. INTRODUCTION

Telecommunication networks currently use SDH/SONET systems for long distance transmission. The traffic is quite varied but mainly consists of voice and data services such as ATM, Frame Relay and IP. Since SDH/SONET systems usually have a ring-based topology, transformation functions that can convert each traffic format into SDH frames are required at network nodes connecting the ring to customer premises equipment. SDH terminators or ATM switches generally implement this function.

ATM systems offer mature traffic management functions and the dominant choice for supporting general high-speed service in telecommunication networks. However, customer interfaces are changing from ATM to Ethernet following the increasing popularity of IP-based systems and service expansion. Most telecommunication companies now provide Ethernet services that accommodate to Ethernet interfaces such as 100Base-TX and GbE. Low cost WAN and MAN networks can only be realized by creating efficient IP multiplexing/demultiplexing schemes without complicated translations between layers and hierarchies. The common expectation is that GbE and 10 GbE systems will be introduced to telecommunication networks systems without any modification.

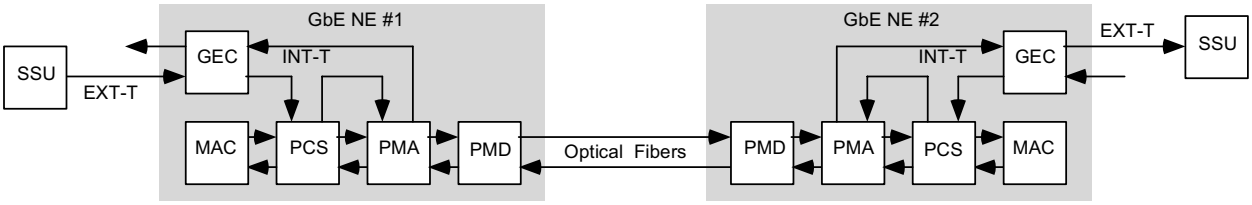
The frequency management function, with which SDH systems are currently equipped, is essential in synchronous transmission protocols, since the function controls synchronous characteristics such as slip and error rates. GbE and 10 GbE systems therefore have to provide the same function, if Ethernet is to coexist with SDH networks. This paper describes a discussion on adding the reference frequency function to GbE systems.

The other reference signal, time, is an additional management function in telecommunication networks. To implement video streaming services, NTT Communications added the reference time function to its WDM-SDH-based ATM switches[1] in its video distribution network; the time function is needed to control the ATM switches. Reference time is also becoming an important tool in computer networks. A transmission characteristic monitor and a data format parameter, both of which are controlled by using a time-stamp based on reference time, determine the distribution time in large file transfers between servers and clients. Precise parameter management based on reference time enables computer networks to transfer large files rapidly. This paper presents a new technique to achieve the reference time transfer function in the GbE physical layer.

2. REFERENCE FREQUENCY CONTROL FUNCTION

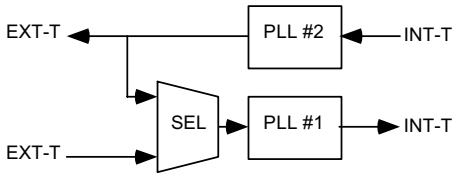
If SDH systems must interwork with GbE systems, the SDH reference frequency distribution paths are cut by the links of the GbE systems. The reference frequency distribution function of GbE systems must allow GbE systems to coexist with SDH systems. The reference frequency transmission method based on the GbE PHY clock frequency of 1.25GHz [2] is promising as SDH systems conventionally adopt the reference frequency distribution system specified in ITU-T G series standards[3]-[6].

The SDH reference frequency distribution system consists of PRC (Primary Reference Clock)、SSU (Synchronization Supply Unit)、and SEC (SDH Equipment Clock) [3]-[6]. The PRC is a master generating the original reference frequency. SSUs are installed as a node system at nodes other than the master, and eliminate jitter components from the transmitted reference frequency. They also hold the regenerated reference frequency while the transmitted reference frequency signal is interrupted due to transmission system failure. This function is called “holdover”. SEC is one of the function blocks forming



MAC: Media Access Control
PCS: Physical Coding Sublayer
PMA: Physical Media Attachment
PMD: Physical Media Dependent
GEC: Gigabit Ethernet Clock

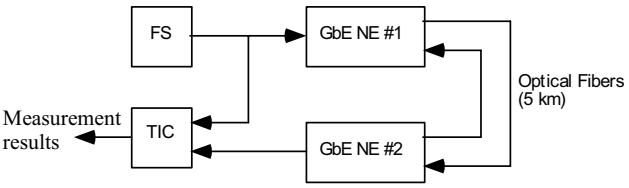
(a) Block diagram of the timing transportation system



PLL: Phase Locked Loop
SEL: Selector
INT-T: Internal Timing
EXT-T: External Timing

(b) Block diagram of the GEC in the GbE

Fig.1 Timing transmission using the GbE NEs



FS: Frequency Standard TIC: Time Interval Counter

Fig. 2 Frequency characteristics measurement configuration

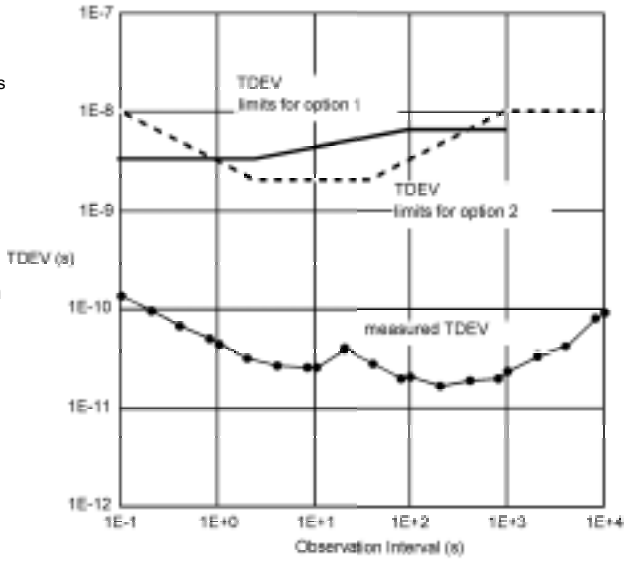


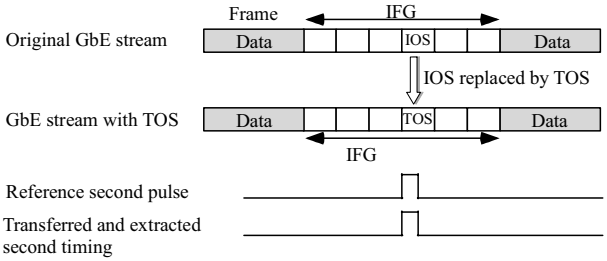
Fig.3 GEC frequency characteristics

the SDH termination equipment. Their roles are mostly the same as those of the SSUs, but there are some differences from the synchronization characteristics specified by ITU-T standards.

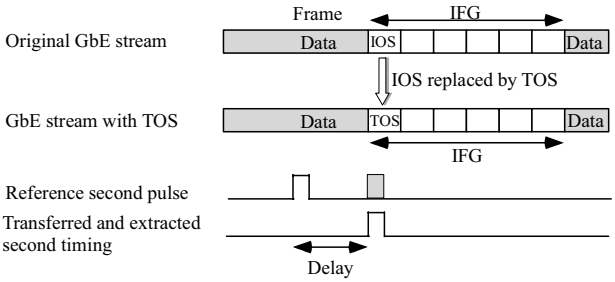
2.1 Reference frequency transfer in GbE systems

The synchronous function is required of GbE systems if they coexist with SDH systems. It should be similar to that of SEC. While the GbE transmission mechanism is asynchronous, its clock frequency, generated at a sender node of the GbE system, is regenerated at the receiver node. SEC functionality can be achieved by using the regenerated clock. The functional block of the clock in GbE-NE, shown in Fig. 1(a), is called GEC (Gigabit Ethernet Clock)(Sublayer functions are explained in 3.2).

This paper examines the case of 1000BASE-LX, which transfers 1300nm optical signals over single fibers, as a MAN/WAN candidate. Since 1000BASE-LX supports full-duplex mode, an Ethernet frame and idle codes are always transmitted. The GEC system shown in Fig. 1(b) is equipped with two frequency conversion circuits to interwork the SDH system and the GbE system timing: PLL#1: from ST (SDH Timing) to GT (GbE Timing), and PLL#2: from GT to ST. Reference frequency characteristics of this GEC system were measured in the test configuration shown in Fig. 2, in which two GbE systems were connected with 5-km-optical fibers, FS (Frequency Standard) supplies the original reference frequency at GbE NE #1, the reference frequency is regenerated by the GEC in GbE NE #2. Figure 3 shows the reference frequency perturbation as evaluated by TDEV (Time Deviation).



(a) IOS replacement with no delay in IFG



(b) IOS replacement with delay in IFG

IOS: Idle Ordered_Set TOS: Timing Ordered_Set

Fig. 4 Time transfer method using IFG

The measured result is much lower than the maximum allowable wander specified by ITU-T G.813; note that this result actually includes double wander due to the two GECs. The GECs examined here used an ordinary VCXO oscillator without temperature compensation as their clock source. The result shows that reference frequency transfer across GbE systems is promising.

2.2 Ring network considerations

Ring networks are often used due to their inherent redundancy which offers enhanced reliability in the face of network malfunction. SDH ring networks adopt the SSM (Synchronization Status Message) transfer mechanism, which can prevent the emergence of reference frequency distribution path loops. The SSM is transmitted by SOH (Section Overhead), a management field of the SDH frame. ITU-T G.781 recommendation specifies SSM transfer times:
(1) Holdover message delay: 500-2000ms
(2) Non-switching message delay: 200ms
(3) Switching message delay: 180-500ms

GbE NEs can be used in the ring network configuration. In GbE ring networks, while SSM can be transferred via the IP layer, it is obvious that its transfer delay will vary with network traffic conditions. The new management information transfer method based on IFGs (interframe gaps) is currently the best approach in GbE systems. The node system in the ring also has to detect input signal down within the specified time to achieve item (1). GbE NEs can detect input signal down through the SD (Signal_Detect) implemented in PMD (Physical Medium Dependent) within much shorter than the specified values. Therefore, these methods enable GbE systems to achieve the SSM function equivalent to that of SDH systems.

3. REFERENCE TIME CONTROL FUNCTION

As we mentioned in 2.2, GbE systems can transfer operation and management information without interrupting GbE data. In the new technique, time information is inserted into IFGs (interframe gaps) that are created by the minimum idle period between adjacent frames. While the IFG function provides only a brief recovery time to permit the reception of the next frame in the original Ethernet framework, our new IFG application transfers time markers that are synchronized to the primary network reference.

3.1 New time transfer method

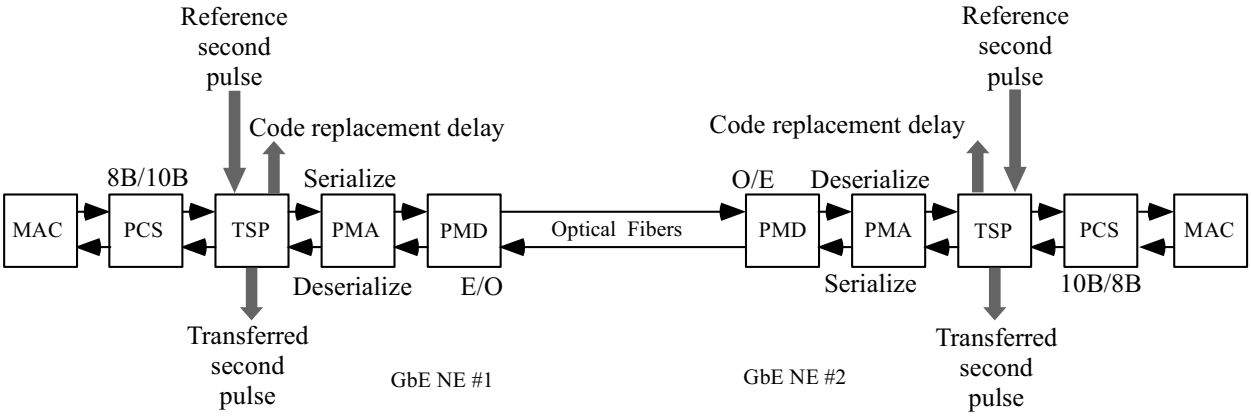
A second pulse is transmitted using special codes in the IFGs of the GbE frame as the reference time in our method, see Fig. 4 (a). However since the second pulse timing is not harmonized with the IFG period, the replacement is delayed until the next IFG if the second pulse is generated at the time when the GbE system is transmitting general GbE data (includes MAC frame). That means that the transferred second pulse timing is delayed as shown in Fig. 4 (b). In our method, this delay is measured at the sender's site, and measurement data are transferred to the other site and used to correct the transferred time as delay data.

3.2 Experimental GbE system

Figure 5 shows the functional blocks of the experimental GbE system. GbE PHY (Physical Layer) is composed of three sublayers: PCS (Physical Coding Sublayer), PMA (Physical Medium Attachment) and PMD (Physical Medium Dependent)[2]. PCSs code to, decode from 8B/10B, and add and delete idle ordered_set consisting of combined special and data codes. PMAs realize SerDes (Serialize and Deserialize). PMDs convert from an electrical signal to an optical signal and vice versa.

A second pulse is transmitted by one of the two proposed codes called TOS (Timing Ordered_Set), /K28.5/D4.7/ and /K28.5/D15.3/, instead of IOS (Idle Ordered_Set), /K28.5/D5.6/ and /K28.5/D16.2/. TOS is chosen so that there is no interference with IOS. The receiver site checks every block of 20 bits in the 10B code stream; if these bits match combinations of /K28.5/D4.7/ or /K28.5/D15.3/, the transmitted second pulse is output after time delay correction.

Figure 6 shows the measured time offset of transferred time at a receiver site relative to the reference time at a sender site. Two 1000BASE-X time transfer boards were connected with a pair of single-mode fibers (each fiber was 5 km long). The two-way time transfer method was used for time synchronization. The resulting time synchronization accuracy was around 2.5 nanoseconds; the delay correction resolution was 6 nanoseconds. Figure 7 shows that TDEV was better



MAC: Media Access Control
PCS: Physical Coding Sublayer
PMA: Physical Media Attachment
PMD: Physical Media Dependent
GEC: Gigabit Ethernet Clock
TSP: Timing Signal Processor

Fig.5 Functional blocks of the experimental GbE system

than 10^{-11} over periods of $10^2 < \tau < 10^3$ sec. The traffic doesn't influence the time transfer performance.

4. CONCLUSION

We suggested a new frequency and time transfer for GbE systems that interwork with SDH/SONET systems. Telecommunication networks are now evolving toward IP-based networks, but most long-distance transmission systems are still based on SDH. Our method will become an important technique to support the coexistence of GbE systems and SDH systems.

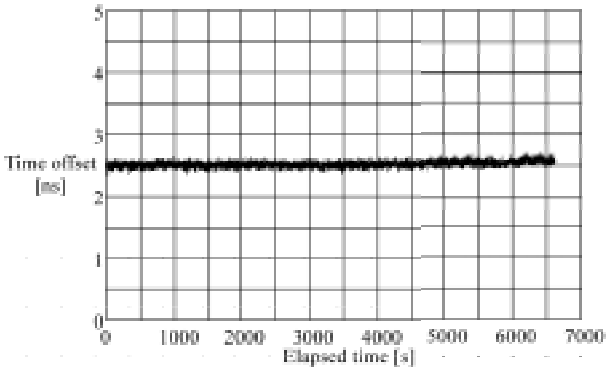


Fig.6 Transferred time offset

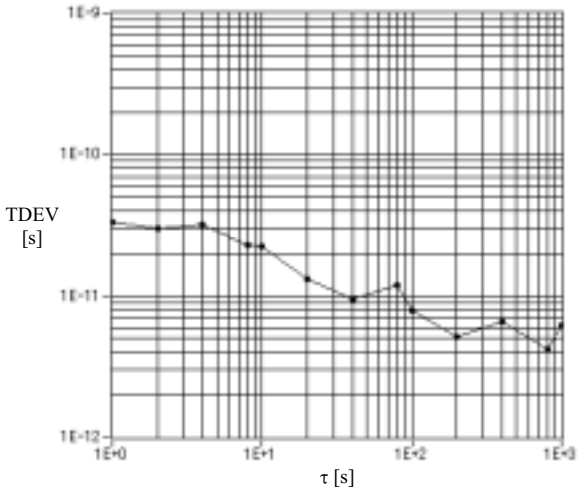


Fig.7 TDEV of transferred time

REFERENCES

[1] M. Kihara, S. Maruyama, K. Hisadome, and T. Hiramatsu, "Time transfer characteristics in a WDM-SDH based ATM system", Proc. European Frequency and Time Forum 2002, pp.A040-A043, (2002)
[2] IEEE 802.3 Std, Part3: Career sense multiple access with collision detection (CSMA/CD) access method and physical layer specifications, 1998.
[3] ITU-T Recommendation G.803, Architecture of transport networks based on the synchronization digital hierarchy (SDH), 2000.
[4] ITU-T Recommendation G.811, Timing characteristics of primary reference clock, 1997.
[5] ITU-T Recommendation G.812, Timing requirement of slave clocks suitable for use as node clocks in synchronization networks, 1998.
[6] ITU-T Recommendation G.813, Timing characteristics SDH equipment slave clocks (SEC), 1996.

Session 3A / Environmental effects on frequency standards

A PHASE LOCKED CRYSTAL OSCILLATOR DESIGN TO WITHSTAND LONG-TERM VIBRATION LOADING

D.Sc Harri Eskelinen, M.Sc Janne-Matti Heinola and Prof. Pertti Silventoinen
Lappeenranta University of Technology, P.O. BOX 20, FIN-53851 Lappeenranta, Finland
Phone: +358-5-6212442, email: harri.eskelinen@lut.fi

1. ABSTRACT

Shocks and vibration are stated to be the worst enemy of precision oscillators [1] and are nowadays very obviously encountered in a portable environment such as in mobile phones and various military gear [2]. Typically a relative frequency change of e.g. 5×10^{-6} is specified or tolerated but very little has been documented about the specific correlation between applied acceleration and induced frequency shifts particularly for complete oscillator circuit designs . In this paper we will concentrate into following new aspects: possibilities to utilize adaptation of the stiffness of the printed circuit board together with the selected joining technology of components and their feeding interfaces and effects of layout of the printed circuit board to the stiffness properties of the mechanical construction. Special attention is paid also to load cases of aircraft applications and to the damping mechanisms for harmful vibrations and impulse forces during different flight operations.

2. INTRODUCTION

In Turin at EFTF 2000 we presented results for a 5 MHz phase locked oscillator, which has been used to evaluate the possibilities to reduce the effects of external shocks by mechanical absorbers. Two different frequency shift types have been recorded. Alternatively the crystal shows a quasi-permanent frequency step or a sharp excursion followed by a decaying ripple. Detailed numerical values are presented in [3]. As we explained in [3] even a slight change in the mechanical arrangement e.g. by immersing the oscillator unit into a viscous fluid moves the natural frequencies or adjusts the damping factor reducing the tendency to long-lasting frequency shifts. In this paper we show the importance of establishing the long time endurance of soldered joints inside the oscillator device and we explain the difficulties to analyze the different load cases during various flight operations (like landing, normal flying, takeoff, turbulence etc.). In this paper we pay special attention to damp the force impulse during the landing to avoid its possible affect to the oscillator performance. We also research the options to damp the harmful vibrations by adjusting the stiffness of the printed circuit board inside the device by tuning its mechanical layout as an individual spring-mass-damper-element.

3. TESTED OSCILLATOR CONSTRUCTION

The tested oscillator construction is a simple 5 MHz phase locked oscillator circuit, which characteristics are earlier presented in [3]. The original crystal oscillator design is modified to better withstand long-term vibrations and individual shocks. The most important changes are: modification of layout of the circuit, use of surface-mounted components and modifications of connector interfaces for providing a spring-mass system. The result of the redesign is decreased weight and size of the oscillator construction. In addition, the construction is more flexible compared to the earlier design.

4. VIBRATION LOAD CASES IN AIRCRAFTS

Aircrafts and their subsystems must be able to withstand the g forces encountered during takeoff, flight and landing. Because the vibration spectrum of frequencies for each type of aircraft and engine are different, it is reasonable to accept that these effects of g forces from impacts and vibrations can be accurately calculated only for simple components and structures. More efficient way is to either measure the vibrations for each particular loading case or try to simulate the critical phenomena by computer aided simulations or use impact tests.

These vibrations during different flight operations range from one vibration per second to thousands vibrations per second. In our case we must therefore take care of three possible risks: At first the impact forces during landing should be limited by a well-designed damping system under the mechanical endurance limit of the device. At second we must ensure that the impact would not disturb the performance of the oscillators. At third there is a self-evident risk of fatigue failure of the soldered joints inside the device due to high frequency vibrations. Because the range of vibrations is varying constantly there is also at least a theoretic risk to meet the resonance frequency of any mechanical component of the device during specific flight operations.

In principal, the structures, which have very low resonant frequencies, are preferred to reduce the g forces and preserve structural integrity. In general, to accomplish this, designing very stiff and lightweight structures is very important in aeronautical engineering. Just like the main landing gear of an aircraft is built to sustain the extreme forces when the aircraft touches down on the runway during landing, all the important devices should be able to withstand that impact force, which is still lead to the aircraft body after the first damping effect. For military purposes it is important to notice that this shock, which is absorbed by the landing gear, is much more extreme than that at touchdown on an ordinary runway. E.g. on aircraft carriers, the landing shock could be up to 5.5 G's

5. MATHEMATICAL ANALYSIS

In [3] we have shown how to analyze the vibration phenomenon with differential equations in those cases where we assume that the aircraft is rolling on the runway. More critical is the load case during landing: it could be regarded more or less as an impulse or impact force, which effects during a very limited and short time period. In theory it would be reasonable to lengthen the possible deflection distance or increase the vibrating mass. However, in practical cases both of these actions are impossible: we would like to keep the size and the mass of the device in minimum. The most effective tool here is therefore the use of suitable damping mechanisms. The value for relative damping factor $\xi = c/c_{cr}$ is the most important one. To analyze the effects of impulse forces ([4] and [5]) we can solve the generalized differential equation by assuming that the other forces than the impulse are meaningless at the effecting time of the impulse, which gives us Eq.1. In this case it will be easy to calculate the amplitude X from Eq.2.

$$x = Xe^{-\xi\omega_n t}\sin\left(\sqrt{1-\xi^2}\omega_n t\right) \qquad \text{(Eq. 1)}$$

$$X = \frac{I}{m\omega_n\sqrt{1-\xi^2}} \qquad \text{(Eq. 2)}$$

However, for more general cases, we do need to use the so-called Duhamel's integral (Eq. 3.) into which we can substitute the real force function F=F(t) caused by landing. The main difference is that instead of maximum impact force F_0 we use the time dependent change of the force $dF=F(\xi)d\xi$. The appropriate test results of various landing cases for different aircraft types under different circumstances can be found from the literature.

$$x = F(0)h(t) + \int_0^t \dot{F}(\xi)h(t-\xi)d\xi \qquad \text{(Eq. 3)}$$

6. VIBRATION ISOLATION SYSTEMS

Three different types of vibration isolation systems can be used according to control schemes: passive, active, and semi-active systems (see Fig. 1).

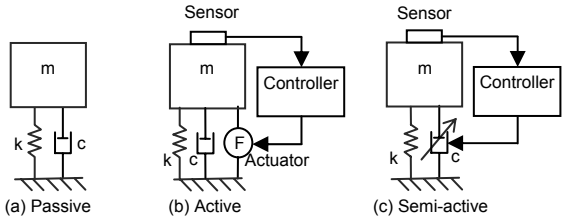


Fig 1. Three different types of vibration isolation systems [6].

In our case the reason of considering the use of either active or semi-active vibration isolation system is their better vibration control of disturbance with varying frequency or large bandwidth. Sensors to activate the system could detect e.g. acceleration of the housing of the oscillator or the position of the landing gear could give the signal to get prepared to the impact force caused by the touch on runway. Active systems are based on systems, which provide an external energy to cancel vibration (e.g. piezoelectric actuators). However, it would make the system even more smart if we do not need to produce any external forces but we could change the systems damping properties according to the vibration loading (ζ or k can be adjusted in real time). MR-dampers are examples of these types of semi-active systems. The basic differences of the damping properties of these three vibration isolation systems are presented in Fig. 2.

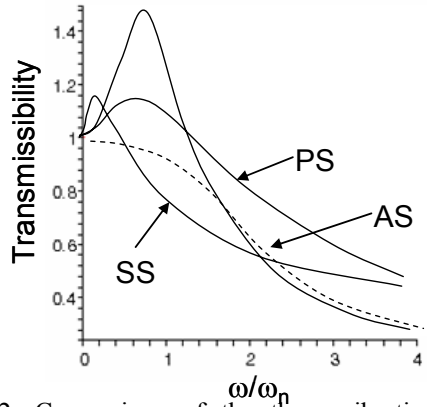


Fig.2. Comparison of the three vibration isolation systems: PS= passive, AS= active and SS = semi-active system (modified from [6]).

The schematic picture of our design is presented in Fig. 3. The supporting mechanics of the printed circuit board is simplified like illustrated in Fig. 4. We wanted to find out, could it be possible to adjust the stiffness of the circuit board according to the other damping properties of the semi-active vibration isolation system. For this purpose we selected two different circuit board materials, which have a ratio of 1 : 0.7 of their spring constants $k_i = [48 \times E_i \times I/l_i^3]$ and modulus of elasticity E_i. A simple impact test and a vibration test were made to check out the possible differences of the oscillator performance and mechanical behavior of the system.

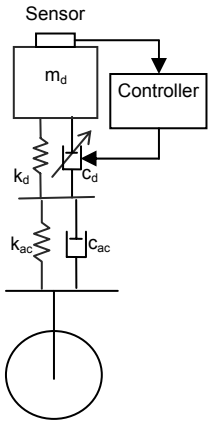


Fig. 3. The schematic picture of our spring-mass-damper design (d= device, ac=aircraft).

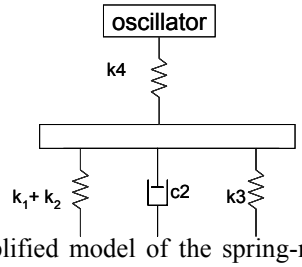


Fig 4. Simplified model of the spring-mass-damper-model of the circuit board ($k_{1...3}$ are the spring constants of soldered connecting pins, k_4 is the spring constant of the circuit board itself and c_2 is the damping factor of the system).

7. VIBRATION ABSORBERS AND DAMPING SOLUTIONS

Application of a tuned vibration absorber (TVA) is sometimes the best option for control of unwanted vibration. This countermeasure is particularly appropriate when the vibration issue occurs for a single frequency, or across a very narrow frequency range. The TVA basically consists of a mass and spring system, which is added to the structure at a location, which exhibits high displacement for the objectionable event. The mass is often selected from considerations of the modal mass of the structure for the resonance of the structure, which contributes to the objectionable vibration. However, performance of the TVA is not strictly limited to resonance response, though it will have its most profound effect for resonance response of a structure. It is sometimes possible to induce a "valley" in the response of a structure for a narrow frequency range using the TVA approach, even when no resonance exists there. The TVA will be most effective approximately where its own (designed) actuation resonance frequency occurs. The tuning frequency is adjusted via the TVA mass and the stiffness of its spring element. This spring element is typically a complex spring having stiffness and damping. A viscoelastic material is typically utilized for the spring element of the TVA. If we know exactly the vibration phenomenon it is possible to design and produce TVA assemblies to within 1% to 2% of the target tuning frequency [7].

In [8] an investigation of a magnetorheological (MR) semi-active system with a MR fluid damper is presented. The study utilizes feedback linearization, which is demonstrated experimentally, to convert the non-linear damper into a linear controllable device. Using a validated model of an MR damper, it is demonstrated that feedback linearization can yield significant performance improvements over more simplistic on/off control strategies. The same strategy could be integrated within larger scale vibrating structures such as aircraft landing gear to implement more complex control strategies, e.g. optimal control. In our case this type of damping technology would solve e.g. the vibration problems during landings and occasional turbulence during flights if the damper is connected directly with the housing of the oscillator device.

In [9] a vibration control of flexible structures using MR and piezoceramic is presented. In that research, two different mounts fabricated from two different smart materials were proposed for the vibration control of a flexible structure subjected to external disturbances. A semi-active magnetorheological (MR) mount was firstly devised and its damping force controllability by the current intensity was evaluated. Subsequently, an active piezoceramic mount was devised and its field-dependent inertial force was experimentally evaluated. Each mount was then installed on the flexible structure and the corresponding control model was established. The LQG controller was designed for each mount system and empirically realized. Control responses of each control system such as acceleration and displacement were investigated in the frequency domain. [9] In our case the whole device could be assembled on MR-gel layer. This would make it possible to use active

damping but the problems with required magnetic field changes should be solved.

8. VIBRATION AND IMPACT TESTS

Calculated values of moments of inertia $I_{x,i}$, spring constants k_i for two different devices equipped with printed circuit board materials having different values of modulus of elasticity E_i are presented in table 1. Also the calculated critical natural frequencies $f_{cr,i}$ for both cases are presented.

Table 1. Calculated and measured values of test cases.

Material i	E_i [MPa]	$I_{x,i}$ [mm^4]	Σm_i [g]	Σk_i* [N/mm]	$F_{cr,i}$ [kHz]
Teflon	75140	1,2555	3,4784	331,909	1,554
FR4	53700	1,0698	3,2461	282,636	1,485

Two interesting results can be discovered: The natural frequencies are set to the critical area of normal vibration spectrum during aircrafts typical flight operations – the result, which we did not expect to achieve. However, we could prove that it is possible to adjust the value of natural frequency simply by tuning the spring constant of the construction either by changing the modulus of elasticity of the printed circuit board material or by changing the supporting mechanism of the board. In this case the difference of +/- 5 % was calculated for two material types.

During the impact tests we tried to increase the value of the impact force up to level about 5 G's. The value which interrupted our tests was about 4.5 G's – not because we would have noticed some malfunctioning of the crystal itself but because of the fracture of the soldered joint between the connecting pin of the crystal enclosure. We repeated the shock test for both construction types and the same phenomenon took place. Later we noticed that the enclosure was assembled to the most critical place on the circuit board: the maximum bending moment did affect on the soldered joints during the impact.

9. CONCLUDING REMARKS

To be able to select the most appropriate damping system for different types of flight operations (landing, takeoff, normal flying and turbulences) it is important to establish the critical vibration load cases. In addition to this the specific analysis for modified impulse force case during the landing is necessary to ensure the performance of the oscillator: It seems fairly reasonable to improve the performance of phase locked crystal oscillators by adding semi-active shock damping structures based on some smart material solutions either to the landing gear construction and/or to the oscillator device and phase locked loop circuit itself. The soldered joints inside the device, though they were properly soldered, could also be critical. Additional to this, the joining pins and the printed circuit board tend to act like a spring-mass-damper-system and that mechanism should be tuned according to exact vibration analysis of the oscillator device.

REFERENCES

[1.] Efratom, Precision Time and Frequency Handbook ,1996
[2.] Neubig B., Das grosse Quarzkochbuch, Franzis´ Verlag, 1997
[3.] Eskelinen H., & al. "Shock tests with a phase locked crystal oscillator", Proc. of the 14th EFTF-conference, Turin 14.-16.2.2000
[4.] Bishop, R. & Johnson D., The Mechanics of Vibration, Cambridge University, 1970
[5.] Timoshenko, S., Vibration Problems in Engineering, Van Nostrand Co., Inc. New York 1966
[6.] Tantanawat T., "Application of Compliant Mechanisms to Active Vibration Isolation Systems", Compliant Systems Design Laboratory, 2005
[7.] Lewis T., "Tuned Vibration Absorbers and Tuned Dampers", Damping Technologies, Inc., 2005
[8.] Batterbee, D.C., Sims, N.D. "Skyhook damping with linearised magnetorheological dampers" Proc. of SPIE - The International Society for Optical Engineering, v 5386, Smart Structures and Materials 2004
[9.] Choi, Seung-Bok, &al..:"Vibration control of flexible structures using MR and piezoceramic mounts", Proc. of SPIE - The International Society for Optical Engineering, v 5390, Smart Structures and Materials 2004

ACKNOWLEDGEMENTS

The authors of this paper are grateful to Professor Pekka Eskelinen (Helsinki University of Technology, Finland) for his important advice and his supporting and encouraging attitude to motivate the authors to continue this research, which has been started under his guidance during year 1999 at Lappeenranta University of Technology, Finland.

Space radiation effects on Ultra Stable Oscillators

Michel Brunet*, Robert Ecoffet*, Guy Rolland*, Gilles Cibiel*
JJ.Boy**

*CNES 18, avenue Edouard Belin, 31401 Toulouse Cedex, France
Michel.brunet@cnes.fr

** FEMTO-ST/LCEP – 26, ch. De l'Epitaphe - 25000 Besançon, France
jjboy@ens2m.fr

ABSTRACT

All satellites use Ultra Stable quartz crystal Oscillators (U.S.O.) and need very stable frequency stability and low phase noise compatible to space environment. Space radiation USO sensitivity should be considered carefully for short term, medium term and long term stability, depending on the use of the USO in the equipment. The paper gives USO dose and dose rate determination method, and examples of frequency variations observed in orbit and during ground experiments. Preliminary analysis is also given.

1. INTRODUCTION

Ultra Stable quartz crystal Oscillators are widely used in space systems, and all space programs need to verify the End Of Life (E.O.L.) frequency stability over the lifetime (few years to 20 years), as well as short term, medium term stability, and phase noise.

Table 1 gives some examples of O.B. USO specifications

Applications	Function	OB USO Stability specs
Telecom-munication	Local oscillators	0.2 to 1 ppm over 20 years
Precise positioning	GPS	1.10^{-10} (1s to 1000s)
	Argos Doppler positioning system	5.10^{-12}/min over 10 min 0.2 ppm over 5 to 10 years
	Doris precise positioning system	4.10^{-13}/min over 10 min 0.2 ppm over 5 to 10 years

***Table 1**: Example of OB USO specifications*

Theses specifications should be compliant with all on board environmental conditions mainly temperature, magnetic field from earth and spacecraft, and radiations.
To verify theses radiation specifications, it is necessary to determine the radiation environment of the USO, and the radiation USO sensitivity.
This paper gives methods to determine the dose and dose rate received by USO in orbit, examples of frequency variations observed in orbit and during ground experiments. Preliminary analysis is also given.

2. RADIATION USO SENSITIVITY DETERMINATION

The radiation USO sensitivity can be determined in 4 steps:

step 1: determination of the dose and dose rate received by each part of the USO in orbit. Method recommended by CNES is given in the paragraph 3.

step 2: determination of the sensitivity of each USO components (crystal resonator, electronic parts) in radiation, versus dose and dose rate received by theses components in orbit. Component radiation sensitivities can be found from existing data bases or from particular radiation experiments.

step 3: determination of the frequency sensitivities versus USO components: it is necessary to use non linear model which gives relationships between USO frequency and all USO components values (resonator, transistor, amplifier, capacitor, varactor…). The results for each components is expressed in $\Delta F/F$ in ppm versus component variation in ppm.

step 4: determination of the USO radiation sensitivity: from step 2 and 3, it's possible to calculate for each components of the USO the frequency variation vs radiation, and then the global effect by cumulating all component variations.
This method can be used for all types of USO (X.O., TCXO, OCXO), for E.O.L, short/medium term stabilities as well as phase noise. Results are

dependent of the accuracy of the model and of the quality of the data used. Experiences show that USO frequency sensitivity to radiation come mainly from the crystal, and from sensitive electronic parts such as thermistor, varactor, and transistor. It is strongly recommended to compare theorical results with experimental data in applying radiation (protons, gamma…) directly to representative USO to measure the frequency shift over radiation in order to verify all specifications.

3. USO DOSE and DOSE RATE DETERMINATION

The radiation USO dose and dose rate are determined from several parameters: duration of the mission, launching date and orbit parameters. The tools used are environment modelling codes and 3D Monte Carlo particles transport codes which need geometrical and physical description of the satellite, equipment, and USO itself.

The dose rate calculations are made at evenly spaced trajectory points. At each of these trajectory points **P,** we can define a local environment spectrum $S_e(\mathbf{P}, E)$, where E stands for the energy of the particles. These local spectrum computations are made thanks to an environment code (SPACERAD, SPENVIS or OMERE for example) which takes into account the orbit parameters. It can be demonstrated that the dose rate $D(\mathbf{P})$, for each component inside the satellite, is deduced by a simple numerical integration from the environment spectrum $S_e(\mathbf{P}, E)$ thanks to a transfer function $F_D(E)$:

$$D(\mathbf{P}) = \int_0^\infty F_D(\xi) \cdot S_e(\mathbf{P}, \xi) \cdot d\xi$$

$$F_D(\xi) = \int_0^\infty Q(E) \cdot R_i(E, \xi) \cdot dE$$

where, $Q(E)$ describes the sensitivity of the component to the dose and $R_i(E, \xi)$ describes the shielding properties of all the materials around the component (component can and the satellite equipment and structure). There is one transfer function for each component.

The transfer function $F_D(E)$ is computed using numerical simulations. These numerical calculations are made by using the Monte Carlo transport code «NOVICE». First, it is necessary to construct a geometrical and physical model of the satellite. This is done with the FASTRAD code which allows us to construct interactively theses descriptions. It is clear that the use of a transfer function saves a lot of computation time because this function is obviously the same for each point **P** and thus it is constructed once only.

The example given hereafter concerns the computation of the dose rate induced by the protons belt at the heart of the quartz inside one of the JASON_1 USO. In **Figure 1**, the trajectory crosses the heart of the SAA and then crosses the antipode of the SAA. Thus, we observe only one intense dose rate peak.

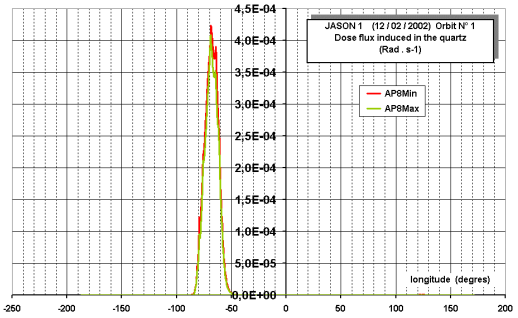


***Figure 1** : Dose rate in the resonator when JASON_1 crosses the heart of the SAA.*

Regardless of the dose rate peaks computations, we are equally interested in the knowledge of the total mission dose integrated at different points inside the USO (on critical components, for example). These doses computations are made by using the environment flux spectrum corresponding to the total mission duration. All the environment particle types and the braking radiation of the electrons have been taken into account. The computations are made by using the Monte Carlo adjoint method of the NOVICE code. To do that, we need again the geometrical and physical description of the USO and the satellite. In **Figure 2** represents an example of such a result. It concerns the ionising dose and the atomic displacement dose at three points of the electronics and in the heart of the resonator.

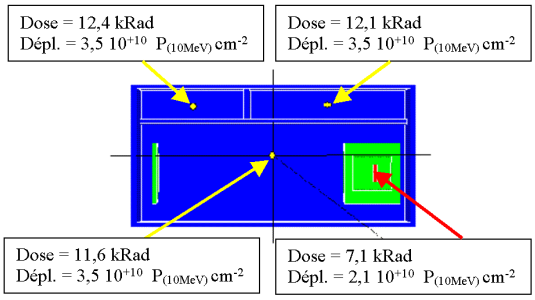


***Figure 2** : Ionizing and atomic displacement total doses inside the US*

4. Radiation effects on USO

4.1 Frequency variations observed in orbit

Figure 3 and **table 2** give the frequency variations observed on several DORIS USO placed on board several satellites.

DORIS (Doppler Orbitography and Radiopositioning Integrated In Space) is a CNES, IGN and GRGS French space program able to determine very accurately (within 1-2 cm) the position of satellites and ground beacons [1]. The DORIS O.B. payload has been integrated into several satellites since 1990, SPOT 2 to 5 series, TOPEX, Jason 1, Envisat, and will be embarked on Cryosat, Jason 2... Spot series and Envisat are at 830 km, TOPEX and Jason 1 at 1300 km. Using several ground beacons synchronized by atomic frequency standard, it is possible to determine the on board frequency USO over 1 pass (10 min) with an uncertainty of few 10^{-12}. The DORIS E.O.L. and mid term performances in orbit are compliant for each USO, except for Jason 1 for which large frequency variations are observed.

Figure 4 gives the frequency variations of the DORIS USO observed in july 2002 and april 2003 on board Jason 1 satellite over a period of one day. We can observe frequency variations of 1 to 6.10^{-11} peak to peak, strongly correlated to the pass of Jason 1 over the S.A.A.

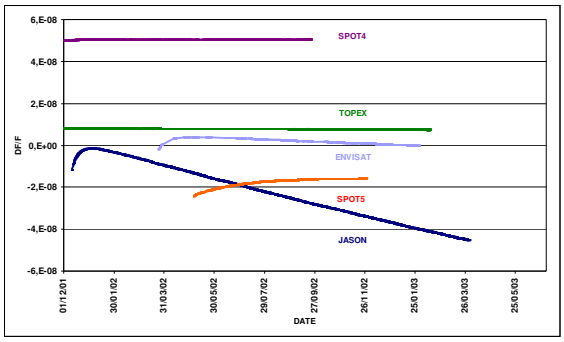


***Figure 3**: 10MHz DORIS USOs frequency drift in orbit for different space programs*

Satellite	SPOT 4	SPOT 5	TOPEX	JASON 1	ENVISAT
T0 : date of USO turn on	March 23 1998	May 4 2002	Dec 23 1998	Dec 8 2001	March 14 2002
Aging per day	+ 1.0 E-12	+ 4.2 E-11	- 1.8 E-12	- 1.0 E-10	-1.7. E-11
After	54 months	4 months	45 months	10 months	6 months

***Table 2** : Aging slope per day of different 10 MHz USO in orbit*

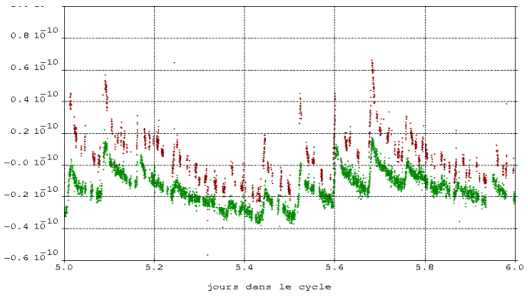


***Figure 4**: Frequency variations in orbit for the Jason 1 DORIS USO*

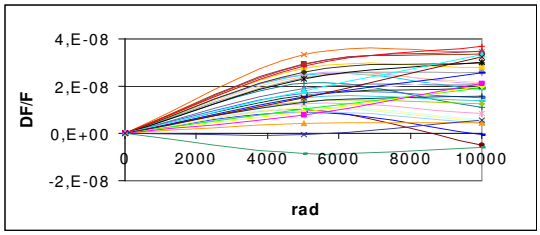
4.2 Frequency variations observed during Gamma ray testing

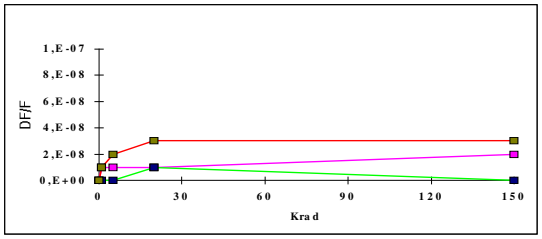
Several experiments were conducted using several type of sources (electron, protons, gamma), at high and low dose rate, applied on crystal resonators and USO. Informations can be founded in several publications of RADECS, IEEE-Frequency Control Symposium, IEEE-UFFC, IEEE-Nuclear Sciences, and proceedings of this Forum.

CNES has conducted several experiments using gamma ray source at high and low dose rate on AT and SC cut quartz resonators.

4.2.1 Frequency variations at high dose rate

Figures 5 and **6** from [4] give results of frequency variations observed on several 10 MHz SC 3rd overtone crystal resonators submitted to gamma rays at 5 krad per hour, from 0 to 150 krad. Frequency variations can be positive or negative, sometimes positive then negative. The major variations arrived during the first 10 to 20 krad. The radiation sensitivity is in the range of 10^{-12}/rad up to 10 krad, and of 10^{-13} /rad around 50 krad.





***Figures 5 and 6:** Frequency deviation of 10 MHz SC 3rd Ov resonators vs irradiated dose at 10 krad, and between 0 to 150 krad*

4.2.2 Frequency variation at low dose rate

Figure 7 represents an example of frequency variations observed on USO when it is applied gamma rays at 4 rad per hour, simulating a satellite pass over the S.A.A.

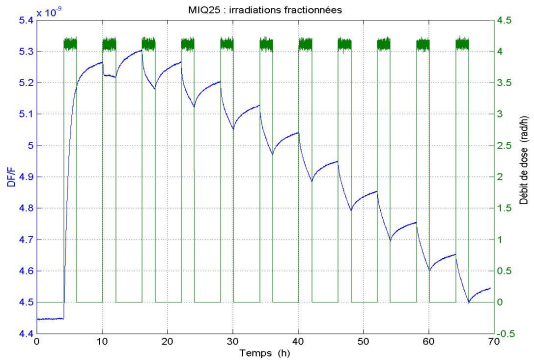


***Figure 7**: Frequency variations versus time observed during discontinous γradiations (4 rad/hour)*

Similar effects were observed by Suter and Norton [9, 10] and Cash [11] using low dose rate protons.
All experiments conclude that the frequency variation is non linear and could be positive or negative. The sensitivity is higher at low dose rate than at high dose rate.
J.Bourrieau [6, 7] had shown that USO frequency variations versus discontinuous radiations can be modelled by the following equation:

$$(1) \quad f(t) = a_0 + a_1 t + a_2 e^{-t/\tau_2} + a_3 e^{-t/\tau_3}$$

where:

a_0 and a_1 are the USO frequency drift coefficients without radiation (long term drift)

a_2 and τ_2 are the amplitude and time constant coefficients of the component 1

a_3 and τ_3 are the amplitude and time constant coefficients of the second component

This model show that the drift observed have two origins, one transient change of the frequency (rapidly saturated and annealed), the other producing a permanent damage (maybe with a very long term annealing).
Amplitude, sign and time constants of the phenomena associated to both components explain the various signatures observed by CNES as well as those reported by Sutter and Norton [9, 10].
Up to now, any physical explanation has been found.

4.3 USO radiation sensitivity analysis

Radiation USO sensitivity can come either from electronic circuit, and crystal resonator.

4.3.1 Component radiation sensitivity

Electronic parts are generally defined and selected in order to minimise the radiation effect. However, the frequency of USO is very dependent of some components' behaviours such as amplifier, transistor or varactor, which can explain part of the USO radiation sensitivity.

4.3.2 Resonator radiation sensitivity

The most important effect is generally attributed to the crystal resonator. Several origins (temperature, mechanical parts, quartz crystal…) could explain the observed frequency variations.

4.3.2.1 Thermal effect

Temperature variation induced by discontinuous gamma rays and measured on the can of the quartz crystal resonator is smaller than 2.10^{-5} °C, and so cannot explain the observed frequency variations. [2].

4.3.2.2 Mechanical radiation effect

Mechanical deformations have been measured on quartz discs before and after several Mrad of gamma dose. Deformations observed by J. Bourrieau [8] cannot explain this phenomenon.

4.3.2.3 Ionizing radiation and material defects in quartz crystal

It is well known that resonance frequency and Q-value of quartz resonators are altered by exposure to X- or γ-ray. Measurements show that these radiations induce on a select resonance frequency long live and transient changes in the acoustic parameters of the equivalent circuit R_m, L, C and, to a less degree, C_0.
Moreover, these measurements show the limitations and, sometimes, large frequency shifts, due to material properties. Indeed, sensitivities to ionising radiation can vary strongly from one to the other one quartz blank, and even, from one growth zone to other ones of same synthetic quartz block. So, we can deduce that the ionic impurities and the dislocations are mainly involved in the degradation of the expected performances of a "time base" quartz oscillator. We report here the first interpretations:

1. the transient frequency change is interpreted to be due to the variation of the elastic modulus of quartz brought by the detrapping of the positively charged compensation ion at substitutional Al^{3+} (or Fe^{3+}) defects in the quartz lattice [12]
2. the large permanent frequency offset experienced by natural quartz or synthetic crystal with high level of impurities was first interpreted by King [13] and treated more generally by King and Sanders in [14]. They explained that the annealing of this frequency change is simply the diffusion controlled

uncorrelated recombination of the cation H^+ (probably), with the Al^{3+} trap.
Therefore, it has been shown that limitations in performances (aging, long term drift, retrace…) of a resonator can be essentially due to the migration of impurities as Na^+, Li^+… or H^+. The frequency shift is correlated to the concentration of these impurities, its variation being related to their motion and nature (in term of migration energy).
Consequently, to use quartz resonators in severe environmental conditions as in space applications, it is necessary to select the raw quartz material (and particularly to control the parameters of the synthetic quartz growth) or to determine treatments to improve frequency stabilities. Of course, the bad performances being due to the diffusion of interstitial impurities, they could be modified using the sweeping process (which is an electro migration under high electric field, at high temperature and under vacuum or not) which theoretically permits the elimination of these mobile impurities.
As it has been illustrated by the **Figure 8** below, AT-cut resonators [16] present frequency changes induced by ionizing radiation due to specific crystal defects which cause:

- stress relaxation
- increasing or decreasing of the mechanical coupling of the resonator with its stress field (coupling between stress and the acoustically active defect).

So, if radiations have altered a defect structure in a way to cause stress, the effective thickness-shear elastic modulus determining the frequency is modified.

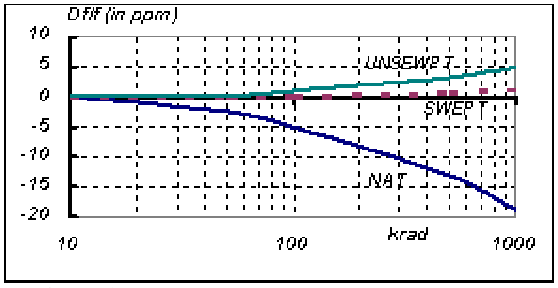


***Figure 8**: Frequency change of 5 MHz USO using natural, cultured unswept and swept quartz.*

4.3.3 Frequency variations at low dose rate in presence of proton and Gamma: preliminary results

Most of USO radiation problems observed in orbits around the earth are due to protons particle. Studies have shown that proton radiation presents a particularly serious threat to piezoelectric resonators. Indeed, frequency shifts in quartz resonators induced by proton radiation are not only dependent on total dose but also on dose rate and proton kinetic energies. J.J. Suter, J.R. Norton [ref 9, 10] and Cash [ref 11] have already worked on quartz resonators which were BVA AT-cut. Then, in order to simulate the proton environment encountered in space by a SC cut resonator built in unswept quartz crystal, we have to model the actual spectrum of this radiation. A quantitative study of a specific low-earth orbit (TOPEX orbit) led to the development of proton simulation tests. A preliminary test has been reached with a synchro-cyclotron at the Center Proton-therapy of Orsay France (CPO). The proton test configuration was the following:

- dose rate: 10 rad/h
- kenetic energy: 110 MeV
- orbit model (= 1 cycle): 10 min radiation on / 110 min radiation off

The **Figures 9** and **10** present the preliminary results of the frequency susceptibility of one specific PFM (Proto Flight Model) USO due to respectively γ rays and proton particles. These results are very surprising. Indeed, the USO seems to be non sensitive to proton particles. However, several points must be treated above to conclude at an eventual non-susceptibility. First of all, the proton beam used was not continued, it was a periodic (22 ms) pulse (20 µs, 10 krad/h) and, more over, the noise floor of the measurements bench was too high and must be improve for next tests.

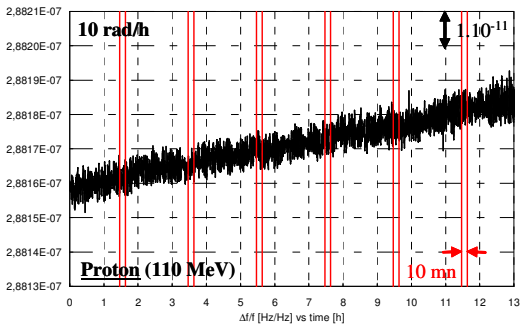


***Figure 9**: Frequency susceptibility of a PFM USO to proton radiation*

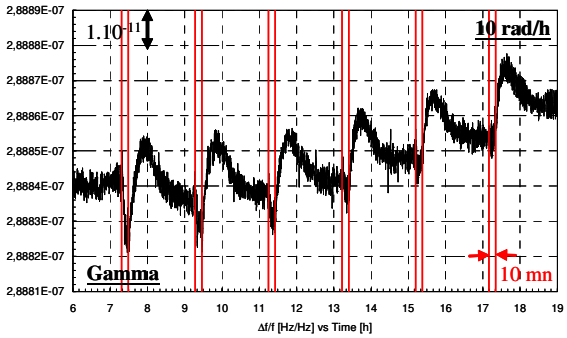


***Figure 10** : Frequency susceptibility of a PFM USO to gamma radiation*

5. Conclusions

All space programs use widely Ultra Stable Quartz crystal Oscillators (U.S.O.) and radiation sensitivity should be considered carefully for short term, medium term and long term stability.

Crystal dose and dose rate should be determined taken into account all parameters, orbit, shielding between spacecraft and components, and duration of lifetime in orbit. Up to 20 to 50 krad at crystal level should be considered.

Variations in orbit over years show drift, including ageing and radiation effects, of several 10^{-8} compatible to the End Of Life typical specifications of 0.2 to 1 ppm over 20 years. For some applications such as sciences or precise positioning systems where medium stability is require, frequency variations when the satellite passes over the SAA show very important variations up to 10^{-10} peak to peak.

Numerous studies and experiments on quartz materials, quartz resonators and USO were conducted using electrons, protons, neutrons, and gamma source. Doses versus frequency change is non linear, and is larger at low doses. Frequency change is quartz impurity dependant, and is not well understood. Quartz material with very low impurity and dislocation densities should be used to reduce radiation sensitivity; pre irradiation at 10 to 20 krad is also recommended by several agencies.

R&T study managed by the CNES with the French experts in quartz material (ICMCB, LCEP, LPMC), in quartz resonator (LCEP), in quartz oscillator (C-MAC, LCEP, LPMO), and in radiation (CPO, LSI, ONERA) has been started in 2004. This study includes 3 stages:

- Bibliography study
- Material study
- Resonator and oscillator study

Five quartz crystals, one natural and four synthetics from the following suppliers: GEMMA (FR), HIRST/ASTRIUM (UK), SAWYER (USA), VNISSIM (R) have been selected.

The goal of this study consists in the determination and understanding of the mechanisms responsible of the radiation sensitivity of quartz material, resonator and oscillator in order to reduce the radiation sensitivity of quartz USO.

ACKNOWLEDGEMENTS

The authors acknowledge the CNES DORIS team, ONERA-DESP and C-MAC for theirs contribution and cooperation.

6. References

[1] M.Brunet et al: Doris precise orbit determination and localization system description and USOs performances; Proceedings of 1995 IEEE FCS

[2] M.Brunet et al: Dérive de fréquence observées sur des OUS soumis à des irradiations continues et discontinues à faible débit de dose, note technique CNES, octobre 1992

[3] J.Delporte: Dérive de fréquence d'oscillateurs équipés de nouveaux résonateurs QAS sous irradiation à faible débit de dose, note technique CNES, juin 2001

[4] V.Candelier et al: Recent progress in ultra stable oscillators for space on board and ground applications; Proceedings of 2001 EFTF, Neuchatel

[5] V.Candelier et al: Space qualified 5 MHz ultra stable oscillators; Proceeding of 2003 IEEE FCS, Tempa, Florida

[6] J.Bourrieau et al: A semi quantative model of the dynamic behaviour of space USOs under fragmented and continous irradiations; EFTF 1993

[7] J.Bourrieau et al: Dynamic behaviour of ultra stable oscillator under irradiation at low dose rate

[8] J.Bourrieau: Etude des variations dimensionnelles du quartz irradié ; Note technique ONERA-DESP, Juillet 1998

[9] JJ.Suter and al: Low and high dose proton irradiation of quartz crystal resonator; PTTI 1987

[10] JJ.Suter et al: Low and medium dose radiation sensitivity of quartz crystal resonators with different aluminium impurity content; FCS 1987

[11] P.Cash et all: Low dose proton radiation sensitivity of quartz resonators; EFTF 1996

[12] J.C. King, H.H. Sander –Proceedings of the 27th AFCS (1973), pp. 113, 119.

[13] J.C. King – Bell system Tech. J. 38, 573 (1959)

[14] J.C. King, H.H. Sander – IEEE Trans. on Nuc. Sci., NS-19, N°6 (1972), pp. 23, 32.

[15] J.J. Suter, A.G. Bates, J.M. Cloeren –Proceedings of the 2d EFTF (1988), pp. 679, 686.

[16] J.J. Boy et al.– Proc. of the 6th EFTF (1992), pp. 93, 98.

[17] M.Brunet et al.– Space radiation effects on Ultra Stable Oscillators, SREC 2004,14-18 june 2004.

[18] G.L.Weaver et al.– Examination of detailed frequency behavior of quartz resonators under low dose exposures to proton radiation, Proc of the 2004 IEEE Frequency Control Symposium.

TEMPORAL STABILITY MEASUREMENT ON OCXO RUNNING UNDER SEVERE ENVIRONMENTAL CONDITIONS

O. FRANQUET, O. ROMAND, L. CAMELIN, A.R.ELECTRONIQUE F-25047 BESANÇON
B. WOLCOFF, A.R.ELECTRONIQUE F-78500 SARTROUVILLE

Espace Valentin Nord – B.P. 3113 – F 25047 BESANCON CEDEX
Tél. : (00.33.3) 81.88.66.77 – Fax : (00.33.3) 81.88.33.55
Email : are@ar-e.com

1- Abstract

Usually, for the evaluation of temporal stability, and in order to know the top of performances on high stability OCXO, the DUT is measured in the best metrological conditions such as :
- oscillator supplied with a source without any ripple or spurious
- climatized laboratory
- oscillator stabilised for a long time in inert conditions
In these conditions, we obtain the well-known Allan Variance curves with a typical floor for τ between 0,1s and 100s about few 10^{-13}.

AR Electronique, involved for a long time in the manufacturing of oscillators for airborne applications has investigated on different ways to characterize oscillators in severe environmental conditions, used in applications like tactical UAV. In order to provide high reliability products, we have to introduce dedicated characterisation methods which well illustrate the final system behaviour.

2- Introduction

The technical requirements for oscillators involved in airborne application like RADAR, GPS, DATA COMMUNICATION lead us to develop high reliability OCXO's with preserved frequency stability performances in spite of the environemental conditions.
Frequency stability is only needed during the flight (few hours) but we meet shocks and high level of vibrations, pressure variation (altitude or velocity), extended and fast temperature changes.
The life time of the equipment is about 10 years, but the real operational running time is about few thousand hours.
According to this special chopped profile of life, long term ageing is not the main parameters. Warm up time, frequency retrace and thermal behaviour are more important in order to reach a stabilized frequency as fast as possible.
The oscillator structure is devoted to tactical applications :

-Very fast warm up
-Temperature stability $<5.10^{-9}$ in the range [-40°C, +85]°C
-Very low noise
-Very low g sensitivity ($<2.10^{-10}$/g worst axis)

An adaptation of performance test procedures to the real need of the final application lead us to choose Allan variance analysis under environnemental conditions.
All the main frequency characteristics vs environment have been performed by Allan Variance analysis under :

- temperature rate and warm-up
- altitude variation effects

Especially regarding vibration effects, a comparative analysis was made with both phase noise direct measurement in all axes and Allan Variance under vibration characterization. The results show a proportional degradation of the temporal stability when G sensitivity or level of vibration increase.

3- Oscillator design

3.1 Electronic circuit :

The 10 MHz crystal is associated to a low noise discrete oscillator electronic. The output is buffered by a FET amplifier. The typical phase noise floor of this structure is about -160 dBc/Hz.
A dedicated ASIC developed by ARE control the thermal regulation of the ovenized part.
In airborne application the external supply are often pollued (more than 100mV peak to peak ripple in a few MHz bandwith). Both, a passive filter and a voltage regulator are placed in order to filter the supply ripple.

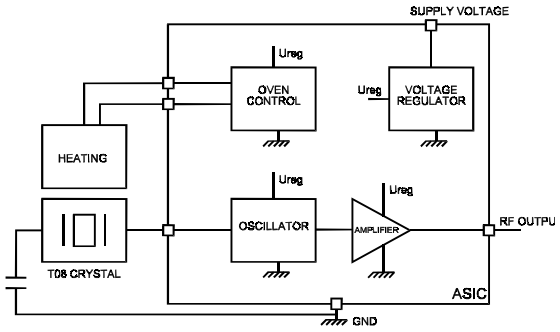


Fig 1: Oscillator design

3.2 Thermal and mechanical structure :

Oscillators for tactical application required very fast warm up. We have choosen to only ovenized the crystal in order to reduce the time to stabilize the frequency. An electronic compensation corrects the temperature sensitivity of the electronic.
All the assembly is designed in order to optimize the reliability with mechanical and thermal shocks.
The oscillator is enclosed in a hermetically sealed case. This is necessary to reduce the effect of altitude variation on frequency. Each manufactured oscillator is unitary controlled under gross leak test.

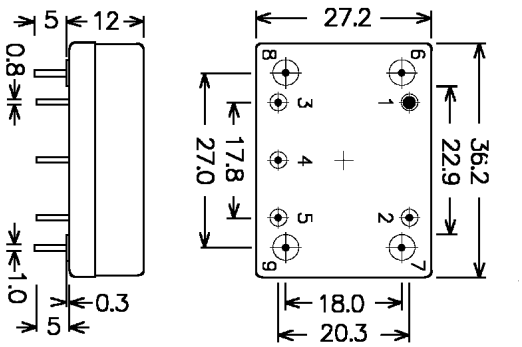


Fig 2: Case dimmensions

3.3 10 MHz crystal :

The resonator used for such tactical application is a 10 MHz SC cut crystal , mounted in four points in a TO8 enclosure. The self mechanical resonances of the structure are rejected up to 6kHz.

AR Electronique has investigated on a new design of the mechanical structure of the 3 rd overtone SC cut 10 MHz crystal resonator. The optimization focused on the minimization of the Gsensitivity in compliance with manufacturing batches of few hundred pieces per month. The result as shown on the following graphs (samples of different months) is more than 95% of the batch under 5 10^{-10}/G.in the worst axis.

Erreur ! Liaison incorrecte.
Erreur ! Liaison incorrecte.

Erreur ! Liaison incorrecte.

Fig 4 a, b, c . Worst axis Gsens on several batches

Note : - All the Xtals are unitary measured in each axis after the manufacturing process, in order to place the crystal, if necessary, in an optimized position in the final application.

- All the values measured < 1 10^{-10} /G are clipped to 1 10^{-10} /G.

All these investigations for few years lead us to divide the average Gsensitivity per batch by a 2 to 3 ratio on 10 MHz SC Xtal and extended these results to 100 MHz AT Xtal families.

AT 100 MHz Xtal oscillators are often the spectral reference in synthesis chain of airborne radar. AR Electronique has focused a part of Xtal developpement for these kind of application.

For the following, study we have choosen two 10 MHz crystals. One with very low Gsens parameters and the other bad in regard with the average Gsens.

The two crystals have been continuously measured in Gsens (per 10° angle) in the X,Y,Z plane.

For instance, we show the measured results / in the XY plane for the first, and XZ plane for the second.

Erreur ! Liaison incorrecte.

Fig 5: Gsens. Cartography XOY plane OSC 51109

Erreur ! Liaison incorrecte.

Fig 6 : Gsens. Cartography XOZ plane OSC51117

The theorical Gsensitivity distribution follows a sinus law (two tangential circle on our graph). But the dispersion in the hardened mounting structure of the plate, and the crystal orientation deform the curve and modify the Gsensitivity repartition.

4- Performances caracterization and experimental results

4.1 Phase Noise :

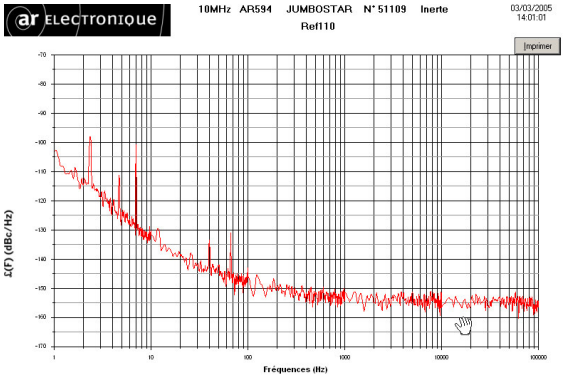
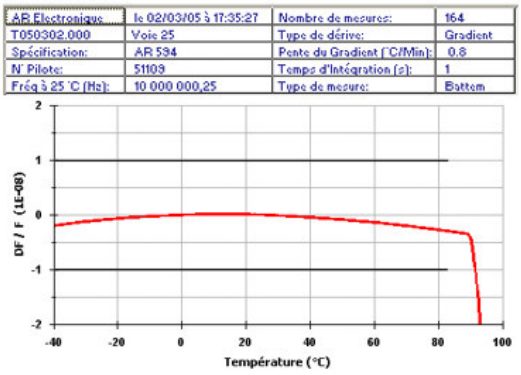


Fig 7 : OSC. 51109 Phase noise inert

4.2 Thermal behaviour :



AR Electronique	le 02/03/05 à 17:35:27	Nombre de mesures:	164
T050302.000	Voie 25	Type de dérive:	Gradient
Spécification:	AR 594	Pente du Gradient (°C/Min):	0.8
N° Pilote:	51109	Temps d'Intégration (s):	1
Fréq à 25 °C (Hz):	10 000 000,25	Type de mesure:	Bottom

Fig 8 : Frequency vs temperature

The slope in the worst condition is about 5 10^{-11}/°C. In order to optimize the ageing / temperature compromise, the Xtal is ovenized on the first turning point. It's the reason why the frequency decreases when the thermal regulation stopped (after 90°C).

.

4.3 warm up :

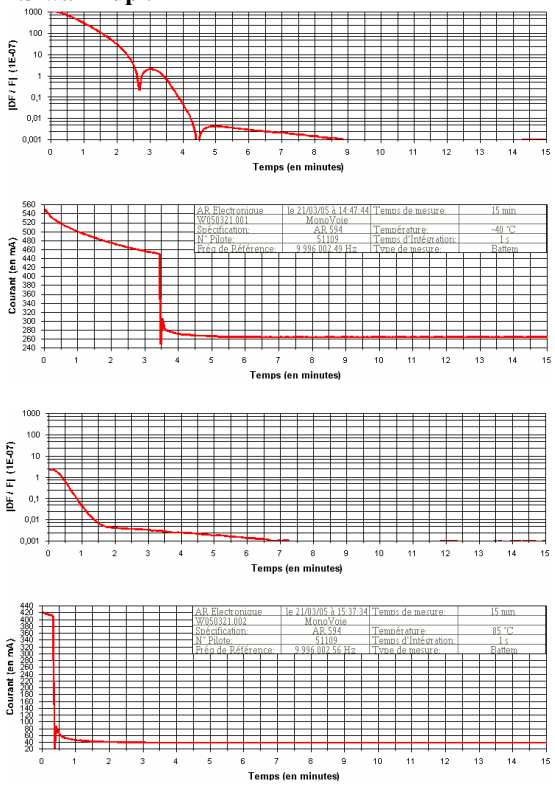


Fig 9a, b, c : Warm up @ -40°C, 85°C

Only the Xtal is ovenized, this provide the very fast warm up. The power consumption at -40°C is about 9W during 150 s.

4.4 Allan variance inert:

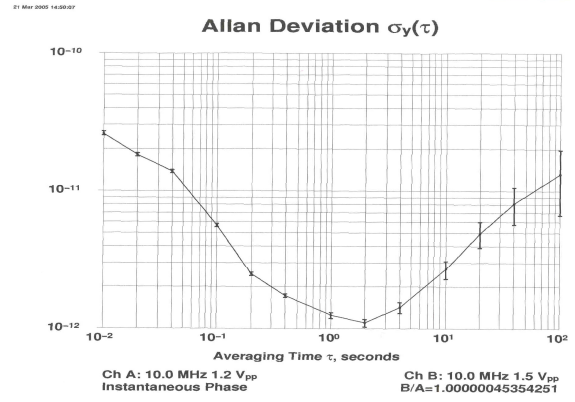


Fig 10 : Allan variance on osc 51109 (inert)

The measure has been performed after 12 hours stabilisation at ambient temperature.

4.5 Phase noise and Allan variance under vibration

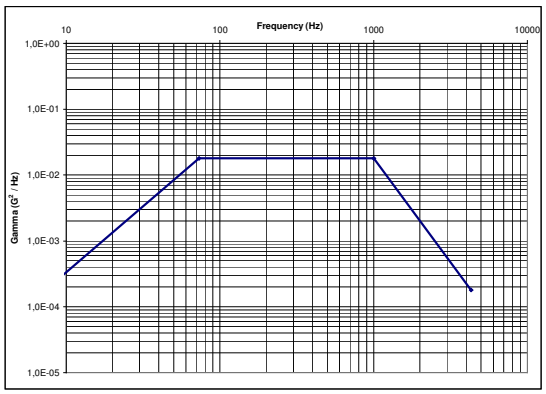


Fig 11 : Vibration excitation profile

This profile is in compliance with norm 'MIL STD 202 method 214 Cond. I Test Cond.A

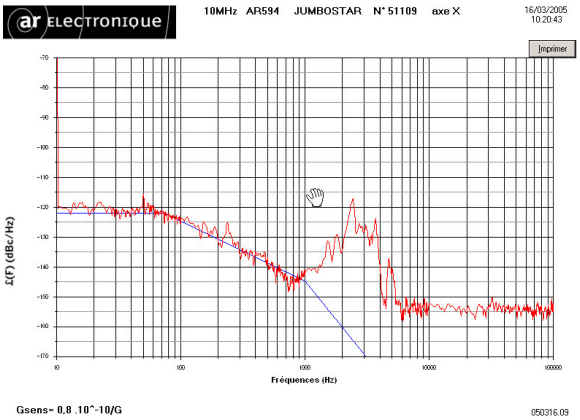


Fig 12 : phase noise degradation for the nominal level

The experimental curve reaches the computed response up to 1 kHz, mechanical structure resonances affect the behaviour between 1 and 4 kHz.

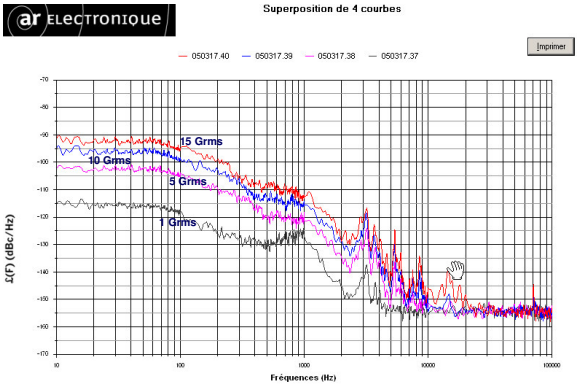


Fig 13 : Vibration level increasement effect

The phase noise degradation is given by the formula :

$$L(F) = 20LOG\frac{Gsens.Fosc.\sqrt{\gamma}}{Fvib.\sqrt{2}}$$

with :
Gsens : G sensitivity
Fosc : Oscillator frequency
Fvib : Vibration frequency
γ : Acceleration g^2/Hz

When the acceleration level increases, we obtain a vertical translation on phase noise. With temporal analysis, we observe the same effect for time constant t comparable to the period of vibration.

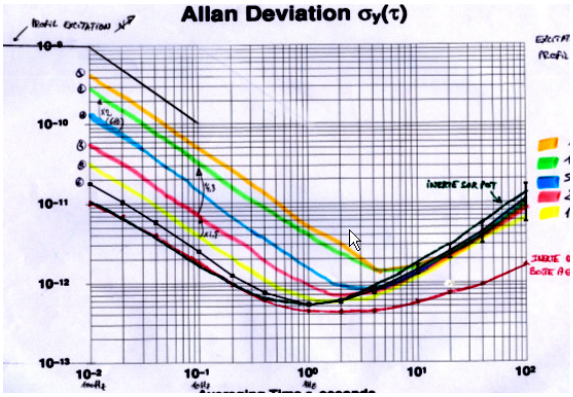
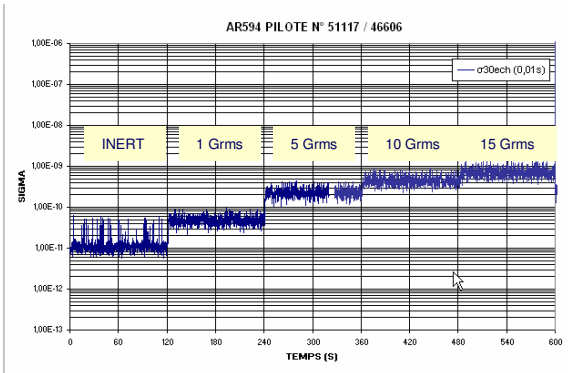


Fig 14 : Allan variance under vibration

The following curves show the comparison between low and standard Gsens oscillators with temporal analysis method :

Osc 51117 : Gsens 8.2 10^{-10}/G



Osc 51109 : Gsens 1 10^{-10}/G

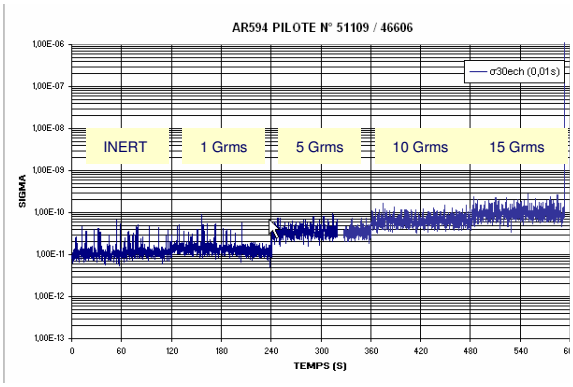


Fig 15 a, b : Allan variance under vibration

When vibration level increases we find the same ratio for the degradation of the variance, and the level of degradation is proportional to the Gsensitivity.
The ratio between the two curves is equivalent of the ratio of the Gsens. of the two oscillators.

4.6 Allan variance with pressure variation :

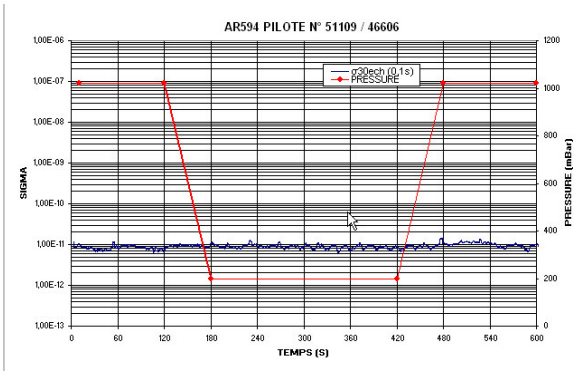


Fig 16 : effect of pressure variation

The oscillator is hermeticly sealed, this limits the direct effect of pressure. Only thermal exchanges modifications are involved in the stability evolution.

4.7 Internal Thermal Gradient Effect cumulated with Gsensitivity :

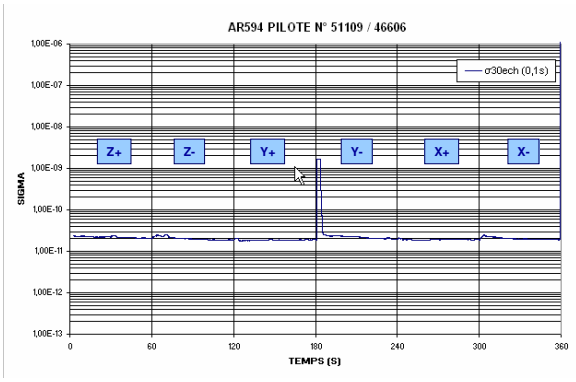
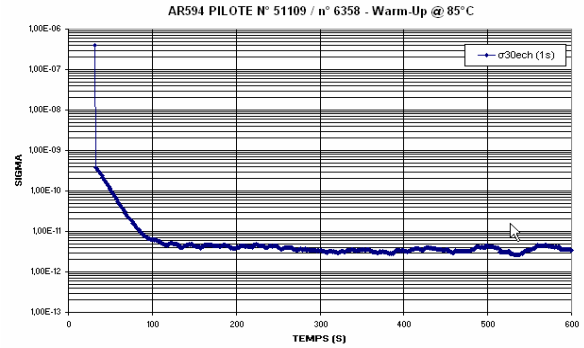


Fig 17 : Osc.51109, ITGE effect on variance

ITGE contribution appears when thermal exchanges are not the same for different position of the oscillator, it's naturally cumulated with G sensitivity. With this oscillator design the thermal behaviour is constant and has no important effect on Allan variance.

4.8 Allan variance during warm up



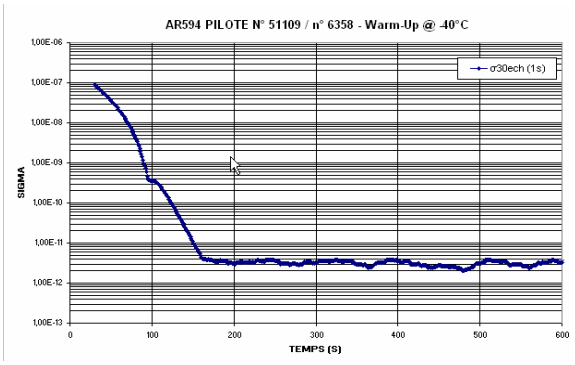


Fig 18a, b : Variance during warm up at extreme temp.

We reach a level of 1 10^{-11} Allan variance (τ=1s) under 100s to 200s, depending of the external temperature

4.9 Allan variance evolution during thermal gradient

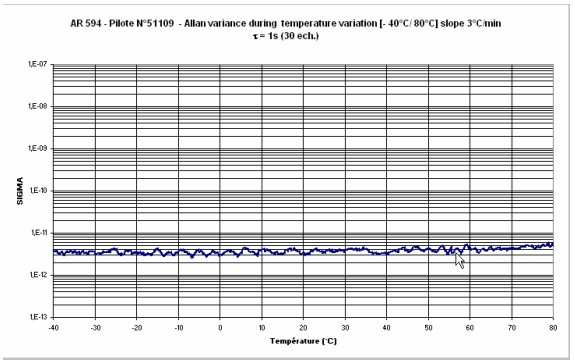


Fig 19 : Allan variance with temperature gradient 3°C/min

The fast temperature evolution don't affect the Allan variance.

4.10 Supply variation (12 V +/- 5 %) :

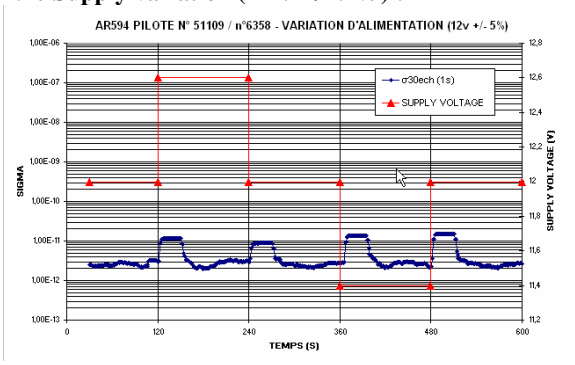


Fig 19 : effect of supply variation

The main contribution of the variation is provided by the modification of power of heat inside the oscillator. The effect on regulated voltage is negligible.

4.11 Load variation (50 Ω +/- 5%) :

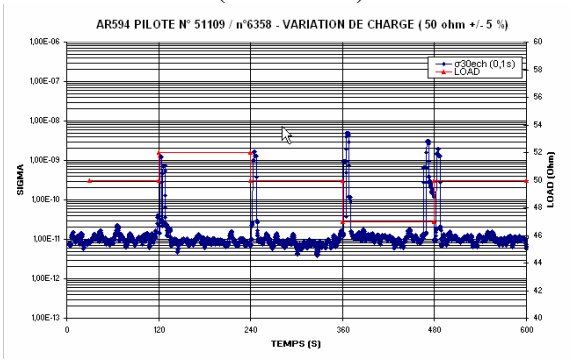


Fig : 20 effect of load evolution

A small variation of load impedance don't affect the variance (efficiency of the output amplifier / buffer).

Note : -All the Allan variance measures have been performed with the Timing Solutions TSC 5110 A system.

-The frequency reference was a 10 MHz OCXO from AR Electronique.

- The phase noise inert and under vibration is performed by a licence LPMO / AR Electronique system

5- Conclusion

We have identified the main factors involved in the frequency stability when the oscillators are perturbed by the environmental condition of tactical applications.
The first is the warm up. If we want to reduce the stabilisation time we have to deliver more and more power at the start up of the oscillator, this can be antagonist with the improvement of the reliability (high thermal shock).
The second is the G sensitivity. It has to be minimized by the optimisation of the Xtal plate mounting structure or additionally by a mechanical filtering of vibration if necessary.
For this kind of oscillator, the others parameters provide a small contribution in the Allan variance contribution.

Even if the temporal analysis gives less information than phase noise analysis , Allan variance under environmental conditions is useful in production rate to characterize the performance of the OCXO's. This method is very efficient when we just need to obtain the global temporal stability aspect. No locked oscillator is needed to perform the measure , however we loose spectral informations like mechanicals resonances detection or spurious localization.

This experimental study is the first part of AR Electronique new way of investigation for developpement of the new OCXO involved in tactical applications.

6- References

R. L. FILLER and J.R. VIG
The acceleration sensitivity of quartz crystal oscillators
Review proc. 41st ASFC (1987)

J. RUTMAN and J. GROSLAMBERT
Characterization and measurement of frequency stability
EFTF(1990)

E. GIRARDET and B. WOLCOFF
High stability, miniature OCXO's for severe environemental conditions
IEEE – IFCS (1999)

O. FRANQUET and B. WOLCOFF
One gigaherz low phase noise oscillator for on board avionic applications

TEMPERATURE REGULATED SAPPHIRE RESONATORS FOR LOW PHASE NOISE OSCILLATORS.

R.Boudot[1], C.Rocher[1], N.Bazin[1], S.Galliou[1], V. Giordano[1], O.Llopis[2]
[1] FEMTO-ST Institute, LPMO and LCEP Department - University of Franche-Comté UMR6174 - 32 avenue de l'Observatoire, 25 044 Besançon Cedex, France
[2] LAAS-CNRS, 7 av.du Colonel Roche, 31077 Toulouse, France.
Corresponding author : rodolphe.boudot@lpmo.edu

ABSTRACT

Technical solutions are proposed to solve the problem of the thermal stabilization of a high-Q whispering gallery mode microwave sapphire resonator. This technology enables the design of very low phase noise microwave sources which can be used as references for phase noise measurements or for high sensitivity radar systems. Nevertheless, the sapphire resonator presents at the ambient temperature a high sensitivity to thermal fluctuations of the order of -70ppm/K limiting the oscillator stability on the long term. An original resonator thermal configuration has been designed and optimized thanks to finite difference technique simulations. The symmetry of the proposed thermal structure greatly limits the influence of the thermal gradients and then increases the thermal regulation efficiency. Coupled to a high-resolution electronic regulation directly embedded on the microwave cavity, this structure presents a very low sensitivity to the environnement thermal fluctuations. This sensitivity is better than -0.05ppm/K.

1. INTRODUCTION

In the microwave frequency range, the sources presenting the lowest phase noise are based on the high-Q Sapphire Whispering gallery Mode Resonator. Combined with ultra-low phase noise amplifier [1] or with interferometric phase noise correction technique [2], these resonators permit at the ambient temperature the achievement of phase noise as low as -140dBc/Hz at 1kHz Fourier frequency. Unfortunately, the sapphire resonator presents a high sensitivity to the thermal fluctuations. This sensitivity (around -70ppm/K) results from the strong dependence of the sapphire permittivity with the temperature. It results that the intrinsic oscillator relative frequency instability of the order of a few 10^{-11} is masked by the thermal fluctuations for integration times higher than 0.1s. A very efficient temperature control has then to be implemented.

The resonator consists in a monocrystal sapphire puck which has typically a diameter of 30 to 50 mm and a thickness of 10 to 30mm. This resonator is placed in the center of a metallic cavity generally made in copper. The thermal regulation of such a large thermal mass reveals itself to be a difficult task. The temperature sensor is usually separated from the crystal resonator whereas this one is the element which has to be thermally controlled. Moreover, the resonator is not surrounded by a perfectly isothermal enclosure. Noncontrolled thermal paths are still found between the outer medium and the sapphire. In other words, the resonator is still sensitive to ambient temperature changes while the sensor is locked at the reference temperature.

Some solutions have been already proposed to stabilize the resonator temperature around 300K using a Peltier element as actuator [3] or a thermal sensitive quartz oscillator as thermal sensor [4]. Nevertheless, at our knowledge, the thermal analysis of the resonator structure has not focused in the litterature at lot of attention.

In the first part of this paper, a commercial finite element simulator enables to foresee the thermal behavior of the system taking into account the exact geometry, the physical characteristics of the materials constituting the resonator and its thermal shielding. In that sense, the importance of the symmetry of the thermal resonator configuration in relation to its electromagnetic one's is demonstrated.

In a second part, we propose a simple electrical analog model for the resonator structure which can be combined, in a CAD software, with the electronic controller. This second approach enables us for different configurations to evaluate the thermal gain, the inner temperature gradients as well as the warm-up time.

Eventually, a high-precision electronic temperature controller has been developped and tested on an optimized symmetrical resonator structure. The thermal behavior predicted by the two models has been confirmed with a great accuracy. Especially, a high rejection of the temperature fluctuations has been obtained leading to a thermal gain higher than 1000.

2. THE SAPPHIRE RESONATOR

The oscillator we developped is based on a $50mm$ diameter and $20mm$ thick cylindrical sapphire resonator excited on the quasi-transverse magnetic whispering

gallery mode $WGH_{5,0,0}$ at 4.8 GHz. Such resonators, characterized by the lowest dielectric losses at microwave frequencies, enable high Q-factors [5] (about 200.000 at 293K). The crystal is placed in the center of a $100mm \times 40mm$ cylindrical metallic cavity to limit radiation losses. The sapphire is supported by one or two 10 mm diameter cylindrical posts as schematized in figure 1. A modal selector is deposited on the top of the resonator as described in [6].

Currently, a copper cavity is used. In order to obtain a portable and lighter structure, we opted for a duralumin cavity. This material presents an easier machining, a cheaper price and interesting thermal characteristics. Indeed, compared to copper, the warm-up time to reach the reference temperature is reduced because the duralumin mass density-specific heat product is smaller. Moreover, simple tests have shown us that Q-factors obtained with a duralumin cavity are the same than those obtained with a copper one.

The possibility to cool the resonator has been discarded in order to keep a portable system even if the resonator Q-factor increases at low temperatures [7]. Around the ambient temperature, the Q-factor thermal sensitivity has been measured to be :

$$\frac{1}{\Delta T}\frac{dQ_0}{Q_0} \simeq -6.25 \times 10^{-3}\ \mathrm{K}^{-1} \qquad (1)$$

We chose to stabilize the resonator at 313K (40°C) where the unloaded Q-factor is still large and equals to 170.000.

3. Thermal finite element model

The analysis is carried out with a commercial finite element software assuming that materials are homogeneous and their thermal properties are independent of temperature in the $273 - 320$K range. The main physical phenomena that participate to heat transfers are conduction and convection. Radiation is neglected because of the working temperature range. Compared to conduction fluxes, its effects are negligible. Also, the microwave dissipated power (10mW) inside the resonator is neglected face to the heating power needed to maintain the steady state (about 2W).

The whole system is divided into a few elementary volumes. Symetries are considered in order to simplify the analysis. Contacts between elements are supposed perfect. Inside the cavity, the air can be considered as quasi-motionless. Therefore, only conduction in the air is taken into account. A temperature of 320K is then applied to some localized areas, represented by solid arrows in figure 1. We consider in those simulations that the cavity is surrounded by an insulating foam whose thermal conductivity is $0.05W.m^{-1}.K^{-1}$. The value attributed to the convection coefficient of the outer ambient air is h=$25W.m^{-2}.K^{-1}$. Indeed, convection fluxes ϕ_{conv} take place outside the structure such as:

$$\phi_{conv} = hS\Delta T \qquad (2)$$

where S is the surface of the area where the convection flux takes place and ΔT is the difference of temperature between the air and the foam's one.
Finite element simulations enable us to show the temperature distribution inside the cavity in steady state for two different configurations drawn in figure 1.

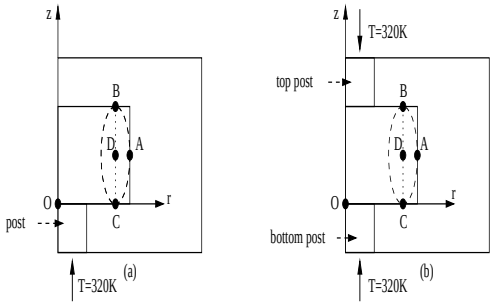


Fig. 1. Different simulated configurations with a finite element software.

The symmetry of the geometry is improved in configuration (b) where the crystal is thermally linked to each lid of the cavity through a duralumin post. In the whispering gallery mode resonator, the electromagnetic fields are confined in a volume near the curved air-dielectric interface as schematized in figure 1. Then, we analysed the temperature in this region for the two configurations.
In term of temperature sensitivity, we first analysed the thermal gain G_{th} of the assembly, that is to say its sensitivity to ambient temperature T_{amb}. The reference temperature is fixed at 320K. Figure 2 shows this result on the test point D.

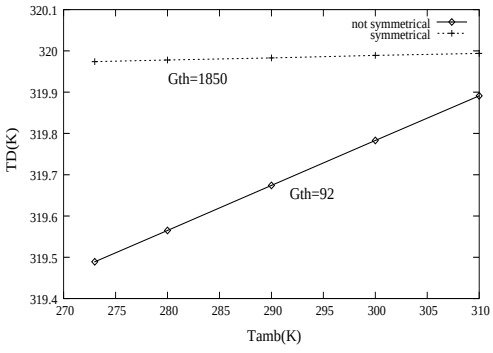


Fig. 2. Simulated thermal gain of the asssembly for the 2 configurations.

We can observe that the thermal gain is divided by 20 in the configuration (a) where the symmetry of the

structure is not optimized. Indeed, in the worst case, simulation results give a change ΔT_D=0.402K at the point D of the sapphire for an outer temperature change ΔT_{amb}=37K, which leads to a thermal gain ΔT_{amb} / ΔT_D=92. In the other case (b), the simulated thermal gain becomes 1850.

Moreover, the temperature gradients into the crystal can be obtained. Figure 3 represents the temperature change along the resonator radius r in the equatorial plane (z=+10mm). The outer temperature is fixed at 300K and the reference temperature is still at 320K. The mean temperature is different for each configuration and then, the results are normalized to the temperature T_A obtained at the point A (r=25mm, z=+10mm). This temperature is equal to 319.783K in the first configuration (a) and 319.989K for the second one (b). The same analysis has been done in the axial direction for r=20mm. It's obvious that temperature gradients are minimized on the symmetrical structure.

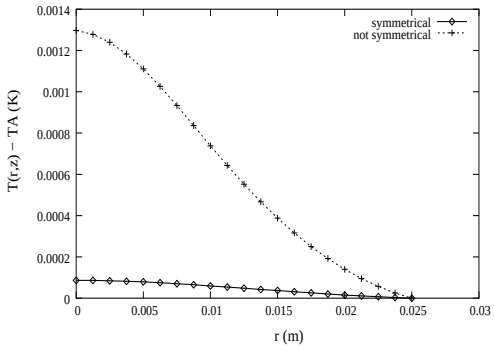


Fig. 3. Temperature distribution in the crystal along its radius (r, z=+10mm) for a 300K outer temperature.

The evolution of these gradients as a function of T_{amb} is also estimated. Figure 4 presents the maximal temperature gradients into the crystal along the axial (r=20mm, z) direction for T_{amb} varying from 273 to 310K.

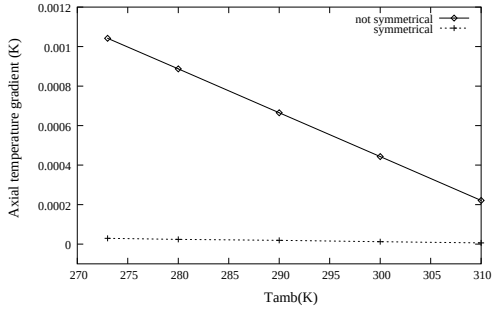


Fig. 4. Maximal innner temperature gradients in the crystal along its height (r=20mm, z) vs T_{amb}.

In the best case (b), for a 290K outer temperature, the maximal temperature gradient along the height of the resonator is about 0.019mK whereas it's about 0.66 mK in the non-symmetrical configuration (a). Along the radius of the resonator (r, z=+10mm), the order of magnitude of maximal inner thermal gradients is higher. For a 290K ambient temperature, the maximal temperature gradient along the radius of the resonator is respectively for the symmetrical configuration and the non-symmetrical one 0.2mK and 2mK. Now, whispering gallery modes propagate at the air-dielectric interface at the periphery of the resonator, therefore it's primordial to limit thermal gradients in this part of the crystal.
Even if these ones seem very small, they are far from being negligible. The temperature gradient effect on the resonator frequency can be calculated by applying the perturbation method to any electromagnetic resonator model such as the matching model [8].

$$\frac{\delta f}{f} = -\frac{\varepsilon_0 \int\int\int \vec{E}[\delta\varepsilon]\vec{E^*}dV}{2\int\int\int \varepsilon_0[\varepsilon]\vec{E^*}dV} \quad (3)$$

where $\vec{E}$ and $\vec{E^*}$ are respectively the electric field and its complex conjugate calculated for a perfectly homogeneous temperature inside the resonator. $[\varepsilon]$ is the permittivity tensor which is a function of the cylindrical coordinates, ε_0 is the vacuum permittivity. $[\delta\varepsilon]$ defines the permittivity shift due to the thermal disturbance.
To give a crude approximation, we can assume a simple model for an axial gradient for which $\delta\varepsilon$ is a non-null constant only in the half-top of the resonator. In this condition, equation (3) gives:

$$\frac{\delta f}{f} = -\frac{1}{4}\frac{\delta\varepsilon}{\varepsilon} \quad (4)$$

Thus, from (4), we can estimate that the resonator frequency thermal sensitivity to an inner thermal gradient δT is:

$$\frac{1}{\delta T}\frac{\delta f}{f} \simeq -33,5 ppm/K. \quad (5)$$

As for the symmetrical configuration, we have δT = -0, 6 μK/K $\times$ ΔT_{amb}, the temperature gradient is responsible for a relative frequency variation of 2×10^{-11} for a 1K ambient temperature variation. Thus, we can consider that the effect on the resonator frequency of inner temperature gradients is negligible face to the change of the crystal mean temperature. Regarding these results, the second configuration (b) has been preferred and chosen.

4. Modeling using the electrical analogy

The principle of this model is inspired of [9]. According to the electrical analogy, temperatures are represented by voltages and heat fluxes by currents. The structure

is divided into 7 elementary volumes: one for each cavity lid (V_1 and V_3), one for the cavity cylindrical wall (V_2), one for each supporting post (V_4 and V_7) and two volumes for the sapphire resonator (V_5 and V_6). Each volume V_i is represented by a node connected to a thermal capacitor C_i such as:

$$C_i = \rho_i \widehat{C_i} V_i \tag{6}$$

where $\widehat{C}_i$ is the specific heat of the material and ρ_i its mass density.
A capacitor C_i is fed by input "currents" flowing through resistors, which link the node representing the volume to neighbouring nodes. The continuity of currents is naturally set between neighbouring volumes belonging to the same solid. Then, a volume V_i is connected to another one V_j through resistors named R_{ij}. In that sense, each resistor covers a surface supposed thermally homogeneous and is defined by:

- for axial fluxes:

$$R_{ij} = \frac{1}{\lambda}\frac{l}{S} \tag{7}$$

  λ is the themal conductivity, l the length of the element and S the section surface.

- for 2π radial fluxes:

$$R_{ij} = \frac{ln(\frac{r1}{r2})}{2\pi\lambda h} \tag{8}$$

  λ is the themal conductivity, h the height of the element. r1 and r2 are the radius of the concentric circular surfaces.

Resistors R_a, R_b, R_c, R_d link nodes belonging to the sapphire crystal. These resistors describe radial conduction fluxes along the radius of the resonator. The equivalent electrical network is set up on figure 5. Important nodes linking the elemental volumes are noted T_a, T_b,...T_j. One heating element is localized on each extremity of the posts. A sensor is close to one of them. The temperature of this one is equal to T_a. The temperature of the sapphire center is equal to T_d.

Inside the cavity, convection fluxes are neglected and only conduction in the air is taken into account. Similarly to the preceding section, the thermal losses are distributed over the entire surface as if there were an insulating foam 5cm thick with a thermal conductivity of 0.05W/m/K.
The main advantage of this model is that it can be combined with the CAD description of the feedback control system. To analyse the closed loop operation, we assume in a first attempt a simple proportionnal correction.

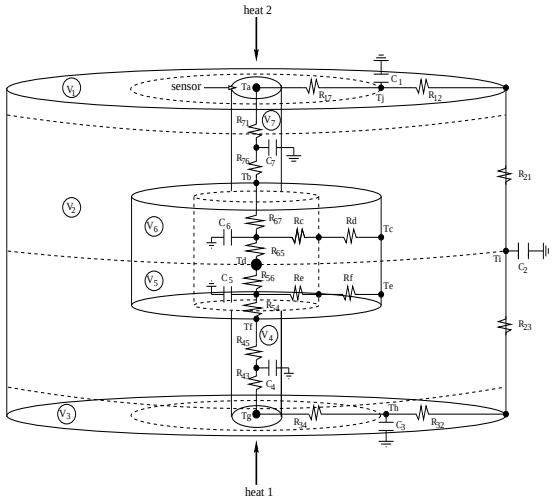


Fig. 5. A simple electrical analog model of the symmetrical cavity with two heating elements (heat1 and heat2).

Figure 6 represents the temperature of the sapphire center test point T_d and the temperature of the sensor T_a during the warm-up period, at ambient temperature 288K. In the same period, the heating power is observed. The temperature reference is fixed at 313K. We observe that the time required to reach the reference temperature is about 1000s for an ambient temperature of 288K. The necessary heating power to maintain the steady state is smaller than 2W.

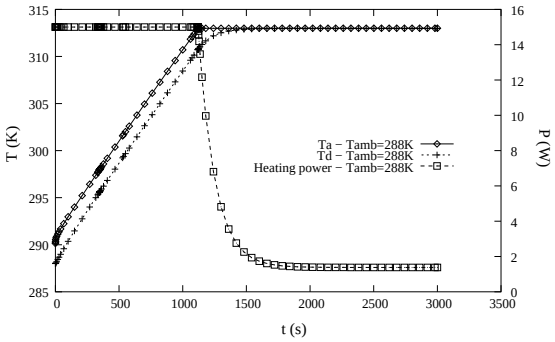


Fig. 6. Temperatures (K) and heating power (W) versus time during the warm-up, when the ambient temperature is +288K.

The thermal gain of the assembly, that is to say its sensitivity to ambient temperature, can be achieved by means of a "DC analysis". Figure 7 shows this result on the sapphire center test point.

We can observe that the thermal gain is divided by 20 if the heating is ensured by a single heating element instead of two ones. Indeed, in the worst case, simulation results give a change ΔT_d=0.6K in the center of the sapphire for an outer temperature change ΔT_{amb}=40K, which leads to a thermal gain ΔT_{amb} / ΔT_d=66. In the other case, with two heating elements, the simulated thermal gain becomes 1212. These results correlate well

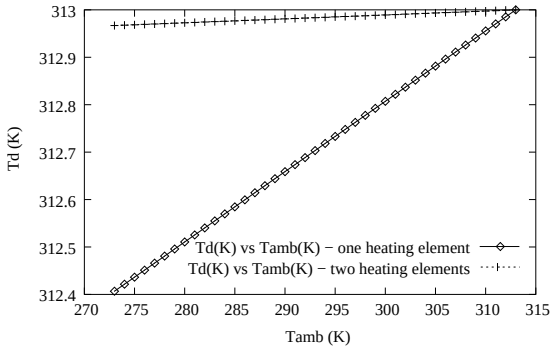


Fig. 7. Simulated thermal gain with one or two heating elements to heat the structure.

with the first model's ones. Regarding these simulation results, we have decided to design an original symmetrical duralumin cavity. The crystal is linked to each lid of the cavity through a duralumin post and a heating element is placed on each extremity of these ones. A hole is foreseen in a post to put a temperature sensor.

5. High-precision temperature regulation

The feedback control system is based on a high-resolution linear Wheatstone bridge, followed by an analog PI correction and a power stage. The temperature of the direct sapphire environment is compared with a reference. It acts on heating elements that are electronically driven into the closed loop.

5.1. Precision thermometer

This is the heart of the electronic regulation. The sensor is a YSI55034 CTN thermistor [10] with high sensitivity of -105Ω/K at 313K. (figure 8).

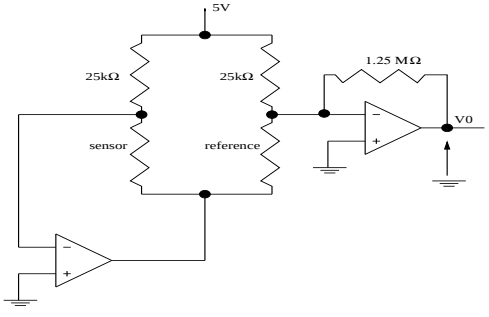


Fig. 8. The thermometer. OP2177 operational amplifiers are used.

V_0 is the error signal. Its sensitivity is around 9.92V/K at 313K.

5.2. PI controller

This stage is a classical parallel PI controller using a Ziegler-Nichols tuning (figure 9).

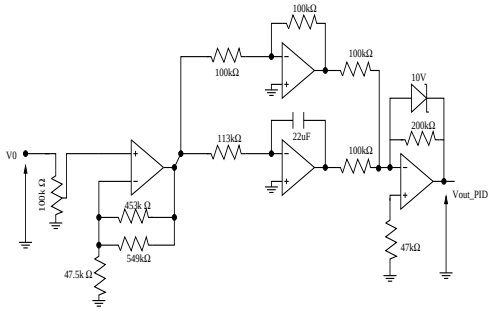


Fig. 9. Ziegler-Nichols tuning and PI controller

5.3. The power stage and control loop

The cavity is heated with two IRF840 power transistors which are parallely monitored. Each one is fixed on one extremity of the duralumin post. The feedback loop ensures the control. The total dissipated power by both transistors is fixed at 15Watts (figure 10).

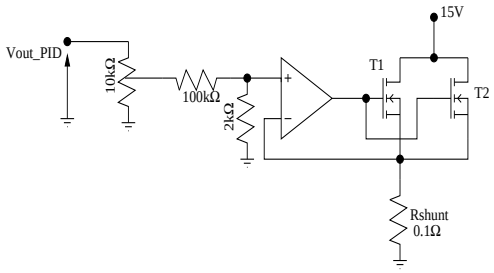


Fig. 10. The power stage.

6. Experimental tests

Our resonator is placed in an original duralumin cavity described in 4. A $WGH_{5,0,0}$ mode (4.866GHz at room temperature) is excited. The high-precision electronic regulation is directly embedded on the cavity. The frequency-temperature dependence of the resonator, shown on figure 11, is measured for 2 different configurations: the structure is heated by a single transistor or two ones.

Experimental results fit pretty well with models (see 4) and show us that the frequency-temperature dependence is strongly reduced when two heating elements are used: the thermal gain is divided by 20 if the heating is ensured by a single transistor instead of two ones. In the best case, measured frequency variations of the resonator for a 273-313K outer temperature range lead to a thermal gain that is around 1360. The temperature coefficient is then reduced at -0.05ppm/K. These results, compared at those obtained in [11], can be considered as satisfying. Afterwards, the evolution of the T_i point (see figure 5) during the warm-up time is measured for a 295K outer ambient temperature. Simulation and experimental results, shown in figure 12, fit well.

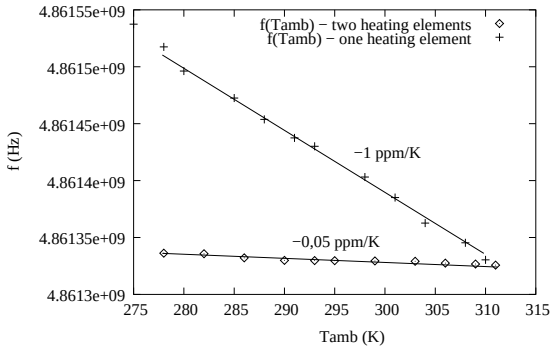


Fig. 11. Resonator frequency versus T_{amb} for the symmetrical (2 heating elements) and the non-symmetrical (1 heating element) confi guration.

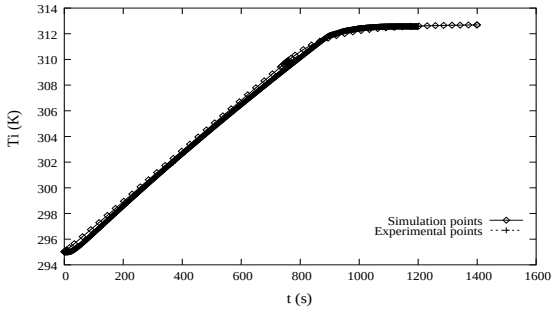


Fig. 12. Temperature (K) of the T_i point versus time (s) during the warm-up at T_{amb}=295K.

The resonator is then implemented in a microwave oscillator. A commercial amplifier compensates losses of the circuit. The corresponding Allan variance, for the temperature stabilized or non-stabilized resonator, is shown in figure 13. In the optimal case, results show us a frequency instability of around 4.10^{-11} for a 1s integration time. For longer integration times, we observe a worst stability because the rest of the circuit wasn't temperature stabilized.

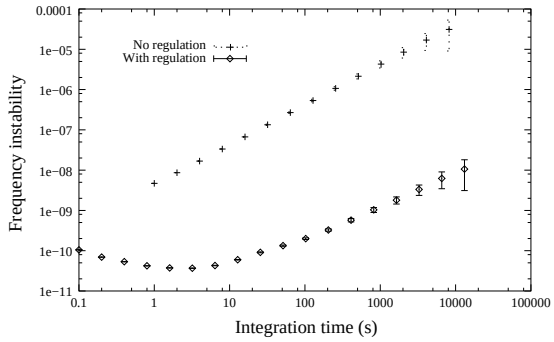


Fig. 13. Allan variance of the temperature stabilized or non-stabilized sapphire resonator based oscillator.

7. Conclusions

A well temperature-stabilized sapphire resonator has been designed at room temperature. Original and simple models have been developped in order to foresee the thermal behavior of the system. The first one has to be reserved to steady state analysis. The behavior of the whole system (cavity + temperature regulation) can be simulated using the electrical model. Principles of this one can be exploited in other kind of configurations.
The necessity to keep the structure as symmetrical as possible has been proved. A high-precision temperature regulation, directly embedded on the cavity, has been developped. The temperature coefficient of the sapphire resonator is reduced at -0.05 ppm/K in the optimal case.

8. Acknowledgment

This research was supported by the Bureau National de Métrologie (BNM) and the Délégation Générale de l'Armement (DGA).

References

[1] G.Cibiel, M.Regis, O.Llopis, Y.Kersalé, V.Giordano, H.Lafontaine, R.Plana, M.Chaubet *Ultra low phase noise SiGe HBT. Application to a C band sapphire resonator oscillator* in *IEEE MTT-S International Microwave Symposium*, pp 691-694, June 2002, Seattle, USA.

[2] E.Ivanov, M.E.Tobar, R.A.Woode *Microwave Interferometry: Application to Precision Measurements and Noise Reduction Techniques* in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol45, n°6, November 1998.

[3] M.E.Tobar, A.J.Giles, S.Edwards, J.H.Searls *High-Q Thermoelectric-Stabilized Sapphire Microwave Resonators for Low-Noise Applications* in *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol41, n°3, May 1994.

[4] Y.Kersale, F.Lardet-Vieudrin, M.Chaubet and V.Giordano *Thermal stabilisation of high-Q sapphire microwave resonator using thermosensitive quartz resonator* in *Electronics Letters*, vol 34, n°18, April 1998.

[5] W.B.Braginsky, V.S.HChenko, K.S.Bagdassarov *Experimental observation of fundamental microwave absorption in high quality dielectric crystal* in *Physics Letters A*, vol 120, n°6, March 1987.

[6] O.Di Monaco, W.Daniau, I.Lajoie, Y.Gruson, M.Chaubet, V.Giordano *Mode selection for a whispering gallery mode resonator* in *Electronics Letters*, vol 32, n°7, 28th March 1996.

[7] V.Giordano, O.di Monaco, Y.Kersale *Résonateur saphir refroidi: Q0=30 millions à 7 GHz* in *Dixièmes Journées Nationales Microondes*, Saint-Malo, pp128-129, 21-22-23 Mai 1997.

[8] E.Ivanov, D.G.Blair and V.I.Kalinichev *Approximate Approach to the Design of Shielded Dielectric Disk Resonators with Whispering-Gallery Modes* in *IEEE Transactions on Microwave Theory and Techniques*, vol41, n°4, April 1993.

[9] S.Galliou *Thermal Behavior of quartz resonators in an oven environment* in *IEEE Transactions on Ultrasonics, FerroElectrics and Frequency Control*, vol 42, n°5, September 1995.

[10] $YSIPrecision^{TM}$ Temperature Group, 2670 Indian Ripple Road, Dayton, Ohio 45440-3605, USA *Datasheet YSI* 55000 *Series GEM Thermistors*

[11] M.E.Tobar, E.N.Ivanov, R.A.Woode, J.H.Searls, A.G.Mann *Low Noise 9-GHz Sapphire Resonator-Oscillator with Thermoelectric Temperature Stabilization at 300 Kelvin* in *IEEE Microwave and Guided Wave Letters*, vol 5, n°4, April 1995.

Session 3D / Oscillators and noise

DEMODULATION OF NOISE VERY CLOSE TO CARRIER AND ESTIMATION OF TOTAL JITTER

Michael J Underhill
University of Surrey and Toric Limited, UK

Abstract

For a symmetrical spectrum at RF at a reasonable distance from carrier the base-band spectrum is the same as the sideband spectrum. For an unsymmetrical RF spectrum this only remains true if the baseband negative and positive frequencies are kept separate and not combined. A single sideband demodulation process is then required. A phase or frequency demodulator processes the RF spectrum into a combined baseband spectrum. If the original RF spectrum is not symmetric the baseband spectrum is a combination of the two sidebands. On the assumption that the AM (arising from the asymmetric sidebands) is removed, the resulting baseband spectrum may then be defined away from the carrier. Very close to carrier the appropriate combination method may depend on the detail of the demodulation process. One purpose of this paper is to examine what the spectrum shape very close to carrier can become after frequency multiplication or division for the case of demodulation of the classic Leeson oscillator spectrum. A high multiplication ratio can cause 'carrier collapse', where a dip or valley appears in the very-close-in noise spectrum.

Once the detail of the noise spectrum has been established it is possible to derive an estimate of total time jitter. The conditions under which an integral down to zero frequency may be used to find the total jitter are examined. A related investigation is what happens in the carrier region when the dominant noise spectrum is 1/f noise.

Introduction

The 2004 EFTF paper "Estimation Of Total Jitter And Jitter Probability Density Function From The Signal Spectrum", [1] showed that:

1. The total jitter power, and the RMS jitter amplitude, can always be calculated from the phase noise spectrum L(f).
2. For sideband spectrum L(f) and time jitter spectrum $\tau(f)$ is $L(f)=2\pi f_o \tau(f)$, where f_o is the carrier frequency.
3. Total jitter can be found directly by integrating the sideband spectrum as it appears on a spectrum analyser.
4. Spur components can be included in total rms jitter if corrected in power level according to the spectrum analyser resolution bandwidth.
5. The phase noise and the total (RMS) jitter are the ultimate determinants of analogue and digital system performance, respectively.
6. In digital and data systems the system Bit Error Rate (BER) determines the final system performance.
7. The clock must sample the bit stream of a data signal correctly within a time window that is a fraction of the bit period.
8. The (raw) Bit Error Rate (BER) is defined by how often this fails to happen.
9. The BER is the proportion of the Probability Density Function (PDF) of the time jitter amplitude that falls outside the prescribed time window.

The questions now to be addressed :-

1. What is the definition of phase noise in the central close-to-carrier region containing 50% of the total signal power?
2. The total signal power can be defined as 'carrier + sideband' power?
3. Assuming a 'Leeson' oscillator spectrum shape for the central region, what is the shape of this region after applying the process of :(a) Frequency Division? (b) Frequency Multiplication?
4. What are the meanings of L(f), $S_\phi(f)$ and $S_\nu(f)$ and modulation index m in this flat central very close to carrier region?
5. Is Differentiation of the baseband signal followed by Phase Modulation the same as Frequency Modulation in the central region?
6. Are these or any processes of Frequency Modulation or Phase Modulation reversible for the central region?

Does 'Carrier Collapse' or 'Coherency Collapse' actually give a central null in the spectrum for typical signals originating from Leeson model oscillators?

Leeson Oscillator Model

D. B. Leeson's Model is a phase noise model of any oscillator [2].

All oscillators have phase noise. In fact the output from an oscillator *can be said to be highly filtered phase noise*. This is one interpretation of D B Leeson's oscillator model.

Fig. 1 shows the Leeson model block diagram. The phase noise half of the thermal noise FkT referred to the input of the amplifier is amplified by positive feedback and filtered until the noise spectrum RMS amplitude reaches the amplifier saturation limit. The limiting action removes the amplitude component of the thermal noise leaving a phase noise component of ½FkT, where F is the noise factor of the amplifier.

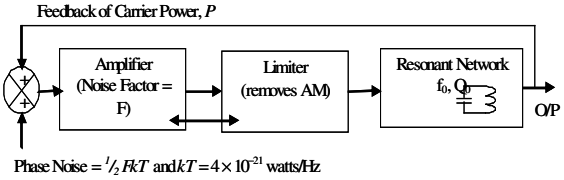


Fig.1 Model of any Oscillator (after D B Leeson)

If the amplification factor is M when saturation power P is reached, the 3dB width of the oscillator spectrum 2α is M times smaller than the 3dB bandwidth 2β of the resonator (when loaded by the rest of the circuitry). Note that by definition

$$2\beta = f_0/Q. \qquad (1)$$

We also have

$$M = \beta/\alpha \qquad (2)$$

The oscillator spectrum is the original phase noise of ½FkT plus the amplified noise which has a $1/f_m^2$ spectrum for sideband frequencies greater f_m than α The total spectrum (from more detailed analysis) is

$$N(f) = ½FkT\,(\beta^2 + f_m^2)/(\alpha^2 + f_m^2) \qquad (3)$$

In the sideband region of interest we have

$$N(f) \approx ½FkT\,\beta^2/f_m^2 \qquad (4)$$

The phase noise L(f) normalised to the total carrier plus sideband power P is

$$L(f) = N(f)/P = (½FkT/P)(\beta^2 + f_m^2)/(\alpha^2 + f_m^2) \qquad (5)$$

In the sideband region we have

$$L(f) = N(f)/P \approx (½FkT/P)\,\beta^2/f_m^2 = (½FkT/P)\,(f_o^2/4\,Q^2 f_m^2) \qquad (6)$$

And this can be simplified to give the usual oscillator phase noise equation:

$$L(f) = N(f)/P \approx (½FkT/4PQ^2)\,(f_o^2/f_m^2) \qquad (7)$$

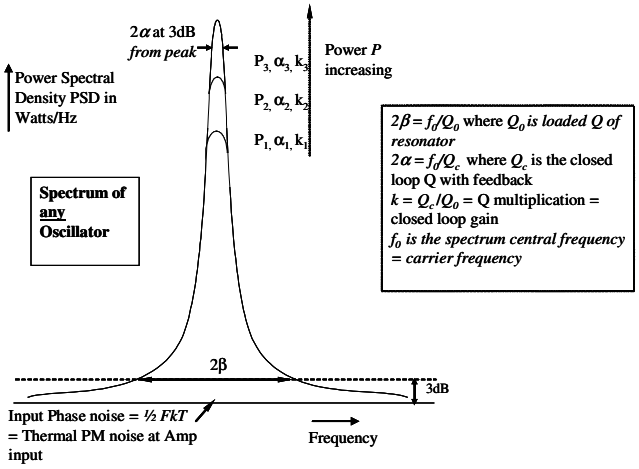


Fig2. Phase Noise Spectrum of any Oscillator

Fig.2 shows the formation of the oscillator spectrum by the positive feedback in an alternative more pictorial way. We have:

1. PSD = Power Spectral Density
2. Plateau Noise phase component = $½FkT$, where F = amplifier noise factor.
3. At $f_m = \pm\beta = \pm f_0/2Q_0$ (where Q_0 = resonator loaded Q) the PSD spectrum lifts by 3dB then rises at 6dB/octave towards the carrier. This directly gives the noise spectrum as in equation (4)
4. $P = P_{av}$ = Output power "available" for feedback to the input. P is the total area under the curve, excluding the $½FkT$ plateau of noise. P determines Q_c the closed loop Q. and height of peak of PSD ≡ Q multiplication). Applying this gives the phase noise equation (7):

$$L(f) = N(f)/P \approx (½FkT/4PQ^2)\,(f_o^2/f_m^2) \qquad (7)$$

Leeson Spectrum and Carrier Bandwidth

The baseband phase noise equation (3) can be plotted on a log-log scale to show the entire noise spectrum as can be seen in Fig. 3.

The SIMetrix SPICE simulator was used to do this by assigning resistor and capacitor values to give a filter with a pole at 16mHz and a zero at 160kHz. 160kHz is the bandwidth of a resonator with a Q = 100 at 16MHz.

The Q multiplication factor in this example is 10^6. A figure of 10^8 is more normal for a typical oscillator. This would give a 'carrier' band width of $2 \times 0.16 =$ 0.32Hz.

Note that it is useful to define 'carrier bandwidth' as the distance between the3dB points on the (oscillator) spectrum. The 'noise sidebands' start at and extend above the 3dB frequencies. This definition recognises that a carrier is actually 'band-limited phase noise'.

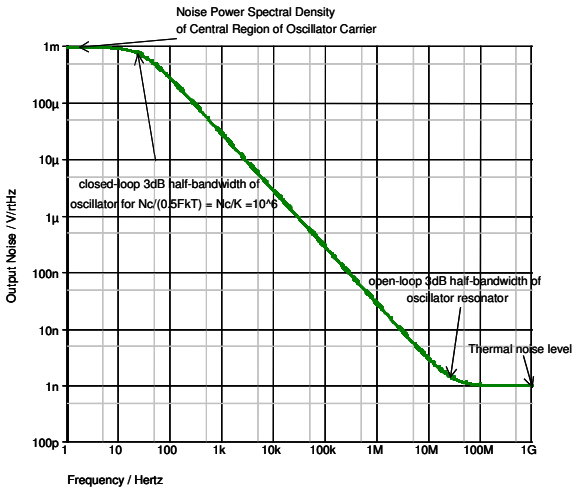


Fig. 3 Baseband Phase Noise Spectrum of Leeson Model of an Oscillator. (The frequency has been scaled up by a factor of 1000.)

'Good' Oscillators

The total power of an oscillator (spectrum) is the power available for feedback to the amplifier input of the oscillator. For a 'good' oscillator the 'dynamic range' at the input before the oscillator 'burns out' or destroys itself should be as high as possible. The oscillator phase noise is reduced by making the oscillator figure of merit PQ^2 as high as possible.

Note that the phase noise improvement obtained by reducing the noise figure to a minimum becomes very small once the noise figure is better than 3dB.

Integration of the Leeson Spectrum to Find the Total Power

For the design of a good oscillator the actual width of the carrier does not have to be known. Nonetheless it is of interest to be able to derive a value for this. The following note shows how this can be achieved.

To find P_{av} = Area under PSD curve = 2 × Area from $f_m = 0$ to ∞, but $|G_{CL}(j\omega)|^2 = \frac{\beta^2}{\alpha^2 + f_m^2}$

therefore $P_{av} = \frac{1}{2}FkT \times 2\int_0^\infty \frac{\beta^2}{\alpha^2 + f_m^2} df_m$. Letting $\tan\theta = \frac{f_m}{\alpha}$ we have

$\int_0^\infty \frac{\beta^2}{\alpha^2 + f_m^2} = \frac{\beta^2}{\alpha}\int_0^{\frac{\pi}{2}} d\theta = \frac{\pi\beta^2}{2\alpha}$ therefore $P_{av} = FkT.\frac{\pi\beta^2}{2\alpha} = \frac{1}{2}FkT.M_Q.2\beta.\frac{\pi}{2}$. Also

$Q_c = \frac{f_0}{2\alpha} = \frac{4P_{av}Q_0^2}{\frac{1}{2}FkT}.\frac{2}{\pi}.\frac{1}{f_0}$ and $M_Q = \frac{Q_c}{Q_0} = \frac{P_{av}}{\frac{1}{2}FkT}.\frac{2}{\pi}.\frac{1}{2\beta}$

Thus the area under the spectrum curve is

$$P_{av} = \tfrac{1}{2}FkT\pi\beta^2/\alpha) \qquad (8)$$

Note that the carrier bandwidth is therefore

$$2\alpha = (\tfrac{1}{2}FkT/4PQ^2)\,(f_o^2\ \pi/2) \qquad (9)$$

This shows that the carrier bandwidth is increased by four times if the frequency is doubled, all else being equal

Frequency Modulation

Frequency modulation by a sine wave produces a series of Bessel Function sidebands at multiples of the modulation frequency. The amplitudes of these sidebands vary with modulation index β as shown in Fig. 4.

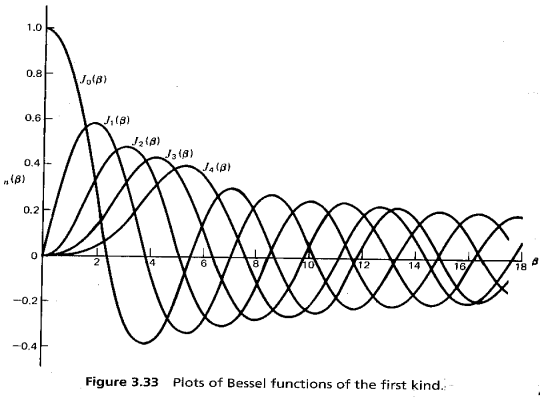


Figure 3.33 Plots of Bessel functions of the first kind.

Fig. 4 Bessel Functions as Amplitudes of Sinusoidal FM Modulation Sidebands (from Haykin [3])

For large modulation indices the energy becomes moved into the sidebands well away from the carrier, which for certain values of modulation index can disappear altogether. This is shown in Fig. 5.

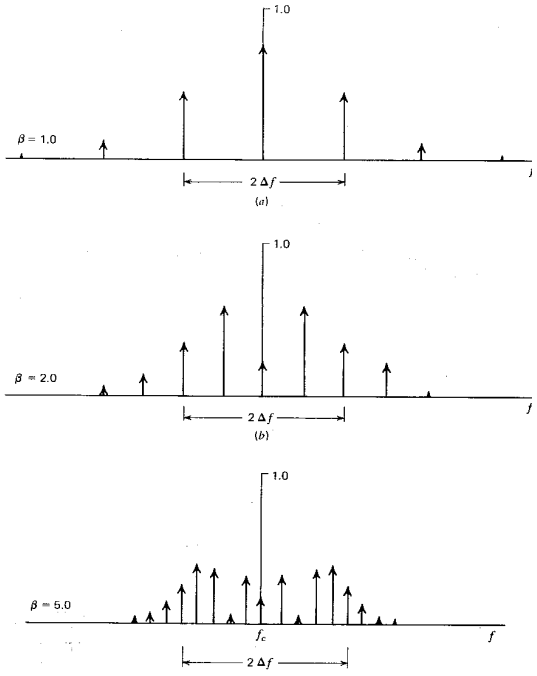


Fig. 5 Bessel FM modulation sidebands for various large modulation indices (from Haykin [3])

Noise can be considered to consist of many independent sine waves at many frequencies and spread to give, when combined, the power spectrum of interest. For white noise the components may be assumed to have similar amplitudes. The amplitude of such a noise signal is known to have a Rayleigh distribution as shown in Fig. 6.

If such noise is used to frequency modulate a carrier the resulting (baseband) spectrum will have a similar Rayleigh shape with a 'hole' at the carrier frequency.

For phase modulation the 'hole' will be filled and the spectrum will have a flat central region.

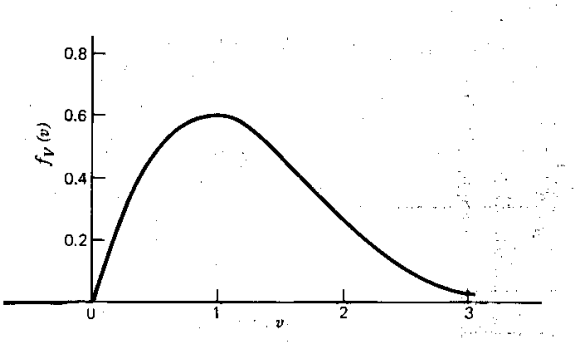


Figure 4.34 Rayleigh distribution.

Fig.6 Rayleigh Distribution of Random Noise with many independent and uncorrelated components (from Haykin [3])

If the modulation spectrum has a dominant sinusoidal component a Rician distribution will apply as shown in Fig. 7.

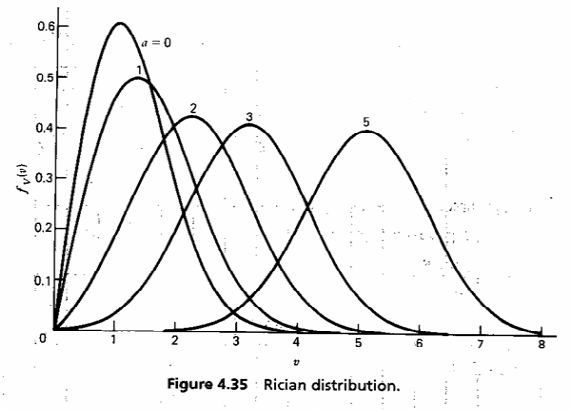


Figure 4.35 Rician distribution.

Fig.7 Rician Distribution of Amplitude for a Carrier Signal with Additive Gaussian Noise for increasing Signal-to-Noise Ratios *a*

As previously shown [1] (in EFTF 2004) a pure sine wave has a PDF as shown in Fig. 8. The FM modulation spectrum can be expected to be similar. The vertical scale is linear for the left hand plot and logarithmic for the right hand plot.

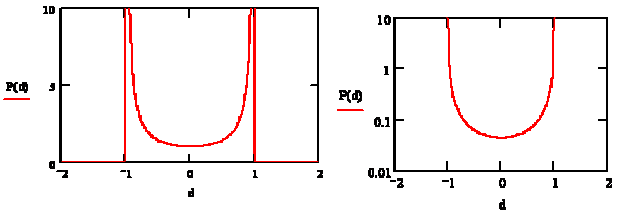


Fig.8 The PDF of a Pure Sine-wave with No Added Noise

If there is some random noise added to the sinusoidal signal we can expect an FM spectrum that looks like Fig. 9.

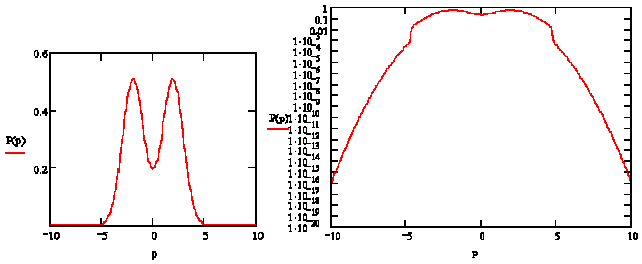


Fig.9 Joint PDF of Sine-wave 'Spur' and 'Random Noise'

All thees spectra have dips in the central region. Wher the dips are particularly pronounced a dip can also appear if phase modulation rather than FM modulation is used.

The Effect of Multiplication or Division on the Phase Noise Spectrum Shape

Fig. 10 shows that dividing the frequency of a source by two reduces the phase noise sidebands by four times but leaves the time jitter unchanged.

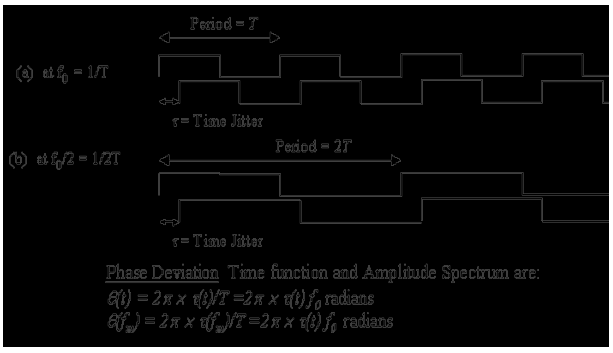


Fig.10 Equivalence of Phase Noise and Time Jitter

Similarly it can be shown that multiplying or dividing by a factor *n* will respectively increase or decrease the

phase noise sidebands by n^2 (and the jitter is left unchanged).

A different argument has to be applied to the central 'flat' region of a Leeson oscillator spectrum. We argue that the 3dB bandwidth of the central region is essentially unchanged by multiplication or division, but its amplitude will change relative to the total power in the sidebands according to the square of the multiplication or division ratio.

In addition the overall spectrum power must remain normalised to unity after multiplication or division has taken place.

Fig. 11 shows the effect of applying the above processes and constraint to multiplication or division by a factor 10 of a Leeson oscillator spectrum. The top leftmost (blue) curve corresponds to a division by 10. The central (green) curve corresponds to the original spectrum. The lower rightmost curve corresponds to a multiplication by 10. The phase noise is given on the vertical axis in dBc/Hz.

Note that the combined effect of the two processes does appear to make the bandwidth of the carrier appear to be a function of the multiplication/division ratio.

Multiplication by 10 does not appear to be enough to cause 'carrier collapse', However it looks as if a ratio of between 100 and 1000 should be enough to initiate carrier collapse.

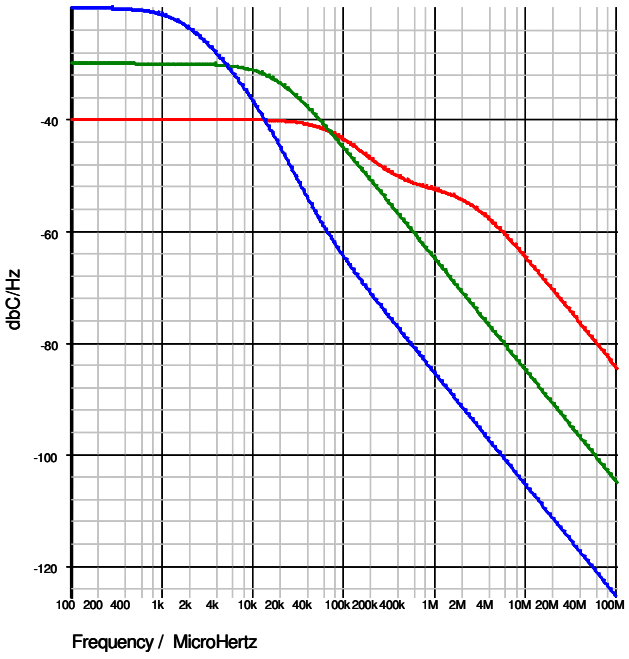


Fig.11 Effect of Frequency Multiplication and Division on a Leeson Oscillator Spectrum

References

[1] M. J. Underhill and P.Brown, "Estimation of Total Jitter and Jitter Probability Density Function from the Signal Spectrum", *Proc. 18th EFTF*, Guildford ,UK, April 2004.

[2] D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Model," Proc. IEEE L., February 1966, pp. 329-330.

[3] S. Haykin. "Communication Systems" Third Edition, 1994 Wiley, ISBN 0-471-30584-7.

REALIZATION OF A 5 AND 10 MHZ HIGH PERFORMANCE ISOLATION DISTRIBUTION AMPLIFIER FOR SHORT TERM FREQUENCY STABILITY MEASUREMENTS OF FREQUENCY SOURCES

P. Salzenstein and F. Lardet-Vieudrin

FEMTO-ST Institute, Department LPMO, CNRS UMR 6174
LNE associate laboratory No 2.13
32 avenue de l'Observatoire, 25044 Besançon Cedex
Phone: +33 (0)3 81 85 39 99, Fax: +33 (0)3 81 85 39 98, e-mail : patrice.salzenstein@lpmo.edu

ABSTRACT

Our Laboratoire National d'Essai (LNE) associate laboratory is accredited to deliver calibration certificates in the short term domain. For spectral density of phase noise and short term frequency stability measurements, it is necessary to have the best reference source signals. However, the noise of the distribution amplifier has to be the weakest possible in order to keep the lowest noise floor. In order to deliver references signals, we achieve a high performances distribution amplifier. In terms of Allan variance, performance of the realized four channels distribution amplifier is respectively better than $\sigma_y(\tau = 1s) = 1{,}4.10^{-14}$ and $7{,}3.10^{-15}$ at 5 and 10 MHz. The intrinsic spectral density of phase noise is then better than $S\varphi(1Hz) = -141{,}4$ dB.rad²/Hz at 5 MHz and –141 dB.rad²/Hz at 10 MHz with a 1/f slope.

KEYWORDS

Distribution amplifier, short term frequency stability, spectral density of phase noise

INTRODUCTION

Any laboratory needs the Allan variance parameter in order to characterize its standards in terms of short term frequency stability for an integration time (τ) typically in the 0.1 s - 100 s range. Of course it is necessary to have the best references.
However, if we use 5 MHz and 10 MHz distribution amplifiers with the lowest phase noise floor, the performances of the reference standard are not limited by the distribution amplifier. For an integration time of 1 s, the frequency stability of the distribution amplifier has to be better than the best available quartz references which present typically 6.10^{-14} frequency stability.

1. DISTRIBUTION AMPLIFIER DESIGN AND REALIZATION

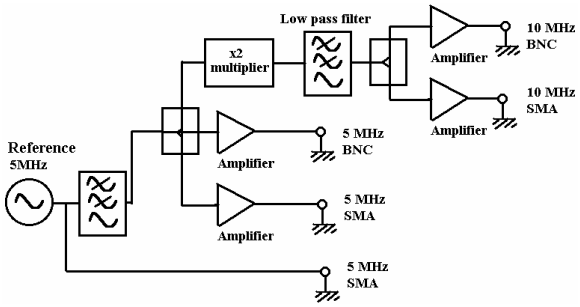


Fig. 1. Electric schematic of the 5 and 10 MHz distribution amplifier

Figure 1 describes the electric schematic of the distribution amplifier. The distribution amplifier performances have been previously evaluated through a *SPICE* simulation. At 5 MHz, the isolation is expected to be equal to –107.9 dBc whereas phase of the signal is much closed to – 180°. We also observe a good output power signal flatness for both 5 and 10 MHz frequencies. Our purpose consists in optimizing the performances of this device in order to integrate it with the best available commercial oscillators. Moreover this distribution amplifier is able to deliver two 5 MHz outputs and two 10 MHz outputs.

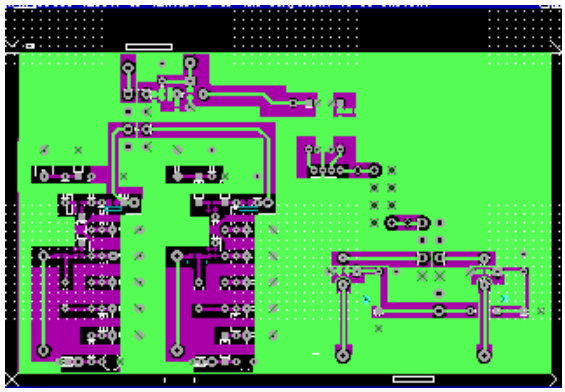


Fig. 2. Lay out of the distribution amplifier

In order to deliver the 10 MHz signals, on one arm, we placed an amplifier and a x2 multiplier based on an hybrid junction. We connected the two 10 MHz outputs after the amplifiers and a power splitter. Between the different output channels, the distribution amplifier to be realized has to reject the signals as good as possible. In the table I, we present the rejection of the harmonic signals. A '4643A' *Hewlett-Packard* synthesizer was connected to the input of the distribution amplifier delivering 7 dBm.

TABLE I
LEVEL OF THE HARMONICS FOR A 7 dBm 5 MHz INPUT SIGNAL

Frequency (MHz)	*OUTPUT (dBc)*
10	-41
20	-56
30	-63
40	-71
50	-91

For an improvement of the performances, we added a filter before the input of the distribution amplifier. We obtained the following results: -41 dBc, -76 dBc, -91 dBc respectively for the 10 MHz, 20 MHz, 30 MHz harmonics at the 5 MHz output.

2. DISTRIBUTION AMPLIFIER PERFORMANCES

Realization of the distribution amplifier allows us to provide several outputs distributed at 5 and 10 MHz. As the reference standard we can possibly use an ultra-stable 5 MHz Quartz Voltage Controlled Crystal Oscillator (VCXO). The distribution amplifier contains x2 multipliers developed in our laboratory. The realized distribution amplifier is shown on the figure 3.

Using a '3048' *Hewlett-Packard* phase noise measurement bench, the spectral density of phase noise characterizing the distribution amplifier has been measured. To achieve such measurement we rejected an *Oscilloquartz OSA 8607* 5 MHz Swiss oscillator known as an ultra-stable source. Typically such quartz oscillators present an Allan variance $\sigma_y(\tau = 1s) = 8.10^{-14}$. As it was mentioned above, the distribution amplifier has to present a better frequency stability performance in order to avoid limitations in the performances for such reference. Applying this experiment, the distributed signal on the 5 MHz and 10 MHz outputs are able to keep the performances of the reference in terms of frequency stability. According the equivalence formula of the frequency Flicker noise, we deduce from the phase noise obtained in the spectral domain, the corresponding Allan variance in the time domain. Then these results have been compared with those obtained by measuring the Allan variance using an time interval analyzer TSC 5110A *Timing Solutions*.

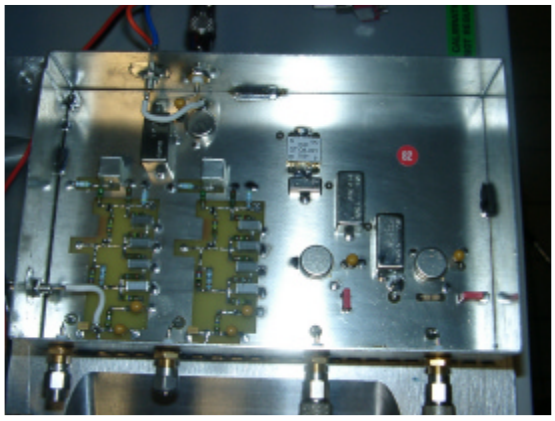


Fig. 3. 5 and 10 MHz distribution amplifier

First of all we measured the spectral density of phase noise of the VCXO on the phase noise measurement bench. We obtained –127 dB.rad²/Hz at 1 Hz from the carrier at 5 MHz. Then we measured the Allan variance on the time interval analyzer as $8.8.10^{-14}$.

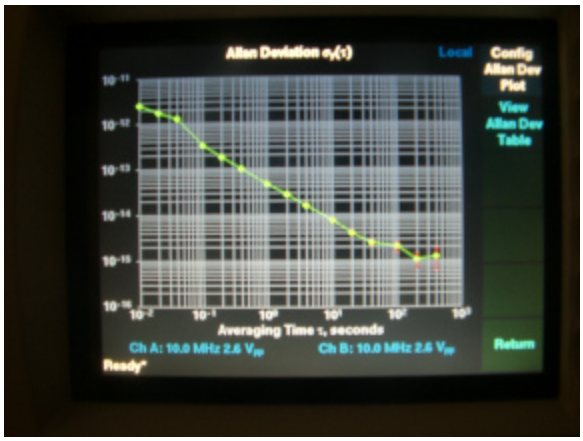


Fig. 4. Distribution amplifier measured Allan variance at 10 MHz on the time interval analyzer

Our results show a rejection of the oscillator noise, through the distribution amplifier. Actually, connecting two outputs of the distribution amplifier on the time interval analyzer, measured Allan variances are respectively equal to $4.8.10^{-14}$ and $2.3.10^{-14}$ at 5 MHz and 10 MHz. Moreover, the obtained Allan variances $\sigma_y(\tau)$ exactly corresponds to the instrument floor given by the manufacturer, *i.e*, 5.10^{-14} and $2.5.10^{-14}$ respectively at 5 MHz and 10

MHz. In our case, the measure of the Allan variance was immediately computed by the system which counts frequencies. However it is limited by the performances of the bench. Phase noise measurement principle is presented on figure 5.

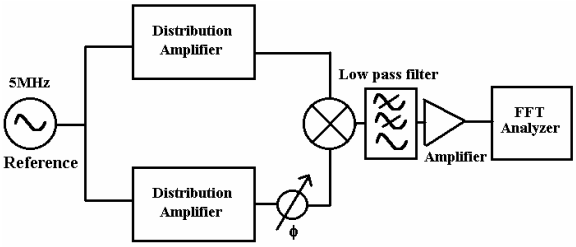


Fig. 5. Phase noise measurement principle

Then we deduced the performances of the distribution amplifier by analyzing the results of the phase noise measurement as presented on the figures 6 and 7. Sφ(1Hz) = – 141.4 dB.rad²/Hz at 5 MHz and –141 dB.rad²/Hz at 10 MHz with a 1/f slope. It was obtained by measuring the phase noise of two similar outputs of the distribution amplifier.

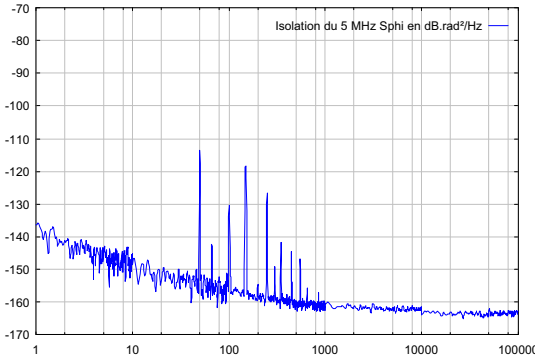


Fig. 6. Phase noise of 5 MHz distribution amplifier measured on the '3048' bench

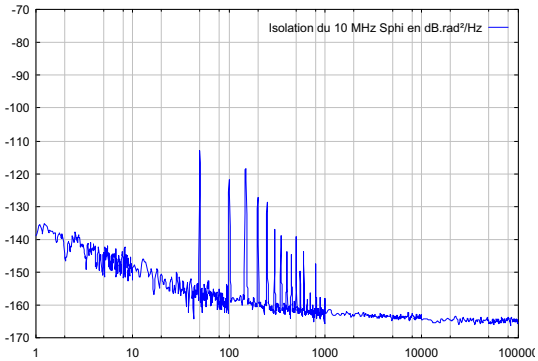


Fig. 7. Phase noise of the 10 MHz distribution amplifier measured on the '3048' bench

We deduced the value considering that the two output have identical contribution to the noise. For a 1/f slope in the spectral domain, we use the following relation [1] that allows to deduce $\sigma_y(\tau)$ versus the spectral density of phase noise :

$$\sigma_y(\tau)^2 = h_1/(4\pi^2.\tau^2) + [9/2 + 3\ \mathrm{Ln}(2\pi F_h\tau) - \mathrm{Ln}2]$$

where h_1 is the coefficient of the slope at 1Hz, τ the integration time, F_h cutoff frequency of the noise.

This equation allows us to estimate the frequency stability in terms of Allan variance, $\sigma_y(\tau) = 1{,}39.10^{-14}$ at 1s for the 5 MHz signal, and $\sigma_y(\tau)=7{,}29.10^{-15}$ at 1s for 10 MHz.

SUMMARY

The distribution amplifier realized in our laboratory presents efficient performances. At 5 MHz and 10 MHz, we obtain better than $\sigma_y(\tau) = 1{,}4.10^{-14}$ and $7{,}3.10^{-15}$ respectively, that corresponds to phase noise floor better than Sφ(1Hz) = –141,4 dB.rad²/Hz at 5 MHz and –141 dB.rad²/Hz at 10 MHz with a 1/f slope. These results are closed to the state-of-the-art NIST results [2].

REFERENCES

[1] C. Audoin, M. Y. Bernard, R. Besson, J. J. Gagnepain, J. Groslambert M. Grandveaud, J.C. Naud, M. Olivier et J. Rutman, La mesure de la fréquence des oscillateurs, CHRONOS, Collection technique et scientifique des télécommunications, MASSON, 1991, pp. 86-89

[2] C. W. Nelson, F. L. Walls, M. Sicarrdi, A. DeMarchi, A new 5 and 10 MHz high isolation distribution amplifier, IEEE Frequency Control Symposium, Boston, USA, 1-3 june 1994, pp. 567-571

DEVELOPMENT OF A PARAMETRIC QUARTZ CRYSTAL OSCILLATOR FOR LOW NOISE APPLICATION

S. GALLIOU*, R. BRENDEL**,
N. RATIER**, M. BRUNIAUX**, P. ABBE*, G. CIBIEL***

FEMTO-ST institute, UMR CNRS 6174,
*Dept. LCEP, 26 chemin de l'Epitaphe, 25030 BESANCON Cedex, France
**Dept. LPMO, 32 rue de l'observatoire, 25044 BESANCON Cedex, France
***CNES
18 avenue Edouard Belin, 31401 TOULOUSE Cedex9, France
Corresponding author: serge.galliou@ens2m.fr

ABSTRACT

Most of the conventional quartz crystal oscillators are using an amplifier to start and sustain the oscillating signal. In this case, the oscillator can be seen as a closed loop including a resonator acting as a passive narrow band filter whose losses are compensated by a non-linear amplifier circuit. Nevertheless, it is well known for a long time that oscillation can also be sustained by using a parametric action. In this case, the losses of the resonant circuit are compensated by a periodic change of an oscillation loop parameter at an appropriate frequency (pump action).
The modulated parameter should be, for instance, the series capacitance of a RLC resonant circuit, and the simplest way to achieve such a modulated capacitance is to use a varactor diode. In fact, this component basically operates as a noiseless reverse-biased junction expected to be a good candidate for a low noise oscillator circuit design.
Although parametric amplifiers are widely used in the microwave domain, they are much less used in the RF circuits and even less for the oscillator circuit design, probably because their development leads to a more involved circuitry and because of the lack of suitable simulation tools.
The paper recall the principle of the parametric oscillator and an example of parametric quartz oscillator (PXO) will be presented. Various methods used to analyse the behaviour of the PXO are described, and comparison between theory, simulation, and experimental results are discussed. Encouraging power spectral density of phase noise fluctuations obtained so far are demonstrated. Simulation tools developed in the frame of this study are currently used to optimise the PXO design.

1. INTRODUCTION

Conventional oscillators convert DC power into signal power, through **resistive** elements. Parametric oscillators make use of a **reactive** element to convert a fraction of an AC power to signal power. This AC power comes from a signal called "pump signal".

Because an ideal reactive element does not exhibit excess noise, parametric oscillators are a potential low noise solution. This is one of the motivations for investigating such a solution for crystal oscillators, and studies on parametric RF oscillators have already been started [1] [2].
Although the idea of parametric operation in oscillator is attractive, its theoretical aspect is nevertheless not so easy and usual simulation tools are quite limited to solve this kind of problem because of its non linear behaviour.

2. FUNDAMENTALS

Let us make reference to the elementary example of the well-known RLC circuit, where a dissipative component, the resistor R, is in series with two reactive components, an inductor L and a capacitor C. When this circuit is simply looped, the current flow i through its elements obeys Equation (1)

$$\frac{d^2 i}{dt^2} + \frac{R}{L}\frac{di}{dt} + \frac{1}{LC} i = 0 \qquad (1)$$

Assuming that the capacitor charge is not zero at time t = 0, Equation (1) leads to a damped sine wave, the damping factor depending on the circuit losses modelled by the resistor R.
To get an increasing sine wave, the second term of Equation (1) must vanish. This can be achieved by generating a negative resistance, whose absolute value is at least equal to R. In this case, oscillations occur in the RLC circuit, even if the capacitor charge is zero at t = 0, (as well as the inductor initial current) because of the inner noise of the circuit. So, the question is: how to generate a negative resistance?
Conventional oscillators are often described from the "two-port" point of view as shown in Figure 1. Then the Barkhausen conditions can be used to extract the gain condition as well as the oscillation frequency of the system. Nevertheless, from the band-pass filter (our previous RLC circuit) the amplifier can also be seen as a negative resistor *Rneg*, according to the relationship given in Fig.1. Then, this new series element suppresses the effect of resistor R.

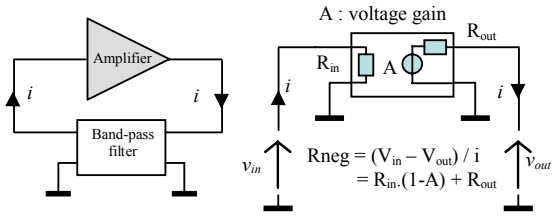


Fig. 1. Left hand: the conventional point of view for describing an oscillator. Right hand: the negative resistor point of view. Rneg is the equivalent negative resistance of the amplifier, seen by the filter.

Parametric operation in a RLC circuit consists in periodically changing one of the reactive elements. Mechanical actions are easy to imagine. The core of an inductor should be moved to modulate the inductance. In the same way, a capacitance should be modulated by changing the distance between both electrodes.

As an example, let us consider a capacitor C modulated by some parameter x (usually called the "pump") according to the simple relationship given by Equation (2)

$$C(x) = C_0.(1+x)\,. \qquad (2)$$

The current i, flowing in this capacitor, is:

$$i = \frac{d}{dt}\left[C(x).u\right], \qquad (3)$$

where u is the voltage across the capacitor terminals. Let us assume that this voltage is a sine wave $u = U_1.\sin(\omega t)$. An interesting and simple case is using a sine pump whose frequency is twice the signal frequency, in this case:

$$x = X_0 + X_2.\sin(2\omega.t + \varphi)\,. \qquad (4)$$

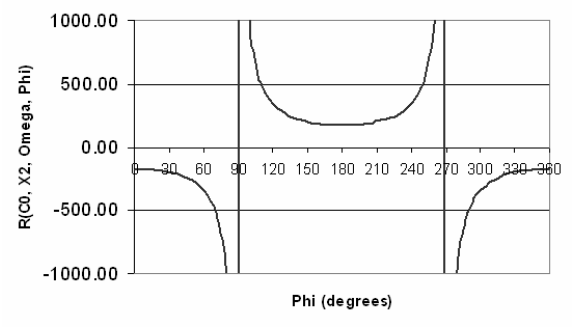
By using Eq. (3) the current i can be written as the sum of three terms:

$$i = i_R + i_C + i_3\,, \qquad (5)$$

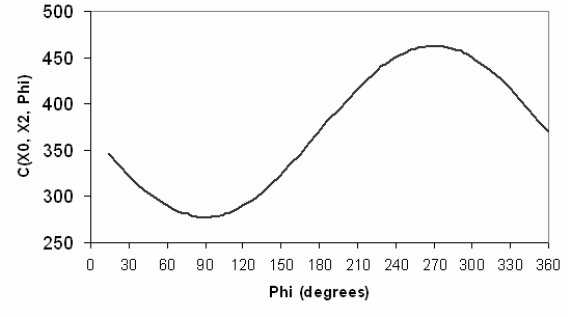
where i_R and i_C can be identified as currents flowing into a resistor and a capacitor, respectively, i_3 being a third harmonic component. Obviously, these equivalent parallel components are function of C_0, X_0, X_2,ω, φ . Fig. 2 (a) and (b) show these equivalent resistance and capacitance versus the phase φ between the signal and its pump, for a set of numerical value of C_0, X_0, X_2 (here, 185 pF, 1, 1), when the signal frequency is 10 MHz.

In Fig. 2, note that the equivalent parallel resistance becomes negative when the phase between pump and signal is within the range [-π/2, π/2].

As a consequence, a capacitor properly pumped according to Eq. (2), is equivalent to a capacitor *Cp* in parallel with a negative resistor *Rpneg*, plus an idler signal. It is often convenient to transform the equivalent parallel-circuit (*Cp*, *Rpneg*) in a series-circuit (*Cs*, *Rsneg*). This is illustrated in Fig. 3, in the case of the series-RLC circuit whose capacitor is pumped. Oscillation occurs if the absolute value of the generated negative-resistance larger than the resistance R of the circuit.



(a)



(b)

Fig. 2. The equivalent parallel resistance expressed in Ohms (a) and the equivalent parallel capacitance expressed in picoFarads (b) of the pumped capacitor *C(x)*, when the pump frequency is twice the frequency of the voltage across *C(x)*. Here, C_0 = 185 pF, $X_0 = X_2 = 1$.

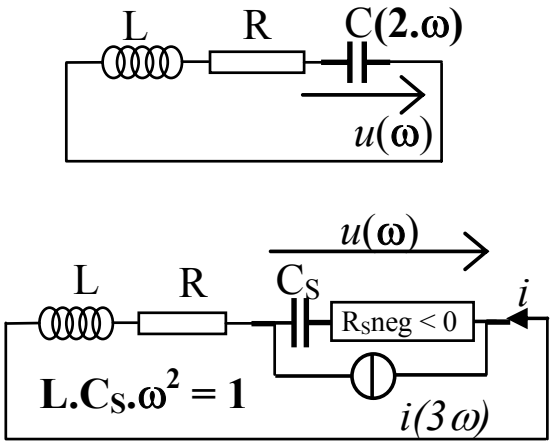
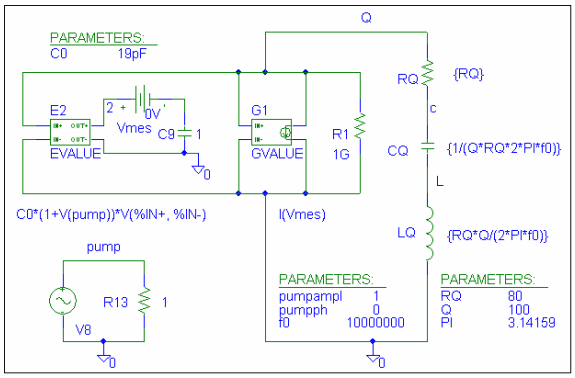


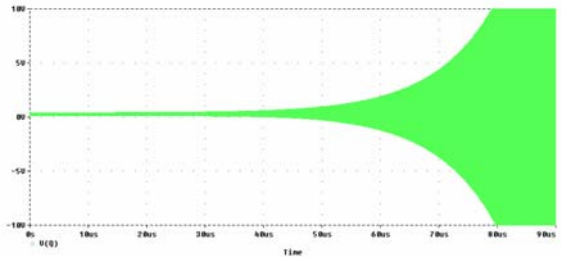
Fig. 3. Illustration of the equivalent circuit of a capacitor pumped at 2ω according to (2). Oscillation occurs at $\omega = 1/\sqrt{LC_s}$, provided that the generated negative-resistance R_{sneg} is larger than R.

We can try to apply the previous parametric effects on a crystal resonator that is equivalent to a series RLC circuit. Obviously, its motional capacitance cannot be modulated (at least, not easily). However, a time-dependent capacitor can be added in series with the resonator. Such a series capacitor will just shift the overall resonant frequency. The larger its value, compared with the crystal capacitance, the smaller the

frequency shift. Fig. 4 is an example of SPICE simulation: here, the controlled capacitor is modelled by a voltage-controlled current source, modulated by the pump voltage. Because only transient analysis can be used to analyse a circuit containing a non-linear capacitance, simulations are quite time- and memory-consuming. Therefore, simulation results shown in Fig. 4 are performed with a crystal quality-factor Q deliberately reduced.



(a)



(b)

Fig. 4. Spice simulation. (a) the analysed network where RQ, CQ and LQ are the motional elements of the crystal resonator. The voltage-controlled capacitor is modelled by {E2, C9, G1}. (b) Voltage at node Q (in volts) versus time (in µs) when the initially charge of CQ is 1V. Oscillation occur.

In the previous example, the pump frequency is twice the signal frequency. Nevertheless, other frequency ratios should be used. Indeed, pumping is still efficient provided that frequencies are related as:

$$\nu = \frac{1}{2}\cdot\frac{\omega_p}{\omega} \quad \text{with} \quad \nu = 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \cdots, \frac{1}{n} \qquad (6)$$

where ω and ω_p are angular frequencies of signal and pump, respectively.

Moreover, oscillation occurs when $Rsneg > R$, that is to say, when $\varphi = 0$:

$$\frac{\Delta C}{C_0'} > \frac{2}{Q}\cdot\frac{1}{\nu^2} \qquad (7)$$

where ΔC denotes the peak to peak swing of the controlled capacitance around its average value C_0'. Q is the quality factor of the circuit.

3. APPLICATION

When attempting to transpose previous ideas to an actual crystal oscillator, the controlled capacitor is implemented with a varactor diode. As a consequence, Eq. (2) is replaced by a more complicated equation:

$$C(u) = C_0 \cdot \frac{1}{(1+\frac{u}{V_J})^M} \qquad (8)$$

for $u < V_J/2$, where u is the voltage across the capacitor terminals. Usually: 0.75 V $\le V_J \le$ 2 V and $0.3 \le M \le 1$, depending on the diode type.

Therefore, the controlled capacitance is no longer independent of the voltage u across it, this voltage being also the controlling parameter x (see Equation (2)).

As a consequence of nonlinearities, assuming that the pump frequency is exactly twice the signal one, u will be something like:

$$u = u_0 + u_1.\sin(\omega t) + u_2.\sin(2\omega t + \varphi_2) + u_3.\sin(3\omega t + \varphi_3) + \ldots \qquad (9)$$

This assumption should be questioned, and it is more rigorous to consider that:

$$u = u_0 + u_s\sin(\omega_s t) + u_p\sin(\omega_p t + \varphi_p) + \sum u_{nm}\sin((n\omega_s + m\omega_p)t + \varphi_{nm})$$
$$\text{with } \omega_p \approx 2.\omega_s \qquad (10)$$

So, taking into account these observations, the problem of generating a negative resistor is not so easy to solve.

First, how to generate a frequency-coherent pump wave? Two solutions can be proposed as illustrated in fig. 5. The first one consists of a pump oscillator phase-locked on the oscillating loop output signal in such a way that $\omega_p = 2\omega$. The second one consists in amplifying the second harmonic of the oscillating loop signal, assuming that it is present.

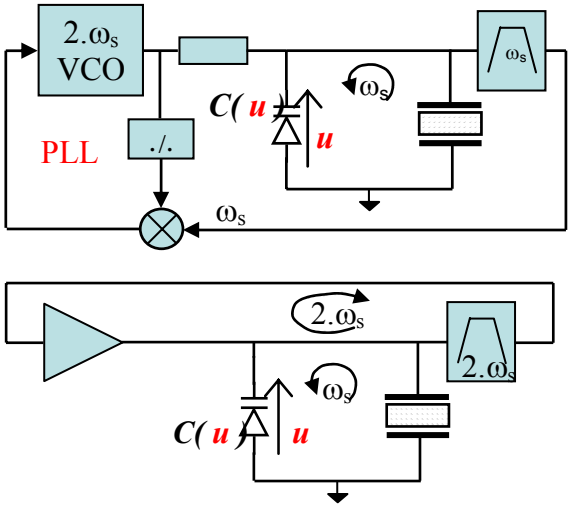


Fig. 5. Two solutions should generate an appropriate signal at twice the signal frequency. The bottom circuit diagram is more closely developed.

The second solution is developed with a 10 MHz quartz crystal resonator. In fact, amplifying the second harmonic of the oscillating loop signal and feeding this

loop with the amplified signal is nothing else than building a second oscillating loop (see Fig. 5).

4. ADJUSTMENTS

At the beginning, it is convenient to analyse the forced-mode behaviour of the oscillating loop. Operating parameters of the varactor diode (pump voltage in relation to signal voltage, phase, varactor type…) are first evaluated with a symbolic mathematics software and next, component values are calculated to generate a negative resistance. Electrical simulations on the overall circuit including the oscillating loop can then be performed. The analysed circuit is shown in Fig. 6. An inductor has been added in the crystal-resonator loop. It is tuned with the varactor diode capacitance, at the resonator frequency (10 MHz) so that the frequency shift induced by the series capacitance is cancelled, and adjustments are easier. It must be pointed out that simulation durations are long because of the large Q factor (time steps are short over a large analysis duration).

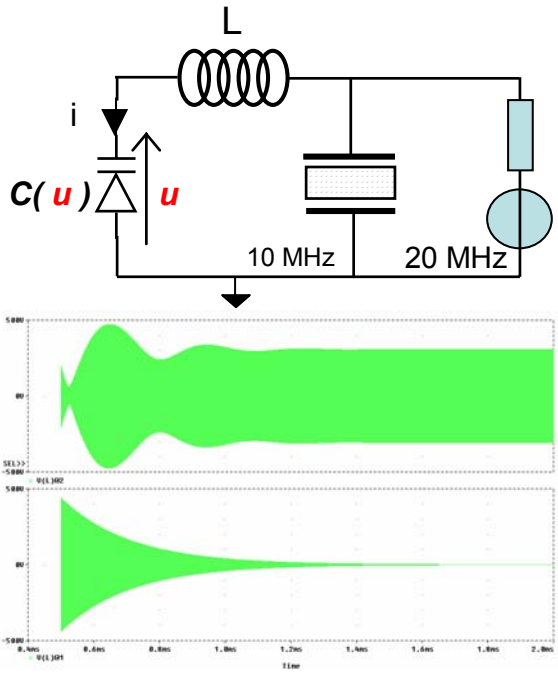
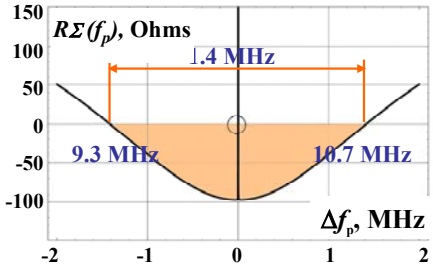
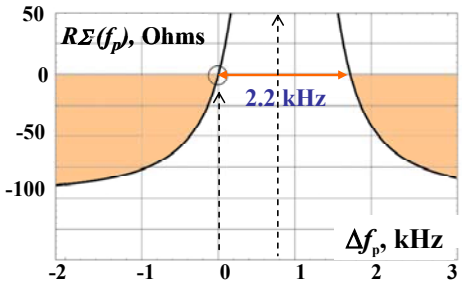


Fig.6. Analysed quartz crystal loop in forced-mode, and simulation results when the motional crystal capacitance is initially charged at 1V. Bottom curve: pump is off. Upper curve: pump amplitude is 4.5V.

Fig 7 shows the forced-mode analytical solution giving the total resistance R_Σ of the loop versus the pump frequency. Here, the Q factor of both elements L and C is about 5. Note that in the case of a C-mode 10MHz SC cut quartz crystal, the temperature sensitive B-mode, located at 10.9 MHz, appearing as a parasitic mode for applications in metrology is automatically rejected without any dedicated filter. This is a remarkable advantage of the PXO, compared with a conventional oscillator.



(a)



(b)

Fig. 7. (a) Overall resistance of the resonant loop versus the pump frequency offset Δf_p (with $f_p = \omega_p/2\pi$) .The axis origin is for the pump frequency twice the signal frequency $f_p = 2f = \omega/\pi$. **(b)** Zoom around the origin at $\Delta f_p = 0$. The maximum positive value of $R_\Sigma(f_p)$ (out of the figure) corresponds to the parallel resonance.

Experiments performed in forced mode are in good agreement with the expected values of pump amplitude, depending on the varactor diode type.

5. CLOSED LOOP SIMULATIONS

Closed loop analysis is the ultimate but also the more critical analysis. Analytical calculation has been performed before implementing the first prototypes. It is based on the principle of harmonic balance and consists in looking for a convergence of the input and the output of the open loop 20 MHz circuit (see Fig. 5).
The varactor-diode voltage found in the previous forced-mode analysis is used as an initial pump voltage, i.e. the input voltage of the 20 MHz open loop. This pump voltage is then filtered and expanded in a non-linear amplifier. The resulting varactor-diode voltage (i.e. the output of the 20 MHz voltage) becomes the new pump voltage of the next step, up to the convergence of both voltages. Experimental results of these first prototypes were in good agreement with theoretical results, as shown in Fig. 8.

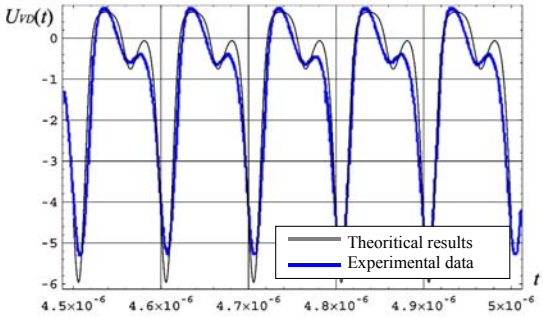


Fig. 8. Experimental and theoretical varactor-diode voltage U_{VD}, versus time. The latter is the result of a step by step convergence of the pump and varactor-diode voltages.

Spice simulations can also complete the closed loop analysis but take a lot of time. An example of analysed circuit is shown in Fig. 9.

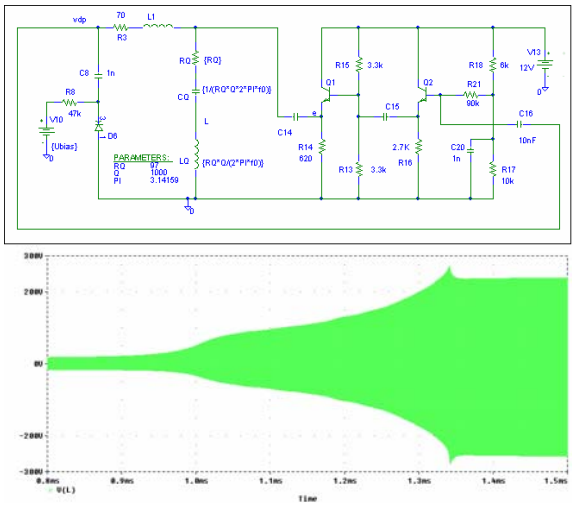


Fig. 9. Schematic of 10 MHz parametric oscillator and transient voltage at node L, i.e. the crystal inductor voltage (Q factor is lowered from 10^6 to 10^3 to save simulation time).

6. RESULTS

A couple of 10 MHz parametric crystal oscillators (PXO) has been built and tested. A record of the power spectral density of phase fluctuations shown in Fig. 10 exhibits a quite good floor noise, because of the high level of crystal excitation. Close to the carrier, the f^{-3} noise is not so good and has to be improved. Note that the loaded Q, given by the f^{-3} and f^{-1} asymptotes intersection is rather low. This could partly explain the poor results close to the carrier.

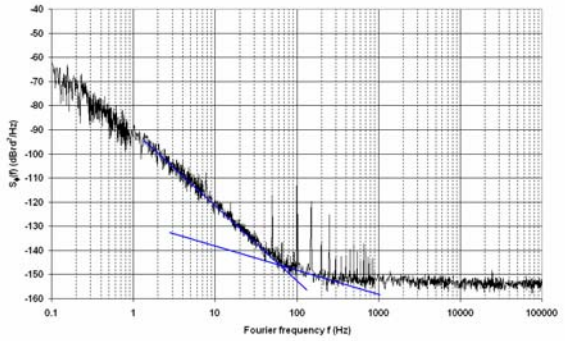


Fig. 10. Power spectral density of phase fluctuations measured on a couple of 10 MHz parametric oscillator. Slopes in f^{-3} and f^{-1} are underlined in the figure.

7. CONCLUSION

Improvements have still to be done on the tested design. Nevertheless, results are very promising, considering that they are measurement made on prototypes of a quite new idea. So, *a priori*, the parametric crystal oscillator (PXO) could be a candidate solution for ultra stable oscillators.

Progress should be made on the simulation aspect. Indeed, the transient analysis is the only type of analysis that can be performed to simulate such a parametric behaviour. Usual simulation tools are not optimized for this operation. Then the development of specific tools [3] becomes a need.

Acknowledgements

This work is supported by the French space agency, CNES (Centre National d'Etudes Spatiales), Toulouse under contract #04/CNES/1859/DCT094.

REFERENCES

[1] V. Komine, S. Galliou, A. Makarov, "Parametric Quartz Oscillators. The Perspectives" *4th International Conference for Young Researchers, Wave Electronics and its Applications in the Information and Telecommunication Systems*. St. Petersburg Russia, May 2001.

[2] V. Komine, S. Galliou, A. Makarov, "A parametric quartz crystal oscillator", *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol.50, no. 12, pp1656-1661, 2003.

[3] N. Ratier, M. Bruniaux, S. Galliou, R. Brendel, J. Delporte, "A very high speed method to simulate quartz crystal oscillator" in this proceeding.

A very high speed method to simulate quartz crystal oscillator

Nicolas Ratier, Mickaël Bruniaux, Serge Galliou, Rémi Brendel
Institut FEMTO–ST, département LPMO, CNRS UMR 6174,
32 avenue de l'Observatoire, F–25044 Besançon, France
nicolas.ratier@lpmo.edu
and Jérôme Delporte
Centre National d'Etudes Spatiales, Toulouse, France

Abstract— **In this paper, we present the SHA method, a Symbolic Harmonic Analysis method to simulate the behaviour of ultrastable quartz crystal oscillators. This nonlinear method is aimed to compute very quickly the steady state as well as amplitude and frequency transients. The ultimate goal is to see instantaneously the influence of a parameter change on the oscillator's features thank to the computation speed. The method proposed here is a mixing of the nonlinear dipolar method previouly developed in our team and the harmonic method. It allows to replace the set of algebro-differential equation of the circuit by a nonlinear system of the Fourier coefficients of the circuit unknowns.**

1. Introduction

We develop a method to analyse the behaviour of quartz crystal oscillators with very high quality factor. The ultimate aim is to perform a real time analysis of this type of oscillators. The method consists in replacing the whole circuit connected to the quartz by a nonlinear dipole whose impedance depends on the current amplitude evaluated at the resonator frequency.

The nonlinear dipolar method [1][2][3] shows that the knowledge of the nonlinear dipole is sufficient to determine the oscillator behaviour. The nonlinear dipolar impedance is computed by replacing the motional part of the quartz by a sinusoidal current source at the resonator frequency. Using an external current source overcomes the drawback of the large resonator time constant thus allowing an efficient harmonic analysis.

The originality of the present approach is the compuation of the dipolar impedance from the circuit equations by using a symbolic calculation method. Our symbolic harmonic analysis replaces the original set of differential equations describing the circuit by a set of nonlinear equations in the Fourier coefficients of the circuit unknows (Node voltages, ...).

Nonlinear dipolar method is first recalled, then, the new method that replaces the dipolar impedance calculation by a symbolic harmonic analysis of the equations is developed. This method is faced with two major difficulties: The processing of the nonlinear components and the management of a very large number of coefficients. The solution proposed to solve these two problems is based on the tree parsing. The method is detailed on a basic oscillator to ease the understanding.

2. Principle of the Nonlinear Dipolar Method

The quartz crystal oscillator is split into two parts: the motional part of the resonator (in short "resonator" by misuse of language) and the amplifier part (including the parallel capacitance C_p), Fig. 1

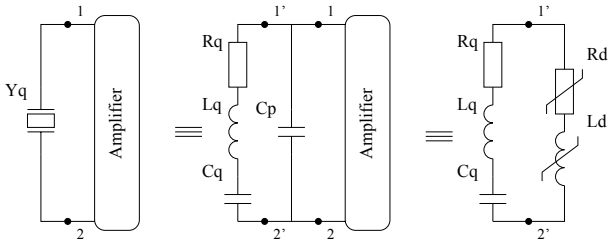


Fig. 1. Dipolar representation

The resonator is regarded as an impedance strongly varying with the frequency and slightly with the amplitude (because of the amplitude–frequency effect) while the amplifier is considered as a nonlinear dipole whose impedance strongly depends on the current amplitude and weakly varying with frequency. These two remarks permit to consider the amplifier part as an nonlinear dipole (R_d, L_d) that depends only on the current amplitude as shown in Fig. 2 right.

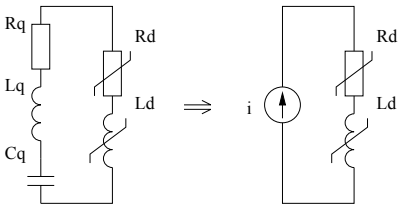


Fig. 2. Amplifier impedance determination

The determination of the equivalent single port network for the amplifier is obtained by replacing the resonator motional branch by a sinusoidal current source of amplitude a and frequency equal to the

resonator's frequency f_q as shown in Fig. 2. Then, a SPICE transient analysis is performed for different current amplitudes. Eventually, for each value of the current amplitude a Fourier transform is calculated to obtain the first harmonic amplitude and phase of the dipolar voltage. The nonlinear dipole R_d et L_d are then obtained as a function of the loop current amplitude a.

The knowledge of $R_d(a)$ and $L_d(a)$ allows us to determine the behaviour of the oscillator. The steady state amplitude a_0 is obtained by solving the implicite nonlinear equation Eq. (1).

$$R_q + R_d(a) = 0. \tag{1}$$

The steady state "frequency" ω_0 of the oscillation is given by Eq. (2).

$$\omega_0^2 = \omega_q^2 \left(1 - \frac{L_d(a_0)}{L_q}\right) \tag{2}$$

The computation time of the curves $R_d(a)$ and $L_d(a)$ is of the order of a few seconds. At each parameter change, for example a change of a resistance value, the computation must be redone. Our aim is to reduce the computation time for the curves $R_d(a)$ and $L_d(a)$ to a few milliseconds allowing to see, in real time, the influence of each circuit component on the behaviour of the oscillator.

3. Principle of the Symbolic Harmonic Method

- Compute the set of ODE describing the circuit in symbolic form, including the nonlinear elements (diode, transistor `NPN`, ...).

$$\begin{aligned} \frac{dV_1}{dt}(t) &= f(V_1(t), \ldots, I_L(t)) && (3)\\ \cdots &= \cdots \\ \frac{dI_L}{dt}(t) &= f(V_1(t), \ldots, I_L(t)) && (4) \end{aligned}$$

- Replace the motional branch of the quartz by a sinusoidal current source.
- Replace all the unknowns $V_1(t)$, ..., $I_L(t)$ by their Fourier series limited to N terms.

$$\begin{aligned} V_1 &= a_0^{V_1} + a_1^{V_1}\cos(\omega t) + b_1^{V_1}\sin(\omega t) + \cdots && (5)\\ \cdots &= \cdots \\ I_L &= a_0^{I_L} + a_1^{I_L}\cos(\omega t) + b_1^{I_L}\sin(\omega t) + \cdots && (6) \end{aligned}$$

- Tranform the set of ODE into a set of nonlinear equations.

$$\begin{aligned} f(a_0^{V_1}, a_1^{V_1}, b_1^{V_1}, \ldots, a_N^{V_1}, b_N^{V_1}) &= 0 && (7)\\ \cdots &= 0 \\ f(a_0^{I_L}, a_1^{I_L}, b_1^{I_L}, \ldots, a_N^{I_L}, b_N^{I_L}) &= 0 && (8) \end{aligned}$$

- Solve numerically this system.
- By using a powerful numerical algorithm, the symbolic algebraic system can be solved very quickly (ms). Because of the quite large Q factor of the resonator, higher order harmonics have no significant contribution on the dipolar current so that only the first order harmonic coefficients a_1 and b_1 are needed to calculate the dipolar impedance.

4. Three main Problems

- How to have a linear complexity? A direct and naive application of the method to build the nonlinear system leads to an exponential growth of the coefficient numbers. That leads to an untractable number of equations, even in the case of basic circuits. So, a method leading to a linear dependence of the coefficient number on the circuit component number has to be found.
- How to manage nonlinear elements? An oscillator is by nature a highly nonlinear circuit. In the case of quartz oscillators with high spectral purity, the nonlinearities usually come from the bipolar transistors. For example, the Ebers-Moll static model transistor (Eq. 9, 10, 11) [4] has an exponential non linearity.

$$\begin{aligned} I_{BE} &= \frac{I_S}{\beta_F}(e^{qV_{BE}/kT} - 1) && (9)\\ I_{BC} &= \frac{I_S}{\beta_R}(e^{qV_{BC}/kT} - 1) && (10)\\ I_{CT} &= I_S(e^{qV_{BE}/kT} - e^{qV_{BC}/kT}) && (11) \end{aligned}$$

- How to solve efficiently the nonlinear system? The resulting system is highly nonlinear with a small density. In the case of a simple oscillator, the classical Newton–Raphson method is sufficient. In the case of an oscillator with a higher number of components, one should develop specific algorithms. Efforts are currently made in this direction.

5. How to achieve a linear complexity?

To ease the understanding, let's demonstrate the method on the simple transconductance oscillator shown in Fig. 3.

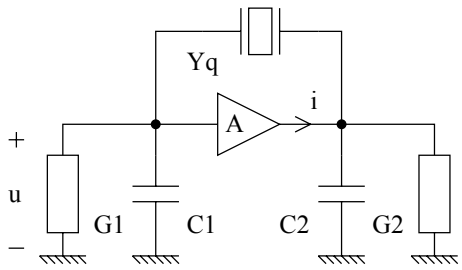


Fig. 3. Transconductance oscillator

The system of equations that models the electronic circuit (Eq. 12 and 13) is build symbolically from a circuit description by using the modifed nodal

method (MNM)[5]. It consists in applying the Kirchoff Current Law (KCL) to each circuit node, and applying the constitutive or branch equations to each circuit element. The programme used is fully home–made and does not call for any external software.

$$-a\sin(\omega t) + G_1 V_1(t) + C_1 \frac{dV_1}{dt}(t) = 0 \quad (12)$$

$$a\sin(\omega t) + G_2 V_2(t) + C_2 \frac{dV_2}{dt}(t) \quad (13)$$
$$+G_1 V_1(t)(1 - \varepsilon V_1(t)^2) = 0$$

The problem is to express each equation in harmonic form, with the constraint that the coefficient number generated during this phase does not increase exponentially.

The key point of the method is to build in a binary tree relative to each equation. Each node represents an algebraic operation. The whole tree can be expressed into an harmonic form as soon as the method to express each leaf and to reduce each node into an harmonic form is known.

The tree will be progressively reduced in harmonic form from the bottom to the top: at each node new coefficients expressed in fonction of the previous one will be generated, so we will have a linear complexity of the generated coefficient. The formal definition of the binary tree is written in OCAML–like syntax [6] below. This kind of definition is generally called "abstract syntax".

```
type equTree =
       SUM  of equTree * equTree
     | PROD of equTree * equTree
     | DIFF of equTree
     | EXP  of equTree
     | COS  of equTree
     | SIN  of equTree
     | UNK  of unknown
     | FCT  of function
```

For example, the binary tree (Fig. 4) represents the first Eq. (12).

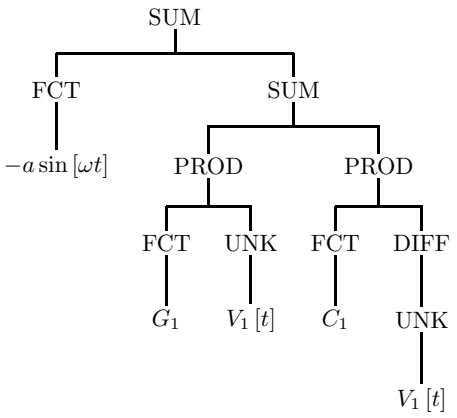


Fig. 4. Binary tree of expression

The representation has many advantages: since abstract syntax has a unique tree structure, it is not necessary to reflect binding strength using different types to avoid ambiguity. Moreover, the abstract syntax shows that the tree is binary, so each node involves only one operation (SUM, PROD, ...).

The abstract syntax processing consists in progressively parse the tree so that, at the final step, it will remain only a Fourier series with manageable coefficient. The processing of the nonlinear functions EXP, COS and SIN is performed in the next section.

- UNK(V1(t)) is replaced by its Fourier series limited to N terms, i.e. it return:

```
A00V1 + A01V1 cos  wt + B01V1 sin  wt
      + A02V1 cos 2wt + B02V1 sin 2wt
      + A03V1 cos 3wt + B03V1 sin 3wt
```

- FCT(G1) is already under the righ form, so it is replaced by itself.
- SUM(S1,S2) with

```
S1 = A00S1 + A01S1 cos wt + B01S1 sin wt
S2 = A00S2 + A01S2 cos wt + B01S2 sin wt
```

is reduced by generating new coefficients

```
TT1 = A00S1 + A00S2
TT2 = A01S1 + A01S2
TT3 = B01S1 + B01S2
... = ...
```

and return

```
TT1 + TT2 cos wt + TT3 sin wt + ...
```

- PROD(S1,S2), DIFF(S1) are reduced like SUM(S1,S2), however, the equations associated with the generated coeffcients are a little bit more complex than the trivial case of SUM(S1,S2).

6. How to manage the nonlinear elements?

The problem is to represent the exponential, cosinus and sinus of a Fourier series into the form of a Fourier expansion. In other words, The problem is to express EXP(S1) with

```
S1 = A00S1 + A01S1 cos(wt) + B01S1 sin(wt)
```

Under the form

```
S2 = EXP(S1)
   = TT1 + TT2 cos(wt) + TT3 sin(wt) + ...
```

The solution of the problem is quite similar to the previous one. It is based on constructing a binary tree and then generating new coeffcient function of the previous one during the tree parsing. In this problem, the concrete grammar to analyse is written in Fig. 5

The definition of the tree is based on the addition formulas. Of course, each transcendental function EXP, COS, SIN will have a different associated tree. The reduction of the tree will be possible thanks to their associated Bessel series.

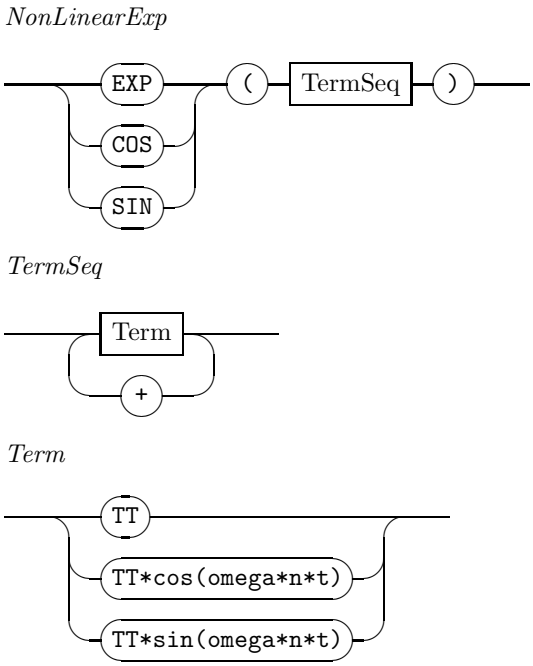


Fig. 5. Concrete grammar

7. Harmonization of EXP function

A recursive application of the addition formula for the exponent function (Eq. 14) allows one to define an abstract grammar that represents a binary tree of EXP.

$$\exp(a+b+c+\cdots) = \exp(a)\exp(b+c+\cdots) \quad (14)$$

For example, the expression

```
EXP(A00S1 + A01S1 COS  wt + B01S1 SIN  wt
          + A02S1 COS 2wt + B02S1 SIN 2wt)
```

is represented by the following binary tree (Fig. 6)

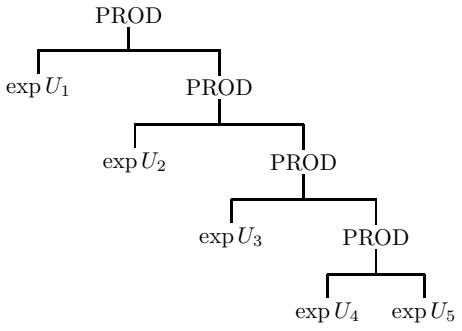


Fig. 6. Binary tree of exponential

with

U_1 = A00S1
U_2 = A01S1 cos(wt)
U_3 = B01S1 sin(wt)
U_4 = A02S1 cos(2wt)
U_5 = B02S1 sin(2wt)

The reduction of the tree is based on the associated Bessel series of the exponential function. Indeed, it is well known that the harmonic expansion of $\exp(z\cos\theta)$ and $\exp(z\sin\theta)$ can be expressed under the form Eq. (15) and Eq. (16) where $I_k(z)$ are the Bessel functions of the second kind and of integer order [7]. This allows us to harmonize the leafs EXP(A01S1 COS(wt)) and EXP(B01S1 SIN(wt)). The terms in EXP(A00S1) are already in the right form.

$$\begin{aligned}\exp(z\cos\theta) &= I_0(z) \\ &+ 2\sum_{k=1}^{\infty} I_k(z)\cos(k\theta)\end{aligned} \quad (15)$$

$$\begin{aligned}\exp(z\sin\theta) &= I_0(z) \\ &+ 2\sum_{k=1}^{\infty}(-)^k I_{2k}(z)\cos(2k\theta) \\ &+ 2\sum_{k=1}^{\infty}(-)^k I_{2k+1}(z)\sin((2k+1)\theta)\end{aligned} \quad (16)$$

8. Harmonization of COS and SIN functions

The method used for EXP can be applied to COS and SIN, it is just a little bit more complicated because the trees for COS and SIN are mutually recursive.

A recursive application of the addition formula for cos (resp. sin), see Eq. (17) and Eq. (18), allows one to define an abstract grammar that represents a binary tree of COS (resp. SIN). Note that the construction of the tree for COS involves the tree of SIN, and conversely.

$$\begin{aligned}\cos(a &+ b+c+\cdots) = \\ &\cos(a)\cos(b+c+\cdots) \\ &- \sin(a)\sin(b+c+\cdots)\end{aligned} \quad (17)$$

$$\begin{aligned}\sin(a &+ b+c+\cdots) = \\ &\sin(a)\cos(b+c+\cdots) \\ &+ \cos(a)\sin(b+c+\cdots)\end{aligned} \quad (18)$$

For example

```
COS(A00S1 + A01S1 cos  wt + B01S1 sin  wt)
```

is represented by the binary tree shown in Fig. 7, with

U_1 = A00S1
U_2 = A01S1 cos(wt)
U_3 = B01S1 sin(wt)

Similary,

```
SIN(A00S1 + A01S1 cos  wt + B01S1 sin  wt)
```

is represented by the binary tree shown in Fig. 8, with

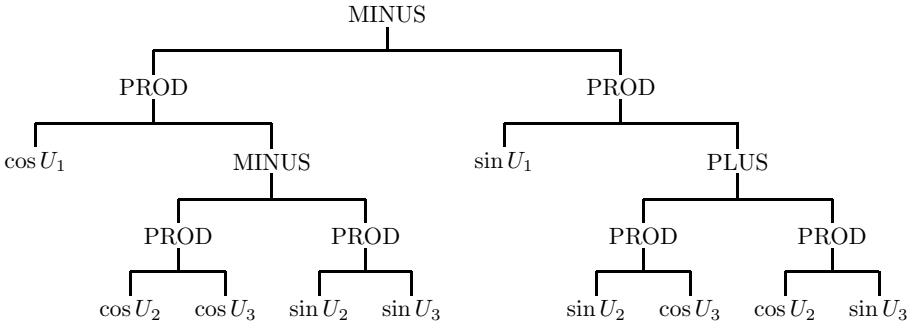


Fig. 7. Binary tree of cosinus

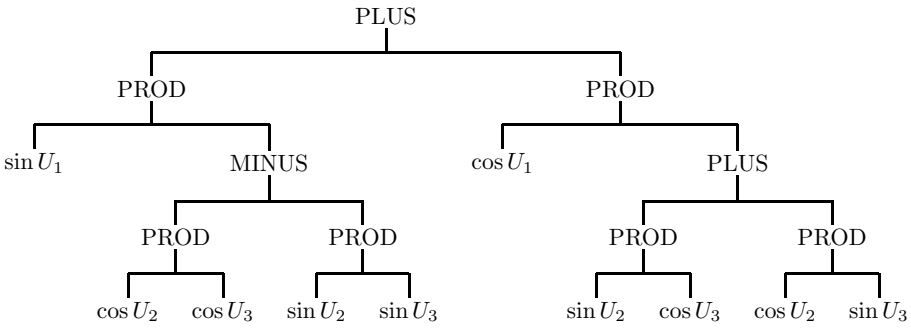


Fig. 8. Binary tree of sinus

$$\begin{aligned} U_1 &= \texttt{A00S1} \\ U_2 &= \texttt{A01S1 cos(wt)} \\ U_3 &= \texttt{B01S1 sin(wt)} \end{aligned}$$

The reduction of the tree is based on the harmonic expansion of $\cos(z\cos\theta)$ and $\cos(z\sin\theta)$. These well known formulas are recalled here for completeness (Eq. (19) and Eq. (20)), where $J_k(z)$ are the Bessel functions of the first kind and of integer order [7]. These equations allow one to reduce the leaf `COS(A01S1 COS(wt))` and `COS(B01S1 SIN(wt))`. The leaf `COS(A00S1)` is already in the right form.

$$\begin{aligned} \cos(z\cos\theta) &= J_0(z) \\ &+ 2\sum_{k=1}^{\infty}(-)^k J_{2k}(z)\cos(2k\theta) \end{aligned} \tag{19}$$

$$\begin{aligned} \cos(z\sin\theta) &= J_0(z) \\ &+ 2\sum_{k=1}^{\infty} J_{2k}(z)\cos(2k\theta) \end{aligned} \tag{20}$$

Similarly, the reduction of the tree is based on the harmonic expansion of $\sin(z\cos\theta)$ and $\sin(z\sin\theta)$. These well known formulas are recalled here for completeness (Eq. (21) and Eq. (22)), where $J_k(z)$ are the Bessel functions of the first kind and of integer order [7]. These equations allow one to reduce the leaf `SIN(A01S1 COS(wt))` and `SIN(B01S1 SIN(wt))`. The leaf `SIN(A00S1)` is already in the right form.

$$\begin{aligned} \sin(z\cos\theta) &= J_0(z) \\ &+ 2\sum_{k=1}^{\infty}(-)^k J_{2k+1}(z)\cos((2k+1)\theta) \end{aligned} \tag{21}$$

$$\begin{aligned} \sin(z\sin\theta) &= J_0(z) \\ &+ 2\sum_{k=1}^{\infty} J_{2k+1}(z)\sin((2k+1)\theta) \end{aligned} \tag{22}$$

9. Harmonization of the other functions

For all the other transcendental functions (log, tan, acos, ...) one proceeds by introducing an additional unknown to the equation system. This transforms the initial differential equation system into an algebro–differential system. This handling is trivial as shown in the following example.

$$\cdots + \tan V_1 + \cdots + acosV_2 + \cdots = 0 \tag{23}$$

becomes

$$\cdots + H_1 + \cdots + H_2 + \cdots = 0 \tag{24}$$

$$H_1\cos V_1 = \sin V_1 \tag{25}$$

$$\cos H_2 = V_2 \tag{26}$$

The latter equations are actually written under the harmonized forms. The same method can be directly applied to the inverse of a function and to the composite functions.

10. Conclusion

In this paper, an extension of the nonlinear dipolar method has been presented to reduce by one order of magnitude the simulation time of ultrastable quartz oscillators. It has been demonstrated that replacing Spice simulations by a symbolic calculation is faced with two difficulties: The processing of the nonlinear components and the management of a large number of coefficients.

A solution to these problems has been found by rewriting the differential equations under the form of a binary tree. Each node of this tree corresponds to an algebraic operation, each leaf is either a parameter of the circuit (R, L, C, I_s, ...) or an unknown of the problem (V_1, V_2, I_l, ...).

This tree is then progressively reduced. At the end of this reduction process, the differential equation is transformed into a nonlinear equation involving only the Fourier coefficients of the unknowns. At each step of a node reduction, a set of equations relating the previous coefficients to the new ones is generated so that the number of equations describing the oscillator increases linearly with the number of components.

References

[1] R. Brendel, D. Gillet, N. Ratier, F. Lardet-Vieudrin, and J. Delporte, "Nonlinear dipolar modelling of quartz crystal oscillators," in *Proc. of 14th EFTF*, 2000.

[2] M. Addouche, N. Ratier, D. Gillet, R. Brendel, F. Lardet-Vieudrin, and J. Delporte, "Adoq: A quartz crystal oscillators simulation software," in *Proc. Int. Freq. Contr. Symp.*, 2001, pp. 753–757.

[3] M. Addouche, R. Brendel, D. Gillet, N. Ratier, F.Lardet-Vieudrin, and J. Delporte, "Modeling of quartz crystal oscillators by using nonlinear dipolar method," *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, vol. 50, no. 5, pp. 487–495, May 2003.

[4] Giuseppe Massobrio and Paolo Antognetti, *Semiconductor Device Modeling with SPICE*, MacGraw–Hill, Inc., second edition edition, 1993.

[5] Jiři Vlach and Kishore Singhal, *Computer methods for circuit analysis and design*, Van Nostrand Reinhold, second edition edition, 1994.

[6] Emmanuel Chailloux, Pascal Manoury, and Bruno Pagano., *Developing Applications With Objective Caml*, Online version (03/2005): `http://caml.inria.fr/pub/docs/oreilly-book/`.

[7] Milton Abramowitz and Irene A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., 1972.

Posters Session 2

NOISE POWER CONVERSION IN OSCILLATORY SYSTEMS: THE METHOD

Yuriy Shmaliy
FIMEE, Guianajuato University, Salamanca, Gto, 36730, Mexico, shmaliy@salamanca.ugto.mx

ABSTRACT

A method is presented to calculate the amplitude and phase power spectral density (PSD) functions (power laws) of an oscillatory system (resonator and oscillator) via the PSDs of its intrinsic noise sources and relevant transformation coefficients. The method description is given for the scalar and vector noises. As an illustration, the noise transformation coefficients are derived and the phase PSD function is studied of a piezoelectric series resonant branch with noisy motional inductance, capacity, and losses.

1. INTRODUCTION

Extensive fundamental studies of noise in oscillatory systems were carried out by Stratonovich [1,2] and Malakhov [3], and then by many other authors. In engineering applications, the most influential has been Leeson's simple but rather heuristic oscillator phase noise model presented in 1966 [4]. Nowadays, we use several mathematical approaches to study oscillatory systems perturbed by noise $v(t)$ depending on its correlation time τ_v and variance σ_v^2, by comparison with the system time constant τ_s and signal power $2P_s$, respectively. Regarding τ_v, τ_s, and the signal-to-noise ratio (SNR) $\gamma = P_s/\sigma_v^2$, the most widely used methods are classified as follows:

Johnson's (thermal) noise ($\tau_v << \tau_s$). If $\gamma >> 1$, then the *statistical linearization* and *harmonic balance* techniques are efficient. If $\gamma \sim 1$, then the *Markov model* is the most accurate.

Excess (flicker) noise of slope f^{-1} ($\tau_v \sim \tau_s$), where f is the Fourier frequency. If $\gamma >> 1$, then the *statistical linearization* and *harmonic balance* techniques are efficient. If $\gamma \sim 1$, then the *Wiener method* (*Volterra series method*) is valid.

Excess (flicker) noise and **random walks** with slope f^{-n}, $1 < n$, ($\tau_v >> \tau_s$). If $\gamma >> 1$, then the *quasi-static method*, the *statistical linearization*, and the *harmonic balance* technique are accurate. If $\gamma \sim 1$, then the *quasi-static method* is the best.

An analysis of several recently published results has shown that, with some modifications, the authors mostly follow either the above listed rigorous theories or Leeson's approximation. It is to be remarked that the *noise conversion method* still remains a relatively young tool for oscillator problems. Notwithstanding this fact, of all the techniques listed above, this is the one that allows the straightforward engineering calculation of the system amplitude and phase power spectral density (PSD) functions via the PSDs of the system intrinsic noises. Furthermore, it clarifies the rules in the PSD shaping, avoiding redundancy of other methods and allowing the derivation of the coefficients of the Leeson approximating polynomial.

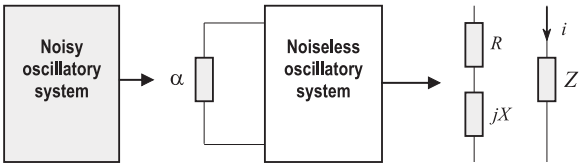


Figure 1: Noisy oscillator system: (a) noisy system, (b) noiseless system with input noise, (c) real and imaginary parts of the system impedance, (d) system noisy impedance with noiseless current

In this paper we give a brief presentation of the *noise conversion method*, which systematic description in given in [5]. We suppose the oscillatory system (e.g. oscillator and resonator) to be noisy, with high quality factor Q, and with large SNR in the signal. Having large SNR, the oscillatory system is linearized and described explicitly (or with some reasonable accuracy) by the complex impedance.

2. THE NOISE CONVERSION METHOD

The essence of the *noise conversion method* may easily be explained by supposing, for the sake of simplicity, that the oscillatory system (Fig. 1a) is the only noise source α that is, for example, in the capacity. We extract α to be the system input (Fig. 1b) (the system is then noiseless) and consider, for example, an imaginary component X (Fig. 1c) of the system impedance Z (Fig. 1d) to be the output.

We suppose that the input has an increment $\delta_\alpha(\mathrm{f})$ that is the root mean square deviation (RMSD) of fractional fluctuations estimated in [6] at an arbitrary Fourier frequency f. This increment, in turn,

induces the output increment $\delta_A(\mathrm{f})$ (that is in X, for example) that basically is positive or negative. In the oscillatory system with high Q, such increments are small, $\sqrt{\delta_\alpha^2} << 1$ and $|\delta_A| << 1$, that allows us to linearize the system and express the output via the input as a linear time invariant operator such as

$$\delta_A(\mathrm{f}) = H_{\alpha A}(\mathrm{f})\delta_\alpha(\mathrm{f}).$$

We then square the left-hand and right-hand sides of this expression, divide them with the bandwidth (BW) of the narrowband fluctuations acting at f (practically [6], the BW is determined by the measurement system), and go to

$$S_A(\mathrm{f}) = H_{\alpha A}^2(\mathrm{f})S_\alpha(\mathrm{f}),$$

where $S_A(\mathrm{f})$ and $S_\alpha(\mathrm{f})$ are the double-sided real PSD functions of the output and input, respectively. The function $H_{\alpha A}^2(\mathrm{f})$ then appears to be the squared real transfer function or the transformation coefficient $K_{\alpha A}^2(\mathrm{f})$ of the input-to-output converter of the PSDs, so that, finally,

$$S_A(\mathrm{f}) = K_{\alpha A}^2(\mathrm{f})S_\alpha(\mathrm{f})\ .$$

For the known (measured [7]) $S_\alpha(\mathrm{f})$, one now needs to determine somehow the coefficient $K_{\alpha A}^2(\mathrm{f})$. Below, we show how to do it employing the *noise conversion method* [5].

2.1. Noise conversion

Aiming to derive generic relations for the system amplitude and phase PSD functions, we replace the system (Fig. 1*a*) by its complex impedance (Fig. 1*d*)

$$Z = R + jX = |Z|\, e^{j\varphi}, \tag{1}$$

$$|Z| = \sqrt{R^2 + X^2}, \tag{2}$$

$$\varphi = \arctan\frac{X}{R}, \tag{3}$$

in which R and X are real and imaginary components of the system impedance, respectively. Thus, the total impedance $|Z|$ (2) and phase φ (3) are associated, respectively, with the system amplitude and phase noises.

It is a typical situation in oscillators and resonators when the acting noise, by nature, is assumed to be Gaussian white and flicker [1–3,6]. Respectively, we assume the noise in (1) to be zero mean stationary Gaussian, additive or colored, low intensity, and with a known PSD $S_\alpha(\mathrm{f})$, or at least with a known part of PSD. We then perform the noisy component α as

$$\alpha = \bar{\alpha}(1 + \delta_\alpha), \tag{4}$$

where $\bar{\alpha}(\mathrm{f})$ is the noiseless value,

$$\delta_\alpha(\mathrm{f}) = \sqrt{\frac{\Delta\alpha^2(\mathrm{f})}{\bar{\alpha}^2(\mathrm{f})}} \geq 0$$

is the noise induced fractional increment, and $\sqrt{\Delta\alpha^2(\mathrm{f})}$ is the RMSD of fluctuations acting in α at f [6].

We now suppose that $\delta_\alpha(\mathrm{f})$ creates small increments, $\delta_r(\mathrm{f})$ and $\delta_x(\mathrm{f})$, in the real and imaginary components, respectively, of the system impedance. To convert δ_α to δ_r and δ_x, we write R and X as in [8], respectively,

$$R = \bar{R}\left(1 + \frac{R_0}{\bar{R}}\delta_r\right), \tag{5}$$

$$X = \bar{X}\left(1 + \frac{R_0}{\bar{X}}\delta_x\right), \tag{6}$$

where R_0 is the system resistance at the resonance frequency (the piezoelectric resonator motional losses with normal drive levels, for example), $\bar{R}$ and $\bar{X}$ are the relevant noiseless values, and

$$\delta_r(\mathrm{f}) = \frac{R[\delta_\alpha(\mathrm{f})] - \bar{R}}{R_0}, \tag{7}$$

$$\delta_x(\mathrm{f}) = \frac{X[\delta_\alpha(\mathrm{f})] - \bar{X}}{R_0} \tag{8}$$

are fractional increments in (5) and (6), respectively. Alternative forms of (7) and (8) are as follows

$$\delta_r(\mathrm{f}) = K_{\alpha r}(\mathrm{f})\delta_\alpha(\mathrm{f}), \tag{9}$$

$$\delta_x(\mathrm{f}) = K_{\alpha x}(\mathrm{f})\delta_\alpha(\mathrm{f}), \tag{10}$$

where $K_{\alpha r}$ and $K_{\alpha x}$ are the transformation coefficients of δ_α-to-δ_r and of δ_α-to-δ_x, respectively. We notice that the linear transformations (9) and (10) make $\delta_r(\mathrm{f})$ and $\delta_x(\mathrm{f})$ to be Gaussian, referring to the Gaussian origin of δ_α.

The next step is to substitute (5) and (6) to (2) and (3) and express either increment in $|Z|$ and φ via the increments of R and X. To be able to do so, we recall that all increments are small, that is $|\delta_{\alpha,r,x,Z}| << 1$ and $|\Delta\varphi| << \pi/2$, and thus the oscillatory system may be linearized. Below we derive the relevant relationships in forms of (9) and (10).

3. AMPLITUDE NOISE

With large SNR, that is $\sqrt{\delta_\alpha^2} << 1$, the signal envelope normalizes [1] and its increments become Gaussian without great loss in accuracy. Using this and acting similarly to (5) and (6), we specify the total impedance of the oscillatory system by

$$|Z| = |\bar{Z}| \left(1 + \frac{R_0}{|\bar{Z}|}\delta_Z\right), \tag{11}$$

where $|\bar{Z}|$ is the relevant noiseless value and $|\delta_Z| << 1$ is the noise induced fractional increment. Exploiting (2), (5), and (6), and neglecting products of small values expresses δ_Z in the following forms

$$\delta_Z(\mathrm{f}) = \frac{|Z[\delta_r(\mathrm{f}), \delta_x(\mathrm{f})]| - |\bar{Z}(\mathrm{f})|}{R_0} \tag{12}$$

$$\cong \frac{\bar{R}\delta_r + \bar{X}\delta_x}{\sqrt{\bar{R}^2 + \bar{X}^2}} = K_{\alpha Z}(\mathrm{f})\delta_\alpha(\mathrm{f}), \tag{13}$$

where, by (9) and (10),

$$K_{\alpha Z} \cong \frac{K_{\alpha r}\bar{R} + K_{\alpha x}\bar{X}}{\sqrt{\bar{R}^2 + \bar{X}^2}} \tag{14}$$

is the transformation coefficient of δ_α-to-δ_Z. We may now square the left-hand side $\delta_Z(\mathrm{f})$ and the right-hand side $K_{\alpha Z}(\mathrm{f})\delta_\alpha(\mathrm{f})$ of (13), divide them with the system BW [6], and go to the amplitude double-sided PSD function

$$S_Z(\mathrm{f}) = K^2_{\alpha Z}(\mathrm{f})S_\alpha(\mathrm{f}) \ , \tag{15}$$

in which the transformation coefficient $K_{\alpha Z}$ is determined by (14) and the spectral density of the acting noise $S_\alpha(\mathrm{f})$ is assumed to be known.

3.1. Multiple noise

For multiple fluctuations, it needs describing the noise induced fractional increment as a vector

$$\delta_\alpha = [\delta_{\alpha 1}\delta_{\alpha 2}...\delta_{\alpha N}]^T \ ,$$

in which $\delta_{\alpha i}$, $i \in [1, N]$, is a generic fractional increment of ith noise source and N is the number of noise sources. We assume that all the transformation coefficients, K_{1Z}, K_{2Z},..., K_{NZ} are determined by (14). That allows one writing the amplitude fluctuations in the linearized oscillatory system as a sum of all the effects:

$$\delta_Z = \mathbf{K}^T_{\alpha Z}\delta_\alpha, \tag{16}$$

where $\mathbf{K}_{\alpha Z} = [K_{1Z}K_{2Z}...K_{NZ}]^T$ is the amplitude noise transformation coefficient matrix. Based upon, we go to the amplitude PSD function

$$S_Z(\mathrm{f}) = \mathbf{K}^T_{\alpha Z}\mathbf{S}_\alpha(\mathrm{f})\mathbf{K}_{\alpha Z}, \tag{17}$$

where

$$\mathbf{S}_\alpha(\mathrm{f}) = \begin{bmatrix} S_{11}(\mathrm{f}) & S_{12}(\mathrm{f}) & ... & S_{1N}(\mathrm{f}) \\ S_{21}(\mathrm{f}) & S_{22}(\mathrm{f}) & ... & S_{2N}(\mathrm{f}) \\ ... & ... & ... & ... \\ S_{N1}(\mathrm{f}) & S_{N2}(\mathrm{f}) & ... & S_{NN}(\mathrm{f}) \end{bmatrix}, \tag{18}$$

is the system noise PSD matrix and $S_{ij}(\mathrm{f})$ is the matrix generic element. If the cross PSD function $S_{ij}(\mathrm{f})$, $i \neq j$ of the intrinsic system noises is unknown then one may use the Cauchy-Schwarz inequality $S_{ij}(\mathrm{f}) \leq \sqrt{S_i(\mathrm{f})S_j(\mathrm{f})}$ and go to the approximation, knowing that, for the uncorrelated noises, (18) takes the diagonal form.

4. PHASE NOISE

With large SNRs, the phase of the oscillatory system normalizes [1] and we thus may specify it by

$$\varphi = \bar{\varphi} + \Delta\varphi, \tag{19}$$

where $\bar{\varphi}(\mathrm{f})$ is a noiseless value at f and the increment $|\Delta\varphi| << \pi/2$ is induced by $\delta_\alpha(\mathrm{f})$.

To convert δ_r and δ_x to $\Delta\varphi$ at an arbitrary f, we first introduce two auxiliary variables

$$x = \frac{X}{R} = \frac{\bar{X}(1 + R_0\delta_x/\bar{X})}{\bar{R}(1 + R_0\delta_r/\bar{R})} \ ,$$

$$y = \frac{\bar{X}}{\bar{R}}$$

and then calculate

$$\Delta\varphi = \arctan x - \arctan y \ . \tag{20}$$

For small noise, there is $xy > -1$ and

$$\Delta\varphi(\mathrm{f}) = \arctan\frac{x - y}{1 + xy}$$

that allows writing

$$\Delta\varphi(\mathrm{f}) \cong R_0\frac{\bar{R}\delta_x - \bar{X}\delta_r}{\bar{R}^2 + \bar{X}^2} = K_{\alpha\varphi}(\mathrm{f})\delta_\alpha(\mathrm{f}) \ , \tag{21}$$

where, by (9) and (10),

$$K_{\alpha\varphi} \cong R_0\frac{\bar{R}K_{\alpha x} - \bar{X}K_{\alpha r}}{\bar{R}^2 + \bar{X}^2} \tag{22}$$

is the transformation coefficient of δ_α-to-$\Delta\varphi$.

The phase PSD is now defined by

$$S_\varphi(\mathrm{f}) = K^2_{\alpha\varphi}(\mathrm{f})S_\alpha(\mathrm{f}) \ , \tag{23}$$

where the transformation coefficient $K_{\alpha\varphi}$ is determined by (22) and the PSD of the acting noise $S_\alpha(\mathrm{f})$ is assumed to be known.

4.1. Multiple noise

Like the amplitude case, the phase PSD, for the multiple noise, calculates by

$$S_\varphi(\mathrm{f}) = \mathbf{K}^T_{\alpha\varphi}\mathbf{S}_\alpha(\mathrm{f})\mathbf{K}_{\alpha\varphi}, \tag{24}$$

where $\mathbf{K}_{\alpha\varphi} = [K_{1\varphi}K_{2\varphi}...K_{N\varphi}]^T$ is the phase noise transformation coefficients matrix, where the coefficients, $K_{1\varphi}$, $K_{2\varphi}$,..., $K_{N\varphi}$, are determined in the same manner as for the one-dimensional case, (22). We notice that, in the one-dimensional case, (24) degenerates to (23).

5. PIEZOELECTRIC SERIES RESONANT BRANCH

It is now in order to consider an example of the piezoelectric resonator. For the sake of simplicity, we neglect here a static capacity C_0, noticing that the common case is discussed in [8].

5.1. System presentation

A series resonant branch of a piezoelectric resonator is modelled with the noisy inductance L_1, capacity C_1, and losses (resistance) R_1, respectively,

$$L_1 = L_{10}(1+\tilde{\delta}_L) \,,$$
$$C_1 = C_{10}(1+\tilde{\delta}_C) \,,$$
$$R_1 = R_{10}(1+\tilde{\delta}_R) \,,$$

where L_{10}, C_{10}, and R_{10} are their noiseless values, and

$$\tilde{\delta}_L(t) = \frac{\Delta L_1(t)}{L_{10}} \,,$$
$$\tilde{\delta}_C(t) = \frac{\Delta C_1(t)}{C_{10}} \,,$$
$$\tilde{\delta}_R(t) = \frac{\Delta R_1(t)}{R_{10}}$$

are their fractional fluctuations, in which $\Delta L_1(t)$, $\Delta C_1(t)$, and $\Delta R_1(t)$ are zero mean stationary random components; that is,

$$\langle \Delta L_1(t) \rangle = \langle \Delta C_1(t) \rangle = \langle \Delta R_1(t) \rangle = 0 \,.$$

We suppose that all fluctuations are measured as in [6,7] and thus their fractional MSDs at `f` are known, namely,

$$\delta_L^2(\mathrm{f}) = \frac{\Delta L_1^2(\mathrm{f})}{L_{10}^2} \,,$$
$$\delta_C^2(\mathrm{f}) = \frac{\Delta C_1^2(\mathrm{f})}{C_{10}^2} \,,$$
$$\delta_R^2(\mathrm{f}) = \frac{\Delta R_1^2(\mathrm{f})}{R_{10}^2} \,.$$

We then specify

$$\omega_0^2 = \frac{1}{L_{10}C_{10}} \,,$$

$$\omega = \omega_0 + \Delta\omega = \Pi_0 \left(Q + \frac{\mathrm{f}}{2\mathrm{f}_r} \right) \,,$$

$$\Pi_0 = \frac{R_{10}}{L_{10}} = \frac{1}{Q^2 R_{10} C_{10}} \,,$$
$$Q = \frac{\omega_0}{\Pi_0} \,,$$

where $\Delta\omega$ is the frequency shift associated with the Fourier frequency f and

$$\mathrm{f}_r = \frac{\Pi_0}{4\pi}$$

is the resonator half 3 dB bandwidth in Hertz.

5.2. Transformation coefficients

In accordance with the method, the real and imaginary parts of the resonator impedance are written as, respectively,

$$R = R_{10}(1+\delta_R), \tag{25}$$

$$X = X_L - X_C = \omega L_{10}(1+\delta_L) - \frac{1}{\omega C_{10}(1+\delta_C)}$$
$$\cong R_{10}\frac{\mathrm{f}}{\mathrm{f}_r} + R_{10}\left(Q + \frac{\mathrm{f}}{2\mathrm{f}_r}\right)\delta_L + R_{10}\left(Q - \frac{\mathrm{f}}{2\mathrm{f}_r}\right)\delta_C. \tag{26}$$

It is seen that the noiseless parts in (25) and (26) are, respectively,

$$\bar{R} = R_{10} \,,$$
$$\bar{X} \cong R_{10}\frac{\mathrm{f}}{\mathrm{f}_r} \,.$$

We now convert each of the increments δ_R, δ_L, and δ_C to the increments of the resonator total impedance (amplitude) and phase. Based upon (9), (10), (25), and (26), we define the relevant increments of real and imaginary components of the system impedance by

$$\delta_{Lr} = \delta_{Cr} = 0 \,,\ \delta_{Rr} = \delta_R \,,\ \delta_{Lx} \cong \left(Q + \frac{\mathrm{f}}{2\mathrm{f}_r}\right)\delta_L \,,$$

$$\delta_{Cx} \cong \left(Q - \frac{\mathrm{f}}{2\mathrm{f}_r}\right)\delta_C \,, \text{and } \delta_{\mathrm{Rx}} = 0 \,,$$

where the double subscript (in δ_{Lr}, for example) means that δ_r is caused by δ_L.

Using (9) and (10), we then determine

$$K_{Lr} = K_{Cr} = 0 \,,\ K_{Rr} = 1 \,,\ K_{Lx} \cong Q + \frac{\mathrm{f}}{2\mathrm{f}_r} \,,$$

$$K_{Cx} \cong Q - \frac{\mathrm{f}}{2\mathrm{f}_r} \,, \text{and } \mathrm{K}_{\mathrm{Rx}} = 0 \,.$$

Finally, employing (14) and (22), we derive the transformation coefficients in question,

$$K_{LZ} \cong K_{CZ} \cong \frac{Q\mathrm{f}}{\sqrt{\mathrm{f}_r^2 + \mathrm{f}^2}}\,, \tag{27}$$

$$K_{RZ} \cong \frac{\mathrm{f}_r}{\sqrt{\mathrm{f}_r^2 + \mathrm{f}^2}}\,, \tag{28}$$

$$K_{L\varphi} \cong K_{C\varphi} \cong \frac{Q\mathrm{f}_r^2}{\mathrm{f}_r^2 + \mathrm{f}^2}\,, \tag{29}$$

$$K_{R\varphi} \cong -\frac{\mathrm{f}_r\mathrm{f}}{\mathrm{f}_r^2 + \mathrm{f}^2}\,, \tag{30}$$

which now allow calculating the amplitude and phase PSDs via (17) and (24), respectively.

5.2. Crystal resonator phase PSD

To determine the resonator phase PSD, we suppose that L_1 and C_1 are purely flicker noisy; that is,

$$S_L(\mathrm{f}) = S_C(\mathrm{f}) = \alpha_X|\mathrm{f}^{-1}|\,,$$

and that R_1 is both thermal and flicker noisy,

$$S_R(\mathrm{f}) = \alpha_R|\mathrm{f}^{-1}| + 2kTP\,,$$

where α_X and α_R are flicker coefficients, k is the Boltzmann constant, T is a temperature in Kelvins, and P is a power dissipated in R_1.

The resonator double-sided phase PSD is then defined, by (24), as

$$S_\varphi(\mathrm{f}) = \frac{2\alpha_X Q^2\mathrm{f}_r^4\,|\mathrm{f}^{-1}|}{\left(\mathrm{f}_r^2+\mathrm{f}^2\right)^2} + \frac{\alpha_R\mathrm{f}_r^2\,|\mathrm{f}|}{\left(\mathrm{f}_r^2+\mathrm{f}^2\right)^2} + \frac{\mathrm{f}_r^2\mathrm{f}^2}{\left(\mathrm{f}_r^2+\mathrm{f}^2\right)^2}\frac{2kT}{P} \tag{31}$$

and we distinguish two limiting cases, namely: (31) may approximately be calculated, within and beyond the half bandwidth, by

$$S_\varphi(\mathrm{f}) \cong 2\alpha_X Q^2\,|\mathrm{f}^{-1}| + \frac{\alpha_R}{\mathrm{f}_r^2}\,|\mathrm{f}| + \frac{2kT}{\mathrm{f}_r^2 P}\mathrm{f}^2, \tag{32}$$

$$S_\varphi(\mathrm{f}) \cong 2\alpha_X Q^2\mathrm{f}_r^4\,|\mathrm{f}^{-5}| + \alpha_R\mathrm{f}_r^2\,|\mathrm{f}^{-3}| + \frac{2kT}{P}\mathrm{f}_r^2\mathrm{f}^{-2}. \tag{33}$$

In (32), two last terms may be neglected [9] and then this approximation reduces to [8, (37)] if the coefficient 4 in [8, (37)] is substituted by 2 (uncorrelated noises) and $\kappa = 0$, having neglected the static capacity. Beyond the half bandwidth, (33), the PSD's slope may vary from f^{-2} (additive thermal noise dominates) through f^{-3} (loss's flicker noise dominates) to f^{-5} (reactance's flicker noise dominates). However, the static capacity of a crystal resonator affects the picture so [8] that the thermal additive noise effect attenuates even more, to the slope f^{-4}.

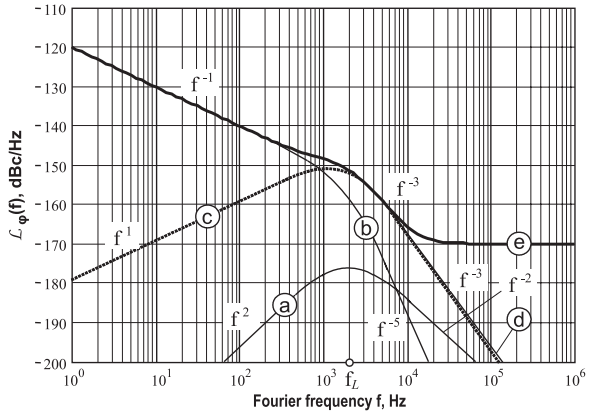


Figure 2: Shaping of the phase PSD function : (a) effect of a resonator additive thermal noise, (b) effect of a motional inductance and capacity flicker noise, (c) effect of a losses flicker noise, (d) total PSD with a noiseless load, (e) total PSD

Here, an increase in the weights of the converted flicker noises forces the slope to range beyond the half bandwidth from f^{-2} to f^{-3}.

6. NUMERICAL EXAMPLE

We now model the phase PSD function of the crystal resonator and compare the result to those measured by Curtis in [10] and simulated in [8] at $f_0 = \omega_0/2\pi = 125$ MHz with $\alpha_X = 5.1\times10^{-22}$ and $\alpha_R = 5\times10^{-12}$. The matched load $R = R_{10}$ in the measurement set doubles the resonator bandwidth, producing $Q_L = Q/2 = 3.13\times10^4$ and $\mathrm{f}_L = 2\mathrm{f}_r = 1997$Hz. The phase noise floor is measured to be -170 dBc/Hz that corresponds to $P = 820\mu W$ with $T = 300$K. To fit the measurement, we operate below with the one-sided PSD and its logarithmic measure [6], respectively,

$$\mathrm{S}(\mathrm{f}) = \frac{1}{2}\left[S(\mathrm{f}) + S(-\mathrm{f})\right], \tag{34}$$

$$L(\mathrm{f}) = 10\log \mathrm{S}(\mathrm{f}). \tag{35}$$

Figure 2 exhibits the shaping of the phase PSD for the examined crystal resonator. It follows that the losses thermal noise makes only an insignificant contribution (a) and may be neglected. The motional inductance and capacity flicker noise (b) appears to be a main shaper within the bandwidth. The losses flicker noise is responsible here for the PSD shape in the transient range (c). Therefore, the phase PSD with the noiseless load looks like (d). If to account for the load thermal noise, then we go to the phase PSD shape (e) that looks like that practically measured. We thus go to the conclusion that,

in the resonator operating at a series resonance frequency, the resonator reactance's flicker noise is a main shaper within the bandwidth, whereas the load thermal noise is a main shaper beyond the bandwidth. In the transient range, all resonator flicker noises may contribute substantially.

5. CONCLUDING REMARKS

In this paper we presented an efficient tool, which we called in [5,11] the *noise conversion method*, for calculating the oscillatory system amplitude and phase PSD functions via the intrinsic system noises with known PSDs. The first significant merit of the approach is a straightforward PSD calculation through the transformation coefficients derived from the system structure. After the latter are obtained in a way similar to that used for the crystal resonator (28)—(31) (or [8]) and oscillator [11], one may merely substitute the system noise PSD into (17) and (24), and calculate the output straightforwardly. The second advantage is that this technique allows the calculation of particular finite ranges of the amplitude and phase PSDs such as those predicted by Leeson in [4] and postulated by the IEEE Standard [6]. Finally, this engineering tool seems to be simplest among all now known.

REFERENCES

[1] R.L. Stratonovich, Topics in the Theory of Ran -dom Noise, vol. I. New York: Gordon & Breach, 1963.

[2] R.L. Stratonovich, Topics in the Theory of Ran -dom Noise, vol. II. New York: Gordon & Breach, 1967.

[3] A.N. Malakhov, Fluctuations in Self-Oscillatory Systems, Nauka, Moscow, 1968.

[4] D.B. Leeson, A simple model of feedback oscillator noise spectrum, Proc. IEEE, vol. 54, no. 2, pp. 329—330, 1966.

[5] Yu.S. Shmaliy, The noise convertion method for oscillatory systems, IEEE trans. on Ultrason., Ferroel. and Freq. Contr., vol. 51, pp. 948-956, Aug 2004.

[6] IEEE SRD 1139-1999, Standard definitions of physical quantities for fundamental frequency and phase time metrology – random instabilities.

[7] H.D. Ascarrunz, A. Zhang, E.S. Ferre-Pikal, and F.L. Walls, PM noise generated by noisy components, in Proceedings of 1998 IEEE Int. Freq. Contr. Symp., 1998, pp. 210-217.

[8] Yu.S. Shmaliy, Conversion of 1/f fluctuations in crystal resonator within an inter resonance gap, IEEE Trans. Ultrason. Ferroelect., and Freq. Contr., vol. 46, pp.61-71, Jan 1999.

[9] G.D. Janushevsky, On fluctuations conversion in passive four-port circuit with crystal resonator, in Proceedings of Frequency Stabil. Conf., Moscow: VIMI, 1986, pp. 246-249.

[10] G.S. Curtis, The relationship between resonator and oscillator noise, and resonator noise measurement techniques, in Proceedings IEEE Annu. Symp. Freq. Contr., 1987, pp. 420-428.

[11] Yu.S. Shmaliy, One-port noise model of a crystal oscillator, IEEE Trans. Ultrason., Ferroelect., and Freq. Contr., vol. 51, pp. 25-33, Jan 2004.

NEW DEVELOPMENT OF PM AND AM NOISE ANALYSIS IN CRYSTAL OSCILLATORS: AN INFLUENCE OF WIDEBAND NOISE SOURCES

V.N.Kuleshov

Moscow Power Engineering Institute (Technical University),
14 Krasnokasarmennaya, Moscow 111250, Russia, KuleshovVN@mpei.ru

ABSTRACT

The Leeson model that is widely used for estimation of PM noise in crystal oscillators (XO) doesn't provide full description of power spectral density (PSD) of output oscillations in the vicinity of the carrier. In the paper of S.Galliou et al. [6] an approach that is aimed to provide more precise calculation of the XO output oscillations PSD and PM noise PSD is developed. The authors paid attention to the oscillation PSD asymmetry with respect to the carrier. The reasons of this asymmetry are discussed.

In this paper we develop a simplified version of the approach offered in by S.Galliou et al. to calculate PSD's of PM and AM noise in XO due to wideband noise sources. Both "linear" and nonlinear models of the XO were analyzed. It is shown that the oscillation PSD asymmetry is closely connected with influence of PM and AM noise and their mutual correlation. The oscillation PSD's at different points of the oscillators are calculated, compared and discussed.

1 INTRODUCTION

At early stages of PM and AM noise analysis in oscillators significant attention was paid to calculation and analysis of power spectral density (PSD) of output oscillations caused by phase and amplitude fluctuations [1,2,3]

As a rule the authors used nonlinear "shortened" equations of the oscillator to calculate PSD's of PM and AM noise and then calculated PSD of the oscillation with fluctuating phase and amplitude. It was found that if there is no correlation between PM and AM noise the PSD of the oscillation is symmetrical with respect to carrier frequency. An influence of correlation between these noises leads to asymmetry of the oscillation PSD [1,2,3].

But the equations that have to be used to calculate PM and AM noise in quarts crystal oscillators (XO) were rather complicated and as a rule they were applicable only to particular circuit. That is one of the reasons why this approach was not widely used in practice. The other reason is connected with the fact that in practice one usually needs the PSD of PM noise. That is why the heuristic Leeson's formula [4] was widely used to estimate PM noise PSD in XO. In the paper of Sauvage [5] the Leesson's formula was obtained using the feedback model of PM transformation in simple oscillator. But this approach doesn't give full picture of the oscillation PSD.

The approach offered in [6] gives more information about the oscillation PSD. But the results presented in [6] need more detailed analysis and clarification of some derivations that follow from this approach.

The aim of this paper is to develop and to do more precise the approach offered in [6] and to discuss the problems of this approach accuracy.

In this paper we analyze only an influence of wideband noise of active device (bipolar junction transistor (BJT)) that is used in the oscillator circuit.

2. THE OSCILLATOR MODEL

To make the following results more clear we shall obtain general formulae basing on particular case of Clapp oscillator [6].

The circuit diagram of the Clapp oscillator is shown in Fig.1.

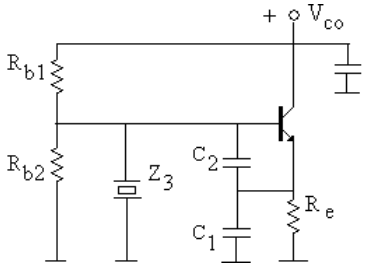


Fig.1. The circuit diagram of the quartz crystal oscillator

To simplify the analysis we use the simplest model of the BJT with sources of wideband noise [3,7]. The high frequency model of the oscillator Fig.1 is shown in

Fig.2. In this model noise sources are presented by two noise current generators.

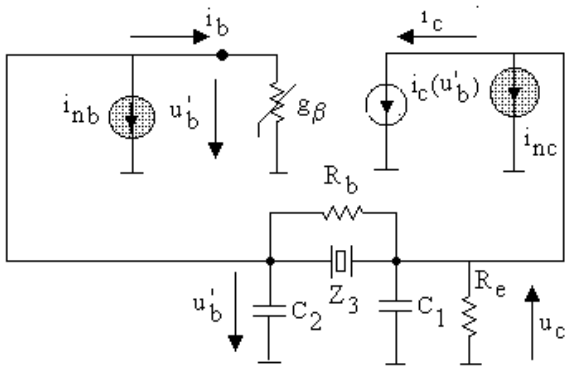


Fig.2. The high frequency model of the oscillator

To calculate a stationary amplitude and frequency of oscillations we use a nonlinear model of the BJT in active mode of it's operation. In this model collector current $i_c(u'_b)$ and base current $i_b(u'_b)$ are defined by formulae [8]:

$$i_c(u'_b) = I_s(e^{u'_b/V_T} - 1), \quad (1)$$

$$i_b(u'_b) = i_c(u'_b)/\beta, \quad (2)$$

where I_s – saturation current, V_T – thermal voltage, β - common emitter configuration current gain.

3. STEADY STATE OPERATION OF THE OSCILLATOR

In the steady state operation of the oscillator

$$u'_b(t) = V_{b0} + V_{b1}\cos[\omega_s t + \varphi_s] \quad (3)$$

where V_{b0} - is the biasing voltage. Complex amplitude of the base voltage

$$\dot{V}_{b1} = V_{b1}e^{j\varphi_s} \quad (4)$$

is connected with complex amplitude of collector current

$$\dot{I}_{c1} = I_{c1}(V_{b1})e^{j\varphi_s} \quad (5)$$

by equation

$$\dot{V}^0_{b1} = Z_m(j\omega_s)\dot{I}_{c1}(V^0_{b1}) - Z_b(j\omega_s)\beta^{-1}\dot{I}_{c1}(V^0_{b1})$$

where $Z_m(j\omega_s)$ – mutual impedance connecting output voltage of the feedback circuit with its input current, $Z_b(j\omega_s)$ – output impedance of the feedback circuit. To simplify the case study we neglect here a contribution of the base current and use the next equation

$$\dot{V}^0_{b1} = Z_m(j\omega_s)\dot{I}_{c1}(V^0_{b1}). \quad (6)$$

From this equation and (4) (5) we obtain

$$X_m(\omega_s) = 0, \quad (7)$$

$$g_{m1}(V^0_{b1})R_m(\omega_s) = 1, \quad (8)$$

where

$$R_m(\omega_s) = \mathrm{Re}\,Z_m(j\omega_s),\ X_m(\omega_s) = \mathrm{Im}\,Z_m(j\omega_s),$$

and

$$g_{m1}(V_{b1}) = I_{c1}(V_{b1})/V_{b1} \quad (9)$$

- an average mutual conductivity of the BJT.

Taking into consideration an influence of the self biasing circuit (R_e, R_{b1}, R_{b2} in Fig.1) and using (1), (2), (3) one can calculate the function $g_{m1}(V_{b1})$ and define the stationary amplitude as a solution of the equation (8). Fig.3 illustrates this solution.

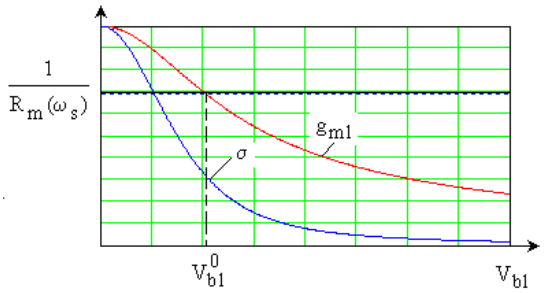


Fig.3. Graphic solution of equation (8)

In the following calculations of PM and AM noise we suppose that the *stationary amplitudes* V^0_{b1}, V^0_{c1} and $I_{c1}(V^0_{b1})$ *are always defined from nonlinear model of the oscillator.*

4. THE OSCILLATION PSD CALCULATION. "LINEAR" MODEL

Let us consider an influence of the additive noise current on this oscillator. Taking into consideration the only noise source $i_{nc}(t)$ we can write an equation (6) as follows

$$Z_m^{-1}(j\omega + j\omega_s)\dot{V}_{b1} = \dot{I}_{c1}(V_{b1}) + \dot{I}_{nc} \quad (10)$$

where $\dot{I}_{nc}$ – complex amplitude of the noise current $i_{nc}(t)$.

The simplest approximate solution of equation (10) can be obtained in the frame of "linear" model [6]. In this model we write the equation (10) as follows

$$[Z_m^{-1}(j\omega + j\omega_s) - g_{m1}(V_{b1})]\dot{V}_{b1} = \dot{I}_{nc} \quad (11)$$

and look for the solution in the next form

$$\dot{V}_{b1} = \dot{V}^0_{b1} + \dot{V}_{n,b}, \quad (12)$$

replacing in equation (11) $g_{m1}(V_{b1})$ by $g_{m1}(V^0_{b1})$. In this case, taking (6) into consideration we obtain from (11)

$$[Z_m^{-1}(j\omega + j\omega_s) - g_{m1}(V^0_{b1})]\dot{V}_{n,b} = \dot{I}_{nc}. \quad (13)$$

Solution of equation (13) gives an explicit expression of Fourier component of the voltage $\dot{V}_{n,b}(j\omega)$ in terms of Fourier component of noise current $\dot{I}_{n,c}(j\omega)$:

$$\dot{V}_{n,b}(j\omega) = \frac{1}{Z_m^{-1}(j\omega + j\omega_s) - g_{m1}(V_{b1}^0)} \dot{I}_{n,c}(j\omega) \quad (14)$$

From (14) we obtain the next formula for PSD of noise component of base voltage

$$S_{vb}(\omega + \omega_s) = \left| \frac{1}{Z_m^{-1}(j\omega + j\omega_s) - g_{m1}(V_{b1}^0)} \right|^2 S_{ic}(\omega + \omega_s) \quad (15)$$

The same way one can obtain from (11) the formula for PSD of noise component of collector voltage $\dot{V}_{n,c}(j\omega)$

$$S_{vc}(\omega + \omega_s) = \left| \frac{Z_m^{-1}(j\omega + j\omega_s) Z_c(j\omega + j\omega_s)}{Z_m^{-1}(j\omega + j\omega_s) - g_{m1}(V_{b1}^0)} \right|^2 S_{ic}(\omega + \omega_s) \quad (16)$$

From (15), (16) we obtain a ratio of noise PSD to the carrier power at the offset ω from the carrier

$$S_b(\omega) = \frac{2S_{vb}(\omega)}{\left[V_{b1}^0\right]^2}; \quad S_c(\omega) = \frac{2S_{vc}(\omega)}{\left[V_{c1}^0\right]^2}. \quad (17)$$

Taking into consideration (6), (7), (8), we can write

$$\dot{V}_{b1}^0 = Z_m(j\omega_s)\dot{I}_{c1}(V_{b1}^0) = R_m(\omega_s)\dot{I}_{c1}(V_{b1}^0), \quad (18)$$

$$\dot{V}_{c1}^0 = Z_c(j\omega_s)\dot{I}_{c1}(V_{b1}^0) = R_c(\omega_s)\dot{I}_{c1}(V_{b1}^0), \quad (19)$$

Using (17), (18), (19), we obtain from (15), (17):

$$S_b(\omega + \omega_s) = \left| \frac{1}{R_m(\omega_s) Z_m^{-1}(j\omega + j\omega_s) - 1} \right|^2 \frac{2S_{ic}(\omega + \omega_s)}{\left[I_{c1}(V_{b1}^0)\right]^2}, \quad (20)$$

$$S_c(\omega + \omega_s) = \left| \frac{R_m(\omega_s) Z_m^{-1}(j\omega + j\omega_s) Z_c(j\omega + j\omega_s) R_c^{-1}(\omega_s)}{R_m(\omega_s) Z_m^{-1}(j\omega + j\omega_s) - 1} \right|^2 \frac{2S_{ic}(\omega + \omega_s)}{\left[I_{c1}(V_{b1}^0)\right]^2}. \quad (21)$$

The formulae (20), (21) are valid both for positive offsets ω from the carrier, and for negative offsets.

5. PM AND AM NOISE. "LINEAR" MODEL

If we present the narrow band stationary noise $\vartheta_{bn}(t)$ in the form

$$\vartheta_{bn}(t) = V_{bII}\cos[\omega_s t + \varphi_s] - V_{b\perp}\sin[\omega_s t + \varphi_s] \quad (22)$$

the sum $\vartheta_{b\Sigma}$ of harmonic oscillation $\vartheta_{b1}(t) = V_{b1}\cos[\omega_s t + \varphi_s]$ and noise $\vartheta_{bn}(t)$ can be written as follows

$$\vartheta_{b\Sigma} = \vartheta_{b1} + \vartheta_{bn} = (V_{b1} + V_{bII})\cos(\omega_s t + \varphi_s) - V_{b\perp}\sin((\omega_s t + \varphi_s) \quad (23)$$

It is known [3,9] that PSD's of $V_{bII}(t)$ and $V_{b\perp}(t)$ can be expressed in terms of PSD $\vartheta_{bn}(t)$ as follows

$$S_{VII}(\Omega) = S_{V\perp}(\Omega) = S_{\vartheta n}(\omega_s - \Omega) + S_{\vartheta n}(\omega_s + \Omega), \quad (24)$$

where Ω - is Fourier frequency that is equal to offset ω from ω_s if ω > 0, and is equal to (-ω) when ω < 0.

The mutual PSD's of $V_{bII}(t)$ and $V_{b\perp}(t)$ look as follows:

$$S_{VII,V\perp}(\Omega) = S^*_{V\perp,VII}(\Omega) = j[S_{\vartheta n}(\omega_s - \Omega) - S_{\vartheta n}(\omega_s + \Omega)]. \quad (25)$$

Assuming that

$$\sqrt{\overline{V_{bII}^2}} = \sqrt{\overline{V_{b\perp}^2}} << V_{b1}, \quad (26)$$

we can present $\vartheta_{b\Sigma}(23)$ in form

$$\vartheta_{b\Sigma} = V_{b1}^0(1 + a_b(t))\cos(\omega_s t + \varphi_s + \phi_b(t)), \quad (27)$$

where

$$a_b(t) = \frac{V_{bII}(t)}{V_{b1}}; \qquad \phi_b(t) = \frac{V_{b\perp}(t)}{V_{b1}} \quad (28)$$

- AM and PM noise of the oscillator.

From (28), (24), (25), (17) we obtain

$$S_{\phi_b}(\Omega) = S_{a_b}(\Omega) = 0{,}5[S_b(\omega_s - \Omega) + S_b(\omega_s + \Omega)], \quad (29)$$

and

$$S_{\phi_b a_b}(\Omega) = -j0{,}5[S_b(\omega_s - \Omega) - S_b(\omega_s + \Omega)]. \quad (30)$$

Thus, from linear model one can calculate PSD's of PM and AM noise, using simple formulae (29), (30).

The main shortcoming of this model is qualitatively incorrect result of AM noise PSD calculation: it goes to infinity when Ω→0. In real oscillators this PSD has a limited value at Ω =0.

To discuss the problem of accuracy of the approach offered in [6], it is necessary to develop more precise approach that takes into consideration nonlinear properties of the oscillator.

6. NONLINEAR MODEL

The main error of the "linear" model is introduced when equation (11) is replaced by equation (13).

To escape this error we return to the equation (10), considering it as a differential equation with respect to slowly varying complex amplitudes [3]:

$$[Z_m^{-1}(p + j\omega_s)\dot{V}_{b1} = \dot{I}_{c1}(V_{b1})] + \dot{I}_{nc} \quad (31)$$

where p=d/dt – symbol of differentiation.

In accordance with (27) we look for the solution in the form

$$\dot{V}_{b1} = \dot{V}_{b1}^0(1 + a_b(t) + j\phi_b(t)). \quad (32)$$

As

$$I_{c1}(V_{b1}) = I_{c1}\left(V_{b1}^0 + V_{b1}^0 a_b(t)\right) = I_{c1}(V_{b1}^0) + \sigma V_{b1}^0 a_b(t)$$

where $\sigma = \left[dI_{c1}/dV_{b1}\right]_{V_{b1}=V_{b1}^0}$, we can write

$$\dot{I}_{c1}(V_{b1}) + \dot{I}_{nc} = \dot{I}_{c1}(V_{b1}^0)\left[1 + \left(\frac{\sigma}{g_{m1}}\right)a_b(t) + j\phi_b(t) + a_n(t) + j\phi_n(t)\right], \quad (33)$$

where

$$a_n(t) = I_{ncII} / I_{c1}(V_{b1}^0); \quad \phi_n(t) = I_{nc\perp} / I_{c1}(V_{b1}^0) \quad (34)$$

- AM and PM noise caused by addition of noise current $\dot{I}_{nc}$ to $\dot{I}_{c1}(V_{b1}^0)$.

Substituting (32), (33) in (31) and taking into consideration (6-9), we obtain the next complex symbolic equation:

$$R_m(\omega_s)Z_m^{-1}(p + j\omega_s)[a_b + j\phi_b] = = \frac{\sigma}{g_{m1}}a_b + a_n + j[\phi_b + \phi_n] \quad (35)$$

To solve equation (35), we present the complex operator $Z_m^{-1}(p + j\omega_s)$ as a sum of real and imaginary parts

$$Z_m^{-1}(p + j\omega_s) = G_m(p, \omega_s) + jB_m(p, \omega_s). \quad (36)$$

Substituting (36) in (35) we obtain two real differential equations that are equivalent to (35):

$$\left[R_m(\omega_s)G_m(p,\omega_s) - \frac{\sigma}{g_{m1}}\right]a_b - R_m(\omega_s)B_m(p,\omega_s)\phi_b = a_n, \quad (37)$$

$$R_m(\omega_s)B_m(p,\omega_s)a_b + \left[R_m(\omega_s)G_m(p,\omega_s) - 1\right]\phi_b = \phi_n, \quad (38)$$

Symbolic solution of these equations looks as follows:

$$a_b = A_{11}(p)a_n + A_{12}(p)\phi_n, \quad (39)$$

$$\phi_b = A_{21}(p)a_n + A_{22}(p)\phi_n, \quad (40)$$

where

$$A_{11}(p) = \frac{\Delta_{11}(p)}{\Delta(p)}; \; A_{12}(p) = \frac{\Delta_{12}(p)}{\Delta(p)}; \; A_{21}(p) = \frac{\Delta_{21}(p)}{\Delta(p)}; \; A_{22}(p) = \frac{\Delta_{22}(p)}{\Delta(p)}; \quad (41)$$

and

$$\Delta(p) = \left[R_m(\omega_s)G_m(p,\omega_s) - \frac{\sigma}{g_{m1}}\right]\left[R_m(\omega_s)G_m(p,\omega_s) - 1\right] - R_m^2(\omega_s)B_m^2(p,\omega_s), \quad (42)$$

$$\Delta_{11}(p) = R_m(\omega_s)G_m(p,\omega_s) - 1; \; \Delta_{12}(p) = R_m(\omega_s)B_m(p,\omega_s); \; \Delta_{21}(p) = -\Delta_{12}(p); \; \Delta_{22}(p) = R_m(\omega_s)G_m(p,\omega_s) - \frac{\sigma}{g_{m1}} \quad (43)$$

Let us suppose that the PSD of $i_{nc}(t)$ is constant in the vicinity of ω_s. Than

$$S_{\phi_n}(\Omega) = S_{a_n}(\Omega) = \frac{2S_{ic}(\omega_s)}{\left[I_{c1}(V_{b1}^0)\right]^2}; \; S_{\phi_n a_n}(\Omega) = 0. \quad (44)$$

Than from (39), (40) we obtain the next formulae for PSD's of AM and PM noise.

$$S_{\phi_b}(\Omega) = \left[\left|A_{11}(j\Omega)\right|^2 + \left|A_{12}(j\Omega)\right|^2\right]\frac{2S_{i_c}(\omega_s)}{I_{c1}^2(V_{b1}^0)}, \quad (45)$$

$$S_{a_b}(\Omega) = \left[\left|A_{21}(j\Omega)\right|^2 + \left|A_{22}(j\Omega)\right|^2\right]\frac{2S_{i_c}(\omega_s)}{I_{c1}^2(V_{b1}^0)}, \quad (46)$$

$$S_{\phi_b a_b}(\Omega) = \left[A_{11}(-j\Omega)A_{21}(j\Omega) + A_{12}(-j\Omega)A_{22}(j\Omega)\right]\cdot\frac{2S_{i_c}(\omega_s)}{I_{c1}^2(V_{b1}^0)}. \quad (47)$$

Thus, in nonlinear model one has to find PM and AM noise PSD's first. After that one can calculate the PSD of the oscillation (27) in the vicinity of ω_s, using the next formula [3]:

$$S_b(\omega_s + \omega) = 0{,}5[S_{\phi_b}(\omega) + S_{a_b}(\omega) - 2\,\mathrm{Im}S_{\phi_b a_b}(\omega)]. \quad (48)$$

In this formula $S_{\phi_b}(\omega)$ and $S_{a_b}(\omega)$ have to be considered as even functions of offset from he carrier ω, and $\mathrm{Im}S_{\phi_b a_b}(\omega)$ is an odd function of ω. From (48) it follows that the PSD $S_b(\omega_s + \omega)$ is symmetrical if $\mathrm{Im}S_{\phi_b a_b}(\omega)$.

It is clear that calculations based on general nonlinear model, providing correct finding both PM and AM noise are not so simple as calculations based on linear model.

7. COMBINED MODEL

In this approach we suppose that nonlinear model is used to find PSD of PM and AM noise at small Fourier frequencies $\Omega < \Omega_{CM}$ (and small offsets from carrier $|\omega| < \Omega_{CM}$). We suppose also that at small offsets $|B_m(j\Omega, \omega_s)| << |G_m(j\Omega, \omega_s|$ and it is possible to simplify equations (37), (38), assuming that $B_m(p, \omega_s) = 0$:

$$\left[R_m(\omega_s)G_m(p,\omega_s) - \frac{\sigma}{g_{m1}}\right]a_b = a_n, \quad (49)$$

$$\left[R_m(\omega_s)G_m(p,\omega_s) - 1\right]\phi_b = \phi_n. \quad (50)$$

Using these equations and (48) we obtain

$$S_{\phi_b}(\Omega) = \frac{1}{\left|R_m(\omega_s)G_m(j\Omega),\omega_s) - 1\right|^2}\frac{2S_{i_c}(\omega_s)}{I_{c1}^2(V_{b1}^0)}, \quad (51)$$

$$S_{a_b}(\Omega) = \frac{1}{\left|R_m(\omega_s)G_m(j\Omega),\omega_s) - \frac{\sigma}{g_{m1}}\right|^2}\frac{2S_{i_c}(\omega_s)}{I_{c1}^2(V_{b1}^0)}, \quad (52)$$

$$S_{\phi_b a_b}(\Omega) = 0, \quad (53)$$

and from (48) we have

$$S_b(\omega_s+\omega)=0{,}5[S_{\phi_b}(\omega)+S_{a_b}(\omega)]. \qquad (54)$$

As a rule in the region of small offsets from the carrier

$$\left|R_m(\omega_s)G_m(j\omega,\omega_s)\right|\approx\sqrt{1+(\omega/\Delta\omega_x)^2}\ , \qquad (55)$$

where $\Delta\omega_x$ is the value of the order of half bandwidth of crystal resonator, and $\sigma/g_{m1}\approx 0{,}3\div 0{,}5$. That is why at the offsets

$$\Delta\omega_x<|\omega|<3\Delta\omega_x \qquad (56)$$

results of calculation $S_b(\omega_s+\omega)$ using (54) and (20) are close. When $|\omega|>3\Delta\omega_x$ an influence of $B_m(p,\omega_s)$ in the sum (36) can not be neglected, but one can neglect the difference between σ/g_{m1} and 1, as $\left|R_m(\omega_s)Z_m^{-1}(j\omega+\omega_s)\right|>>1$. That is why the ratio of noise power to carrier power can be calculated using (20), and PSD of PM and AM noise can be calculated from (20), (29), (30).

8. PM AND AM NOISE OF CLAPP OSCILLATOR

Let us consider the oscillator shown in Fig.1. Circuit model of the quartz crystal resonator is shown in Fig.4.

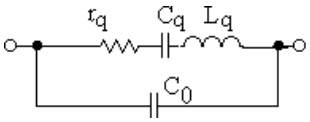


Fig.4. Circuit model of the crystal resonator

In the case, studied here, we suppose: $f_q=\omega_q/2\pi=10\text{MHz}$, r_q = 100 Ohm, $Q_q=10^6$, C_0 = 1 pF. The steady state operation mode is chosen so that the power, dissipated on the crystal, is equal $P_q=10\cdot10^{-6}\,\text{W}$, and oscillation frequency f_S is rather close to f_q $(f_s-f_q)/\Delta f_q=2{,}4$, where $\Delta f_q=f_q/2Q_q$ - half bandwidth of the resonator. Assuming $C_1=C_2$, we obtained $C_1=C_2=136\,\text{pF}$; $V_{b1}^0=0{,}051\text{V}$, $R_m(\omega_s)=110\,\text{Ohm}$. To provide the normalized amplitude $V_{b1}^0/V_T=2{,}06$ in accordance with Fig.3 we have to take $g_{m1}(V_{b1}^0)/g_m=(R_m(\omega_s)g_m)^{-1}=0{,}69$, and $g_m=(0{,}69R_m(\omega_s))^{-1}=0{,}013\,\text{A/V}$. From the BJT model (1) $g_m=I_{c0}/V_T$, and we have to use collector current I_{c0} = 0,33 mA, and I_{c1} = 0,47 mA. From Fig.3 one can see that at this operation mode

$$(\sigma/g_{m1})=0{,}45. \qquad (57)$$

To calculate noise to signal ratio at the input of the BJT one has to find $Z_m^{-1}(j\omega+j\omega_s)$. For the circuit Fig.1 symbolic mutual impedance looks as follows

$$Z_m^{-1}(p+j\omega_s)=R_{m0}^{-1}(\omega_s)\left[-j\xi_0+\bar{Z}_q(p+j\Delta\omega_s)\right], \qquad (58)$$

where $R_{m0}^{-1}(\omega_s)=\omega_q^2C_1C_2r_q$, $\xi_0=(C_1+C_2)/\omega_qC_1C_2r_q$,

$$\bar{Z}_q(p+j(\omega_s-\omega_q))=\frac{1+(p+j\Delta\omega_s)/\Delta\omega_q}{1+j\lambda[1+(p+j\Delta\omega_s)/\Delta\omega_q]}. \qquad (59)$$

$\Delta\omega_q=\omega_q/2Q_q$, $\lambda=\omega_qC_0r_q$, $\Delta\omega_s=\omega_s-\omega_q$.

To calculate $S_c(\omega+\omega_s)$ one has to find the ratio:

$$\frac{R_m(\omega_s)Z_m^{-1}(p+j\omega_s)}{R_c(\omega_s)Z_c^{-1}(p+j\omega_s)}=\frac{-j\xi_b+\bar{Z}_q(p+j\Delta\omega_s)}{-j\xi_b+\bar{Z}_q(j\Delta\omega_s)}, \qquad (60)$$

where $\xi_b=(\omega_qC_2r_q)^{-1}$.

To use (58), (59), (60) in (20), (21) we need just to put p = jω. Taking into consideration only the shot noise of collector current as the main source of PM and AM noise we find $S_{ic}=2qI_{c0}$, where q is a magnitude of electronic charge, and for the chosen operation mode we have $2S_{ic}(\omega+\omega_s)/\left[I_{c1}(V_{b1}^0)\right]^2=1{,}2\cdot10^{-15}\ 1/\text{Hz}$.

Results of calculations $S_b(\omega_s+\omega)$ and $S_c(\omega_s+\omega)$ based on linear model (20), (21) are presented in Fig.5.

One can see the asymmetry of upper and lower bands of PSD's both base voltage and collector voltage. In the PSD $S_c(f_s\pm f)$ lower band is less than the upper one, and there is minimum of this PSD at the offset from the carrier close to $[-(f_s-f_q)]$. This result is similar to the one obtained in [6]. In the PSD $S_b(f_s\pm f)$ upper band is less. But the difference between these bands becomes significant only near the offsets $\pm[f_A-f_s]$, where f_A – is antiresonant frequency of the resonator. At the offset $[f_A-f_s]$ the upper bans of this PSD has a deep minimum.

Using (29), (30), we calculated PSD's of PM and AM noises corresponding to (20), (21) (Fig.6). One can see from Fig.6 and Fig.5 that the asymmetry of PSD's shown in Fig.5 is caused by the influence of mutual PSD's of PM and AM noises. The asymmetry is most significant in the regions where absolute value of mutual PSD is close to PSD's of PM and AM noise.

From Fig.6 one can see that the PSD of AM noise goes to infinity, when $(\Omega/2\pi)\to 0$. Basing on nonlinear model we calculated $S_{\phi_b}(\Omega)$, $S_{a_b}(\Omega)$ and $S_b(\omega_s+\omega)$

using (51), (52), (53) in the region of rather small values of $(\Omega/2\pi)$ where an influence of mutual PSD is small. Results of these calculations are shown in Fig.7.

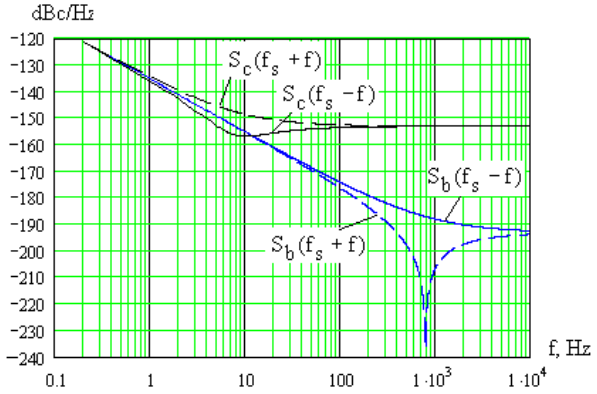


Fig.5. Upper and lower bands of the PSD's $S_b(\omega_s+\omega)$ and $S_c(\omega_s+\omega)$ in Clapp crystal oscillator (Linear model).

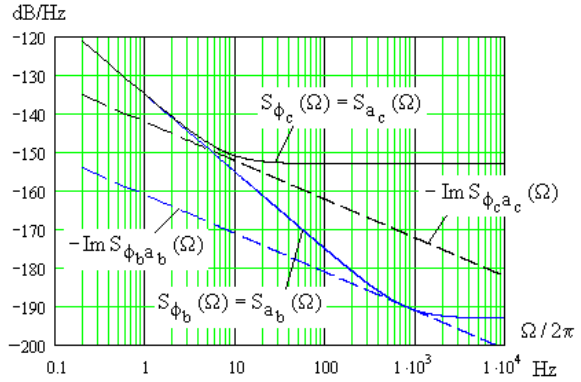


Fig.6. PM and AM noise PSD's and mutual PSD's of base voltage and collector voltage in Clapp crystal oscillator. (Linear model).

One can see that the difference between $S_{a_b}(\Omega)$ and $S_{\phi_b}(\Omega)$ becomes significant when $(\Omega/2\pi) < f_q/2Q_q = 5\,\mathrm{Hz}$. At the frequencies (10…30) Hz results of calculation $S_{a_b}(\Omega)$ and $S_{\phi_b}(\Omega)$ using linear and simplified nonlinear model are almost equal. At higher values of $(\Omega/2\pi)$ results of calculations using complete nonlinear model and linear model practically coincide.

9. CONCLUSION

The results, presented in this paper, show the advantages and shortcomings of the approach to PM and AM noise analysis in oscillators that was offered in [6]. We paid main attention to the approach to

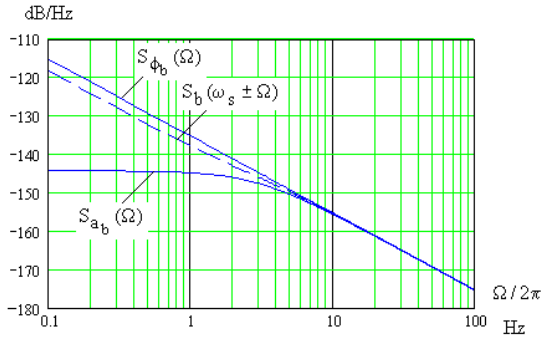


Fig.7. PM and AM noise PSD's of base voltage and noise to signal ratio $S_b(\omega_s\pm\Omega)$, calculated using nonlinear model of Clapp oscillator.

calculation. That is why we considered only one noise source in the oscillator model. Taking into consideration an influence of other wideband noise sources (thermal noise of base resistance, recombination noise, thermal noise of the resonator) is possible in the frame of the approach offered in [6] and developed in this paper. It would give more accurate estimation of PM and AM noise of particular oscillator at all offsets from the carriers that are interesting for applications.

REFERENCES

[1] Bershtein I.L."Fluctuations of amplitude and phase of electronic tube oscillator", Izv. AS USSR Phyziks, 1950, v.14, N2, pp.145-173 (in Russian),

[2] Malakhov A.N. "Fluctuations in selfoscillating systems", M., Science, 1968 (in Russian).

[3] V.Zalud, V.N.Kuleshov, "Noise in semiconductor devices". Soviet Radio, Moscow, 1977 (in Russian).

[4] D.B.Leeson. A Simple Model of Feedback Oscillator Noise Spectrum. Proccedings of IEEE, 1966, vol. 54, N.2, pp. 329-330.

[5] G.Sauvage, "Phase noise in oscillators: A mathematical analysis of Leeson model", IEEE Trans. Instrum. Meas. 1977, vol. IM26, N4, pp. 408-410.

[6] S.Galliou, F.Sthal, and M.Mourey, "New Phase Noise Model for Quartz Crystal Oscillator: Application to the Clapp Oscillator", IEEE Trans.Ultrason., Ferroelect., Freq. Contr., 2003, vol.50, pp.1422-1426.

[7] F.N.H. Robinson, "Noise in junction diodes and bipolar transistors at moderately high frequensis", Electronic Engineering, Febr., 1969, pp. 218-220.

[8] P.E.Gray and C.L.Searle, "Electronic Principles: Physics, Models and Circuits", 1969, John Wiley and Sons, N.Y.

[9] Levin B.R. Theoretical background of statistical radioengineering. V.1. Moscow. Sov.radio. 1969. (In Russian).

NEW DEVELOPMENT OF PM AND AM NOISE ANALYSIS IN CRYSTAL OSCILLATORS: AN INFLUENCE OF 1/f NOISE SOURCES

T.I.Boldyreva

Moscow Power Engineering Institute (Technical University),
14 Krasnokasarmennaya, Moscow 111250, Russia, BoldyrevaTI@mpei.ru

ABSTRACT

The approach to analysis of power spectral density (PSD) of crystal oscillators (XO) output oscillations and PM noise PSD, developed in the paper [1], is applied to calculations of PM and AM noise caused by 1/f noise sources of bipolar junction transistor used in the XO circuit.

The calculations gave the next results. The output oscillation PSD (noise-to-signal ratio) is as a rule an asymmetric function of the frequency offset from the carrier. The type of asymmetry depends on the output point of the oscillator. The PM noise PSD of the XO is defined by PM noise of the BJT amplifier when Fourier frequency is less than a half bandwidth of the crystal resonator, in accordance with Leeson formula. When Fourier frequency is more than a half of the resonator bandwidth the PM noise of the XO is defined by combination of AM and PM noise of the amplifier.

Two models of the XO were considered: a quasilinear model and a nonlinear model. Results of their comparison are presented.

The examples of calculation of PM and AM noise caused by 1/f noise sources in 10 MHz XO with quality factor of $1 \cdot 10^6$ are presented.

1 INTRODUCTION

Most of the papers devoted to calculation of 1/f noise influence on PM noise of quartz crystal oscillators (XO) are based on Leeson's model. In the paper [1] more general approach to PM noise analysis in XO was offered and applied to the Clapp oscillator. But to apply this approach to prediction of PM noise caused by 1/f noise in active device (bipolar junction transistor (BJT)) one has to develop and modify it.

In the paper [2] such development was made to improve prediction of PM and AM noise caused by wideband noise sources. In this paper the approach developed in [2] is applied to prediction of PM and AM noise in XO caused by 1/f noise in BJT. This development is based also on the 1/f noise model of BJT that was used to analyze PM and AM noise in BJT amplifiers [3-5].

2. THE OSCILLATOR MODEL AND ALGORITHM DESCRIPTION

Following the papers [1,2] we use analysis of Clapp oscillator (Fig.1) both to develop an algorithm of PM-AM noise calculation and to illustrate its application numerically.

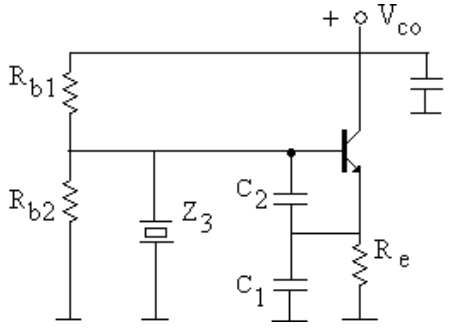


Fig.1. The circuit diagram of the Clapp oscillator

Circuit model of this oscillator, which is used in this paper, is shown in Fig.2.

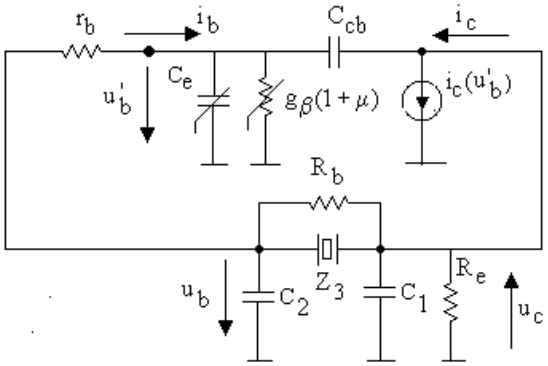


Fig.2. Circuit model of the Clapp oscillator used to analyze 1/f noise influence

In this model fluctuation of recombination conductivity $\mu(t) = \Delta g_\beta(t) / g_\beta$ is considered as a primary source of

1/f noise in BJT [3-5]. An influence of μ(t) causes phase and amplitude modulation of output current of the BJT and input current. This modulation can be calculated, supposing that there is no feedback circuit but the voltages at the input and at the output of the BJT are equal to ones that exist in steady state operation of the XO. Then we consider this modulation influence on the XO using its equations with closed feedback circuit and express PM and AM noise of the oscillator in terms of PM and AM noise of the BJT without feedback circuit. Using these equations we express power spectral densities (PSD) of PM and AM noise of the XO in terms of PSD $S_\mu(\Omega)$ of primary source of 1/f noise.

3. EQUATIONS OF STEADY STATE MODE OF XO OPERATION AND BASIC ASSUMPTIONS

Suppose that the base voltage

$$u_b(t) = V_{b0} + V_{b1}\cos[\omega_s t + \varphi_s]$$

has harmonic component with complex amplitude

$$\dot{V}_{b1} = V_{b1}e^{j\varphi_s}. \quad (1)$$

Assume that using the BJT model shown in Fig.2 we calculated the first harmonic components of collector current $\dot{I}_{c1}$ and base current $\dot{I}_{b1}$

$$\dot{I}_{c1}(V_{b1}) = I_{c1}(V_{b1})e^{j\varphi_s - j\Delta\varphi_c}, \quad \dot{I}_{b1}(V_{b1}) = I_{c1}(V_{b1})e^{j\varphi_s + j\Delta\varphi_b}, \quad (2)$$

Here phase shifts ($-\Delta\varphi_c$) and $\Delta\varphi_b$ are not equal to zero as it was supposed in [2], because PM noise of the BJT is connected with fluctuations of the ($-\Delta\varphi_c$) and $\Delta\varphi_b$ caused by fluctuations of recombination conductivity $g_\beta\mu(t)$.

An equation of the XO steady state operation looks as follows

$$\dot{V}^0_{b1} = Z_m(j\omega_s)\dot{I}_{c1}(V^0_{b1}) - Z_b(j\omega_s)\beta^{-1}\dot{I}_{c1}(V^0_{b1}), \quad (3)$$

where $Z_m(j\omega_s)$ – mutual impedance connecting output voltage of the feedback circuit with its input current, $Z_b(j\omega_s)$ – output impedance of the feedback circuit.

To simplify analysis we assume that BJT is chosen so that

$$\left|\frac{\dot{I}_{b1}}{\dot{I}_{c1}}\right| \approx \beta^{-1} << 1,$$

where β - common emitter configuration current gain of the BJT. Under this assumption instead of (3) we have the next equation

$$\dot{V}^0_{b1} = Z_m(j\omega_s)\dot{I}_{c1}(V^0_{b1}). \quad (4)$$

Substituting (2) in (4) and supposing, that

$$Z_m^{-1}(j\omega_s) = G_m(\omega_s) + jB_m(\omega_s), \quad (5)$$

we obtain

$$B_m(\omega_s) = -g_{m1}(V^0_{b1})\sin\Delta\varphi_c, \quad (6)$$

$$G_m(\omega_s) = g_{m1}(V^0_{b1})\cos\Delta\varphi_c. \quad (7)$$

where

$$g_{m1}(V_{b1}) = \frac{I_{c1}(V_{b1})}{V_{b1}} \quad (8)$$

- an average mutual conductivity of the BJT.

From (6), (7) the next equation for oscillation frequency can be obtained:

$$\frac{B_m(\omega_s)}{G_m(\omega_s)} = -\mathrm{tg}\Delta\varphi_c. \quad (9)$$

It is assumed further that the BJT is chosen so that $\Delta\varphi_c<0{,}2$. Then $\cos\Delta\varphi_c \cong 1$ and steady state amplitude is practically the same as in the case, when $\Delta\varphi_c = 0$ [2].

4. CALCULATION OF NOISE TO SIGNAL RATIO USING QUAZILINEAR MODEL OF XO

Let us suppose that we have calculated PM and AM noise modulation of the current $i_c(u_b)$ caused by μ(t) with ideal harmonic voltage at the BJT input. Then complex amplitude of the first harmonic of collector current can be presented in the form

$$\dot{I}_{c1,f}(V^0_{b1}) = I_{c1}(V^0_{b1})(1 + a_i(t))e^{j\varphi_s - j\Delta\varphi_c + j\phi_i(t)}, \quad (10)$$

where $a_i(t)$ and $\phi_i(t)$ –relative amplitude and phase shift fluctuations caused by μ(t). Suppose that they are expressed in the form

$$a_i = T_a\mu(t), \quad (11)$$

$$\phi_i = T_\phi\mu(t), \quad (12)$$

where T_a, T_ϕ - transformation coefficients.

Supposing that $\sqrt{\overline{a_i^2(t)}} << 1$ and $\sqrt{\overline{\phi_i^2(t)}} << 1$, we can write instead of (10)

$$\dot{I}_{c1,f}(V^0_{b1}) = \dot{I}_{c1}(V^0_{b1}) + \dot{I}_{nc}, \quad (13)$$

where

$$\dot{I}_{nc} = (a_i(t) + j\phi_i(t))\dot{I}_{c1}(V^0_{b1}) \quad (14)$$

- equivalent noise current caused by 1/f noise of the BJT.

Using (13) we can write an equation of the oscillator that is similar to equation (10) from the paper [2]:

$$Z_m^{-1}(j\omega + j\omega_s)\dot{V}_{b1} = \dot{I}_{c1}(V_{b1}) + \dot{I}_{nc}. \quad (15)$$

When we look for the solution of this equation in the form

$$\dot{V}_{b1} = \dot{V}^0_{b1} + \dot{V}_{nb}, \quad (16)$$

where

$$\dot{V}_{nb} = (a_i(t) + j\phi_i(t))\dot{V}^0_{nb} \quad (17)$$

and replace in equation (15) $\dot{I}_{cl}(V_{bl})$ by $g_{ml}(V^0_{bl})\cos\Delta\varphi_c\dot{V}_{nb}$ we obtain the formula connecting Fourier components of $\dot{I}_{nc}$ and $\dot{V}_{nb}$ in the frame of quazilinear model:

$$\dot{V}_{n,b}(j\omega) = \frac{1}{Z^{-1}_m(j\omega + j\omega_s) - g_{ml}(V^0_{bl})\cos\Delta\varphi_c}\dot{I}_{n,c}(j\omega) \quad (18)$$

This formula is similar to formula (14) from [2]. It can be used to express the PSD of the base voltage (noise to signal ratio) in terms of PSD of the normalized noise current with complex amplitude

$$S_b(\omega+\omega_s) = \left|\frac{Z^{-1}_m(j\omega_s)}{Z^{-1}_m(j\omega+j\omega_s) - g_m(V^0_{bl})\cos\Delta\varphi_c}\right|^2 S_{inc}(\omega+\omega_s), \quad (19)$$

where

$$S_{inc}(\omega_s+\omega) = 0{,}5\left[S_{\phi_i}(\omega) + S_{a_i}(\omega) - 2\,\mathrm{Im}S_{\phi_i a_i}(\omega)\right] \quad (20)$$

is calculated using formula (48) from [2]. From (11), (12) it follows that

$$S_{\phi_i}(\omega) = T^2_\phi S_\mu(\omega); \quad S_{a_i}(\omega) = T^2_a S_\mu(\omega); \quad (21)$$

$$\mathrm{Im}S_{\phi_i a_i} = 0 \quad (22)$$

and

$$S_{i_{nc}}(\omega_s+\omega) = 0{,}5\left[T^2_\phi + T^2_a\right]S_\mu(|\omega|). \quad (23)$$

As $S_\mu(|\omega|)$ in (23) is an even function of the offset ω from the carrier the PSD $S_{i_{nc}}(\omega_s+\omega)$ is symmetrical. An influence of the XO transfer function can cause asymmetry of the PSD $S_b(\omega_s+\omega)$ [2]. It can lead to transformation of AM noise to PM noise and vice versa.

To analyze this transformation we need to express PM and AM noise of the voltage $\dot{V}_{nb}$ in terms of PM and AM noise of the current $\dot{I}_{nc}$. In the paper [2] the PSD's of PM and AM noises were calculated directly from PSD $S_b(\omega_s+\omega)$ (formulae (29), (30) in [2]). But these formulae are valid for stationary additive noise $\dot{I}_{nc}$. So they are not applicable in this case.

5. PM-AM NOISE CALCULATION BASED ON QUASILINEAR MODEL

To express $\phi_b(t)$, $a_b(t)$ in terms of $\phi_i(t)$, $a_i(t)$ we have to use symbolic equation from which an equation (18) was obtained [2,6]

$$[Z^{-1}_m(p+j\omega_s) - g_m(V^0_{bl})\cos\Delta\varphi_c](a_b + j\phi_b) = \\ = Z^{-1}_m(j\omega_s)(a_i + j\phi_i) \quad (24)$$

In this equation p=d/dt. As in [2] we put

$$Z^{-1}_m(p+j\omega_s) = G_m(p,\omega_s) + jB_m(p,\omega_s). \quad (25)$$

To simplify calculations we put here $\Delta\varphi_c = 0$. Than $Z^{-1}_m(j\omega_s) = g_m(V^0_{bl}) = G_m(0,\omega_s)$. Using (25) we obtain from (24) two equations for real variables $\phi_b(t)$ and $a_b(t)$:

$$\left[\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1\right]a_b - \frac{B_m(p,\omega_s)}{G_m(0,\omega_s)}\phi_b = a_i, \quad (26)$$

$$\frac{B_m(p,\omega_s)}{G_m(0,\omega_s)}a_b + \left(\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1\right)\phi_b = \phi_i, \quad (27)$$

If $Z^{-1}_m(p+j\omega_s)$ is real function of p, than $B_m(p,\omega_s) = 0$ and from (26), (27) it follows:

$$\left[\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1\right]a_b = a_i, \quad (28)$$

$$\left(\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1\right)\phi_b = \phi_i, \quad (29)$$

Than for PSD's of AM and PM noises we have

$$S_{a_b}(\Omega) = \left|\frac{G_m(j\Omega,\omega_s)}{G_m(0,\omega_s)} - 1\right|^{-2}\cdot S_{a_i}(\Omega), \quad (30)$$

$$S_{\phi_b}(\Omega) = \left|\frac{G_m(j\Omega,\omega_s)}{G_m(0,\omega_s)} - 1\right|^{-2}\cdot S_{\phi_i}(\Omega), \quad (31)$$

It means that at the region of offsets ω where $\left|Z^{-1}_m(j\omega+j\omega_s)\right|$ is symmetrical function of ω AM modulation of the BJT current is transformed only in AM modulation of the XO oscillations, and PM is transformed in PM.

General symbolic solution of equations (26), (27) looks as follows:

$$a_b = A_b(p)a_i + A_{mb}(p)\phi_i, \quad (32)$$

$$\phi_b = -A_{mb}(p)a_i + A_b(p)\phi_i, \quad (33)$$

where

$$A_b(p) = \Delta_b(p)/\Delta_L(p), \quad (34)$$

$$A_{mb}(p) = \Delta_{mb}(p)/\Delta_L(p), \quad (35)$$

$$\Delta_L(p) = \left[\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1\right]^2 + \left(\frac{B_m(p,\omega_s)}{G_m(0,\omega_s)}\right)^2; \quad (36)$$

$$\Delta_b(p) = \frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1; \quad \Delta_{mb}(p) = \frac{B_m(p,\omega_s)}{G_m(0,\omega_s)}. \quad (37)$$

Substituting (11), (12) in (32), (33), we obtain

$$a_b = [T_a A_b(p) + T_\phi A_{mb}(p)]\mu(t), \quad (38)$$

$$\phi_b = [T_\phi A_b(p) - T_a A_{mb}(p)]\mu(t), \quad (39)$$

From (38), (39) we find PSD's of $a_b(t)$ and $\phi_b(t)$:

$$S_{\phi_b} = \left|T_\phi A_b(j\Omega) - T_a A_{mb}(j\Omega)\right|^2 S_\mu(\Omega), \quad (40)$$

$$S_{a_b} = \left|T_a A_b(j\Omega) + T_\phi A_{mb}(j\Omega)\right|^2 S_\mu(\Omega), \quad (41)$$

$$S_{\phi_b a_b} = [T_\phi A_b(j\Omega) - T_a A_{mb}(j\Omega)] [T_a A_b(-j\Omega) + T_\phi A_{mb}(-j\Omega)] S_\mu(\Omega). \quad (42)$$

Using formula

$$S_b(\omega_s + \omega) = 0{,}5\left[S_{\phi b}(\omega) + S_{ab}(\omega) - 2\,\mathrm{Im} S_{\phi b a b}(\omega)\right], \quad (43)$$

we obtain result that coincide with the one calculated by means of (19). But one can't obtain (40), (41) and (42) directly from (19).

6. PM AND AM NOISE CALCULATION BASED ON NONLINEAR MODEL

The main shortcoming of the approach, based on quasilinear model, is a fundamental error in prediction of AM noise PSD at low Fourier frequencies ($\Omega \to 0$). As a consequence of this error one obtains an error in prediction of PSD of the oscillation $S_b(\omega_s + \omega)$ at small offsets from the carrier.

To escape these errors one has to use a nonlinear model. Using the approach, developed in [2], we obtained the next symbolic equations

$$\left[\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - \frac{\sigma}{g_{ml}}\right] a_b - \frac{B_m(p,\omega_s)}{G_m(0,\omega_s)}\phi_b = a_i, \quad (44)$$

$$\left[\frac{B_m(p,\omega_s)}{G_m(0,\omega_s)}\right] a_b - \left(\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1\right)\phi_b = \phi_i, \quad (45)$$

where $\sigma = \left[dI_{c1}/dV_{b1}\right]_{V_{b1}=V_{b1}^0}$, and in the stable steady state mode of operation

$$(\sigma/g_{ml}) < 1. \quad (46)$$

These equations differ from equations (37), (38) in the paper [2] only by sources of noise a_i, ϕ_i.

When $Z_m^{-1}(j\omega + j\omega_s)$ is a real function of p a nonlinear model gives $S_{\phi b}(\Omega)$ that coincides with (31). But to calculate AM noise PSD one has to use the formula:

$$S_{a_b}(\Omega) = \left|\frac{G_m(j\Omega,\omega_s)}{G_m(0,\omega_s)} - \frac{\sigma}{g_{ml}}\right|^2 S_{a_i}(\Omega). \quad (47)$$

It predicts a limited value of the transfer function, connecting $S_{a_i}(\Omega)$ and $S_{a_b}(\Omega)$, at Ω = 0. It corresponds to real situation with AM noise.

Symbolic solution of equations (44), (45) can be presented as follows

$$a_b = A_{a,b}(p)a_i + A_{t,b}(p)\phi_i, \quad (48)$$

$$\phi_b = -A_{t,b}(p)a_i + A_{\phi,b}(p)\phi_i, \quad (49)$$

where

$$A_{a,b}(p) = \frac{\Delta_{a,b}(p)}{\Delta(p)}, \qquad A_{a,b}(p) = \frac{\Delta_{\phi,b}(p)}{\Delta(p)},$$

$$A_{t,b}(p) = \frac{\Delta_{t,b}(p)}{\Delta(p)},$$

$$\Delta(p) = \left[\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - \frac{\sigma}{g_{ml}}\right]\left[\frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1\right] - \frac{B_m^2(p,\omega_s)}{G_m^2(0,\omega_s)},$$

$$\Delta_{a,b}(p) = \frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - 1, \quad \Delta_{\phi,b}(p) = \frac{G_m(p,\omega_s)}{G_m(0,\omega_s)} - \frac{\sigma}{g_{ml}},$$

$$\Delta_{t,b}(p) = \frac{B_m(p,\omega_s)}{G_m(0,\omega_s)}.$$

Using (11), (12), we obtain from (48), (49)

$$a_b = [T_a A_{a,b}(p) + T_\phi A_{t,b}(p)]\mu(t), \quad (50)$$

$$\phi_b = [T_\phi A_{\phi,b}(p) - T_a A_{t,b}(p)]\mu(t). \quad (51)$$

From these equations we obtain the formulae for PSD's of $a_b(t)$ and $\phi_b(t)$ that follow from nonlinear XO model

$$S_{\phi_b} = \left|T_\phi A_{\phi,b}(j\Omega) - T_a A_{t,b}(j\Omega)\right|^2 S_\mu(\Omega), \quad (52)$$

$$S_{a_b} = \left|T_a A_{a,b}(j\Omega) + T_\phi A_{t,b}(j\Omega)\right|^2 S_\mu(\Omega), \quad (53)$$

$$S_{\phi_b a_b} = [T_\phi A_{\phi,b}(j\Omega) - T_a A_{t,b}(j\Omega)] \cdot [T_a A_{a,b}(-j\Omega) + T_\phi A_{t,b}(-j\Omega)] S_\mu(\Omega). \quad (54)$$

The PSD of the oscillations at the input of the BJT can be calculated from (52-54) and (43).

7. 1/F PM AND AM NOISE OF CLAPP OSCILLATOR

As an example of general approach application to particular circuit we consider Clapp oscillator shown in Fig.1. The resonator model parameters are same as in the paper [2]: f_q = 10MHz, r_q = 100 Ohm, Q_q = 10^6, C_0 = 1 pF. We suppose that in the steady state operation mode the power, dissipated in the crystal, is $P_q = 10 \cdot 10^{-6}$ W.

To analyze 1/f noise influence on the oscillator we consider the BJT model shown in Fig.2. In the case considered here we took the BJT that has f_T = 600 MHz, r_b = 150 Ohm, C_C = 3 pF, C_e = 6 pF. An in [2] we have chosen I_{c0} = 0,33 mA. At this collector current β = 13. The phase shift between first harmonics of base voltage and collector current is $(-\Delta\varphi_c)$ = -0,095.

The transformation coefficients in (11), (12) were calculated supposing that the average emitter current $I_{c0} + I_{b0} = \text{const}$. The calculation gave:

$$T_\phi = 0{,}013,\ T_a = -0{,}21. \qquad (55)$$

Power spectral density $S_\mu(\omega)$ was obtained from results of measurements [7]. At the operating point $I_{c0} = 0{,}33$ mA.

$$S_\mu(\omega) = 5{,}5\cdot 10^{-10}\left(\frac{\omega}{2\pi}\right)^{-1}. \qquad (56)$$

To use quasilinear model (19), (23) we need the next formulae, that were obtained for Clapp oscillator:

$$Z_m^{-1}(p + j\omega_s) = R_{m0}^{-1}(\omega_s)[-j\xi_0 + \hat{Z}_q(p + j\Delta\omega_s)], \qquad (57)$$

where

$$\hat{Z}_q(p + j\Delta\omega_s) = \frac{1 + (p + j\Delta\omega_s)/\Delta\omega_q}{1 + j\lambda[1 + (p + j\Delta\omega_s)/\Delta\omega_q]} \qquad (58)$$

and

$$\Delta\omega_s = \omega_s - \omega_q;\quad \Delta\omega_q = \omega_q/2Q_q;\quad \lambda = \omega_q C_0 r_q,$$

$$R_{m0}^{-1}(\omega_s) = \omega_q^2 C_1 C_2 r_q,\quad \xi_0 = (C_1 + C_2)/\omega_q C_1 C_2 r_q.$$

Results of base voltage PSD calculation using (19), (23), (55) – (58) are shown in Fig.3. Using Fig.3 one can find noise to signal ratio at negative ($\omega < 0$) and positive ($\omega > 0$) offsets from the carrier. It is caused both AM noise and PM noise. One can notice an asymmetry of upper and lower bands, and deep notch at the offset $(\omega_A - \omega_S)/2\pi$ due to rejection of noise Fourier components close to antiresonance frequency ω_A.

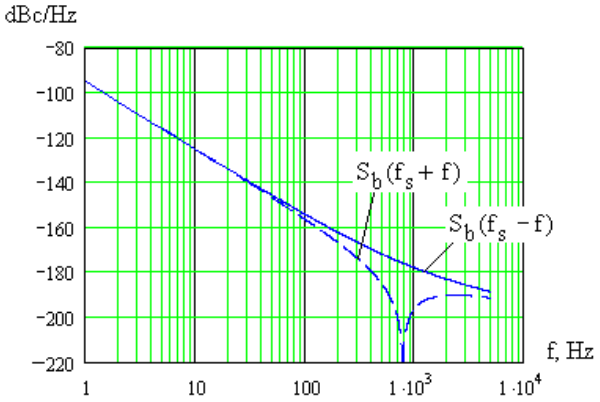


Fig.3. Upper and lower bands of the PSD $S_b(\omega+\omega_S)$ in Clapp crystal oscillator. (Linear model).

To calculate PSD's of PM and AM noise contributing to $S_b(\omega+\omega_S)$ (Fig.3) we use the formulae (40), (41). The curves shown in Fig.4 present PSD's $S_{\phi_b}(\Omega)$ and $S_{a_b}(\Omega)$. One can see that at low Fourier frequencies $S_{\phi_b}(\Omega)$ changes as Ω^{-3}. But at higher frequencies the slope of the curve decreases up to Ω^{-1}. At low Fourier frequencies the main contribution to $S_{\phi_b}(\Omega)$ is provided by $S_{\phi_i}(\Omega)$ (31). At higher Fourier frequencies a contribution of transformation $[a_i(t) \to \phi_b(t)]$ predominates. At $\Omega = \omega_A - \omega_s$ the PSD's $S_{\phi_b}(\Omega)$ and $S_{a_b}(\Omega)$ are equal and the notch shown in Fig.3 is caused by mutual correlation of $\phi_b(t)$ and $a_b(t)$. One can notice also that the PSD $S_{a_b}(\Omega)$ calculated from linear model grows as Ω^{-3} when $\Omega \to 0$. As it was mentioned earlier, it is not consistent with real AM noise PSD in the stable oscillator.

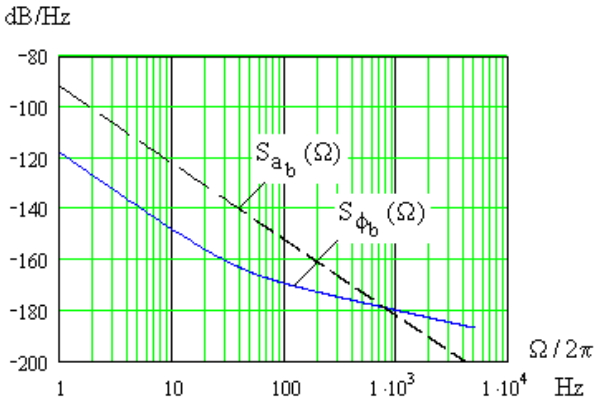


Fig.4. PM and AM noise PSD's of base voltage in Clapp crystal oscillator. (Linear model).

Correct results of both AM and PM noise and noise/signal ratio calculation were obtained using nonlinear model (44), (45), (46), (52), (54) and (43) with $(\sigma/g_{m1}) = 0{,}45$.

Results of PM and AM noise calculation are shown in Fig.5. It follows from Fig.5 that at low frequencies $\Omega < \Delta\omega_q/2$ the PSD $S_{a_b}(\Omega)$ grows as Ω^{-1} when $\Omega \to 0$. It is consistent with nonlinear theory of AM noise in oscillators [6] and experimental results.

One can see that at sufficiently low Fourier frequencies the PSD $S_{a_b}(\Omega)$ of AM noise becomes less than $S_{\phi_b}(\Omega)$ (in Fig.5 it is at $(\Omega/2\pi) < 0{,}2$ Hz). So at low Fourier frequencies PM noise $S_{\phi_b}(\Omega)$ predominates and it gives the main contribution to the PSD of the oscillation $S_b(\omega+\omega_S)$. This conclusion is confirmed by results of $S_b(\omega+\omega_S)$ calculations that were fulfilled using (44) and (52-54). They are shown in Fig.6. One can see from Fig.6 and Fig.3 that at rather large offsets from the carrier ($(|\omega|/2\pi) < 10$ Hz in Fig.3, 6) results of $S_b(\omega+\omega_S)$ calculations based on nonlinear and linear

models coincide. The same is true with respect to PSD's $S_{\phi_b}(\Omega)$ and $S_{a_b}(\Omega)$.

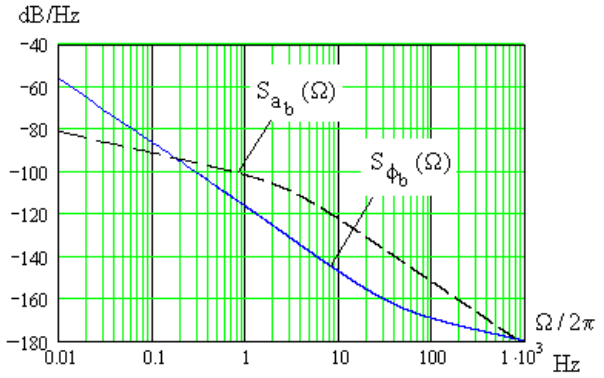


Fig.5. PM and AM noise PSD's of Clapp crystal oscillator base voltage. (Nonlinear model).

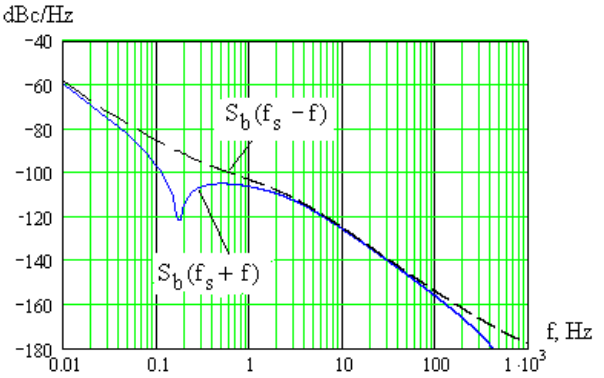


Fig.6. Upper and lower bands of the PSD $S_b(\omega+\omega_S)$ in Clapp crystal oscillator. (Nonlinear model).

But at lower offsets from the carrier nonlinear model predicts results that are quite different from prediction of linear model (Fig.6 and Fig.3). One can notice a very significant asymmetry of $S_b(\omega+\omega_S)$ dependence on the offset from carrier due to correlation between PM and AM noise caused by 1/f noise of the BJT collector current. It is predicted by nonlinear model and can not be predicted by linear model.

9 CONCLUSION

The main result of this paper is development of approach to quantitative analysis of 1/f noise of active device influence on PM and AM noise in quartz crystal oscillator.

An algorithm of PM and AM noise calculation based on results of linear analysis of 1/f noise in the oscillator is offered and applied to Clapp crystal oscillator noise analysis.

It is shown that the linear model well predicts noise to signal ratio at rather large offsets from the carrier ($|\omega| > \omega_q / Q_q$). At small offsets the error can be very large and one needs a nonlinear model to provide good prediction of noise to signal ratio. Basic equations of nonlinear model are also given in this paper.

In particular example we considered only an influence of 1/f noise modulation of collector current on PM and AM noise of base voltage. To obtain the complete PSD's one has to take into consideration 1/f noise modulation of base current and its contribution to total PM and AM noise PSD's. This can be done using the same approach, but it is out of the frame of this paper.

ACKNOWLEDGMENTS

The author would like to thank prof. V.Kuleshov under guidance of whom this research was fulfilled.

REFERENCES

[1] S.Galliou, F.Sthal, and M.Mourey, "New Phase Noise Model for Quartz Crystal Oscillator: Application to the Clapp Oscillator", IEEE Trans.Ultrason., Ferroelect., Freq. Contr., vol.50, pp.1422-1426, Nov.2003

[2] V.N.Kuleshov, "New Development PM and AM Noise Analysis in Crystal Oscillators: an Influence of Wideband Noise Sources", in Proc. EFTF-2005.

[3] V.N.Kuleshov and T.I.Boldyreva, «1/F AM and PM Noise in Bipolar Transistor Amplifiers: Sources, Ways, of Influence, Techniques of Reduction», Proc. 1997 IEEE Int. Freq. Control Symp., 1997, pp.446-455.

[4] T.I.Boldyreva, "Comparative study of phase and amplitude noise of CE, CE-CB and CC-CB bipolar transistor amplifiers", Proc. 1998 IEEE Int. Freq. Contr. Symp. 1998, pp.218-255.

[5] V.N.Kuleshov. "1/F Models of Bipolar Junction Transistor and Their Application to PM and AM Noise Calculation", Proc. 1998 IEEE Int. Freq. Contr. Symp., 1998, pp.167-171.

[6] V.Zalud, V.N.Kuleshov, "Noise in semiconductor devices". Soviet Radio, Moscow, 1977 (in Russian).

[7] V.N.Kuleshov, G.D.Yanushevsky, V.N.Legotin, E.N.Pigalina, "Experimental study bipolar junction transistor flicker noise", Frequency stab. and precision radioengineering , v.2, Moscow, 1983, pp.76-78,

Predictability of a Hydrogen Maser Time Scale

Laurent-Guy Bernier

METAS - Swiss Federal Office of Metrology and Accreditation
Lindenweg 50, CH-3003 Bern-Wabern, Switzerland
e-mail: laurent-guy.bernier@metas.ch fax: +41-31-3233573

Keywords: time scales, hydrogen maser, statistical prediction, optimal linear prediction

Abstract

The METAS Time & Frequency metrology laboratory has acquired a hydrogen maser in 2002. The maser was commissioned at the beginning of 2003. In June 2003 the maser started to contribute to the generation of the UTC(CH) and TA(CH) time scales. The Swiss time scales are currently generated from 3 cesium clocks and 1 hydrogen maser. The clock results are sent monthly to BIPM for contribution to the generation of TAI and UTC. The hydrogen maser is also used in combination with a DDS synthesizer to generate UTC(CH.R), a real-time realization of UTC(CH). The daily steering of UTC(CH.R) is based on a prediction of the hydrogen maser time scale over 1.5 d. Although the hydrogen maser does not have an Automatic Cavity Tuning (ACT) system, we have found that, despite the frequency drift, the time scale is extremely predictable.

1 Introduction

A constant drift is not a problem with respect to predictability. It is the presence of non-stationary discontinuities, such as spontaneous rate steps, that makes prediction difficult. In principle the time scale from an atomic clock affected by the classical set of noise processes (polynomial model) is a non-stationary process with Stationary Second Increment (SSI). SSI processes constitute the class of well behaved noise processes typical of atomic clocks for which the Allan deviation statistic is stationary. SSI processes are also tractable regarding prediction because it is possible to construct linear prediction operators for which the error on the prediction is stationary. Unfortunately atomic clocks are sometimes affected by spontaneous rate steps. It is necessary to detect, estimate and remove the rate steps from the time scale before attempting to apply a prediction algorithm because each step constitutes a discontinuity in the statistical properties which is disruptive to the statistical prediction algorithm. In the case of our hydrogen maser we have detected 4 spontaneous rate steps over a 590 d observation interval. The steps are of the order of 1 ns/d and their origin is unknown. One possible source could be the occasional sudden mechanical relaxation of the glass-ceramic microwave cavity. The present paper focuses on the outstanding frequency stability and predictability of our hydrogen maser.

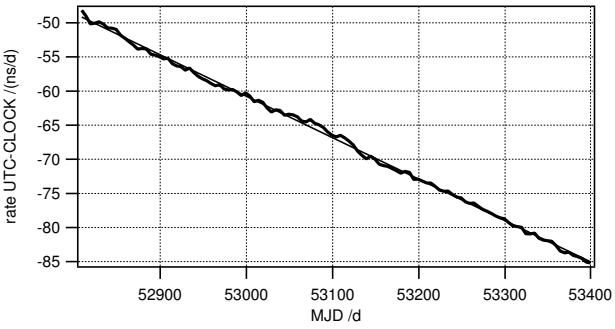


Figure 1: 5 d moving average rate of UTC-CLOCK for hydrogen maser ID 1405701.

2 Basic Time Scale Data

In the BIPM data base our hydrogen maser is known as clock ID 1405701. The basic time scale data is UTC-CLOCK /ns as published by BIPM. Changes in the calibration of the TAI common view GPS link have produced artificial time steps of -50 ns on MJD 53129 and +8 ns on MJD 53385. These steps were removed before analysis. The recorded span is from MJD 52809 to MJD 53399 (590 d) with a 5 d sampling interval. The 5 d moving average rate is shown on Figure 1. The average drift is -0.0607 ± 0.0002 (ns/d)/d. The observed stability includes both the intrinsic hydrogen maser frequency stability and the GPS remote comparison noise.

MJD	ns/d
52857	-0.85
53037	-0.92
53111	-0.85
53352	-0.38

Table 1: Rate steps removed from UTC-CLOCK process (clock ID 1405701).

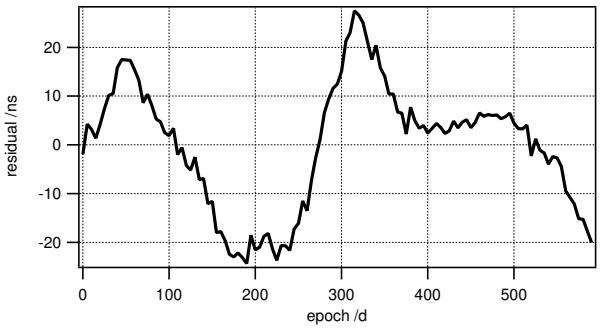


Figure 2: Drift removed residual of UTC-CLOCK before removal of accidental rate steps.

3 Rate Steps Removal

The approach to statistical prediction used in this paper follows a sequence of three operations: pre-processing, prediction and post-processing. The pre-processing consists into estimating and removing the spontaneous rate steps present in the original process. The post-processing consists into adding back the removed rate steps to transform the prediction of the pre-processed time scale into a prediction of the original time scale.

The accidental rate steps are detected and estimated by looking at the drift removed residual of the time scale process. A diverging residual is the sign of an undetected rate step. A rate step has two parameters: the epoch of occurrence and the rate offset. To estimate each rate step the pair of parameters is optimized to minimize the RMS drift removed residual.

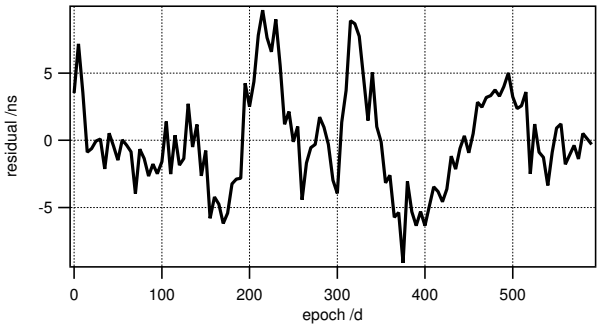


Figure 3: Drift removed residual of UTC-CLOCK after removal of accidental rate steps.

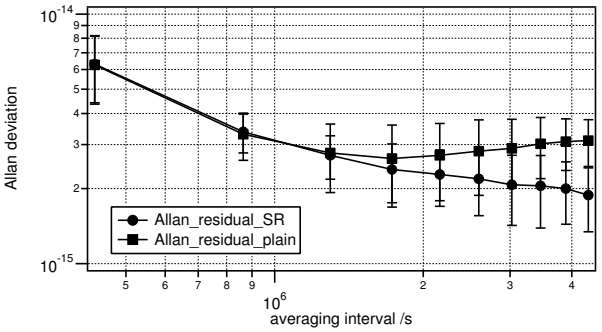


Figure 4: Allan deviation of drift removed residual of UTC-CLOCK with (Allan residual SR) and without (Allan residual plain) rate steps removal.

The four detected and removed rate steps are listed in Table 1. Figure 2 shows the drift removed residual before removal of the rates steps and Figure 3 shows the drift removed residual after removal of the rates steps. The drift removed residual before steps removal has a RMS deviation of 13 ns and is clearly not random. After steps removal the residual is more random and the RMS deviation goes down to 4 ns. Figure 4 shows the Allan deviation of the drift removed residual with and without steps removal. The Allan deviation after step removal is only 2×10^{-15} for an averaging time of 50 d.

The limited accuracy of the estimation of the rate steps does not affect the accuracy of the prediction. Each rate step removed from the original process to generate the residual is compensated by an identical rate step term used in the post-processing after the prediction. In this sense the separation of the original process into a rate steps part and a noise residual part is arbitrary. The requirement is that the residual process must be as close as possible to a SSI random process so that statistical prediction can be applied successfully.

4 Prediction Method

We use the DGSF-1 prediction algorithm for the prediction of the steps removed residual. The prediction algorithm is reported in details in [3]. The DGSF-1 prediction operator is a generalization of the simple second difference SF-1 prediction operator [2]

$$\hat{x}(t+\tau_1) = x(t) + \tau_1\, y(t,\tau_1), \qquad (1)$$

where the time process $x(t)$ is the integral of the instantaneous frequency process $y(t)$

$$x(t) = \int_{-\infty}^{t} y(u)du, \qquad (2)$$

and where $y(t,\tau)$ is the moving average of $y(t)$ over an averaging interval τ

$$y(t,\tau) = \frac{1}{\tau}\int_{t-\tau}^{t} y(u)du. \quad (3)$$

In the SF-1 prediction the prediction $\hat{x}(t+\tau_1)$ is simply the present state $x(t)$ of the time scale process extrapolated into the future, assuming that the initial average rate $y(t,\tau_1)$ will not change during the prediction interval τ_1. In the SF-1 prediction the averaging interval of the rate is identical to the prediction interval and the RMS error on the prediction as a function of the prediction interval is given by the Allan deviation of the time scale process

$$\text{rms}\,\{\epsilon(t,\tau_1)\} = \sqrt{2}\tau_1\sigma_y(\tau_1). \quad (4)$$

The DGSF-1 prediction operator, on the other hand, is a simple generalization of the SF-1 prediction

$$\hat{x}(t+\tau_1) = x(t) + \tau_1\, y(t,\tau_2) + \frac{1}{2}\, d\tau_1^2 \left(1+\frac{\tau_2}{\tau_1}\right), \quad (5)$$

where the averaging interval of the rate τ_2 is different from the prediction interval τ_1. This allows an optimization of the averaging interval τ_2 to minimize the RMS error on the prediction. d is a drift parameter that is also optimized in order to minimize the RMS prediction error. Its function is to compensate a deterministic drift in the time scale process.

5 Prediction Results

Figure 5 shows the sequence of all the individual prediction errors, for a prediction interval τ_1 of 50 d, that can be computed on the basis of the recorded steps removed residual. For a given prediction interval it is the RMS value of all past predictions that is minimized for the estimation of the optimum prediction parameters. In the present case the optimum prediction parameters are $d = -6\times10^{-3}$ (ns/d)/d and $\tau_2 = 115$ d. Using the optimal prediction parameters an RMS value of the prediction error of 5.5 ns is obtained.

6 Generation of UTC(CH.R)

UTC(CH.R) is a real-time realization of UTC(CH) which is performed by means of the daily steering of a DDS synthesizer which is driven by the free running hydrogen maser. The predictive steering method is reported in [4] and involves a prediction of the hydrogen maser time scale over a prediction interval of 1.5 d. Using the optimized DGSF-1 prediction algorithm, the RMS prediction error obtained over this prediction interval is 1 ns. A recording of UTC(CH.R)-UTC(CH)

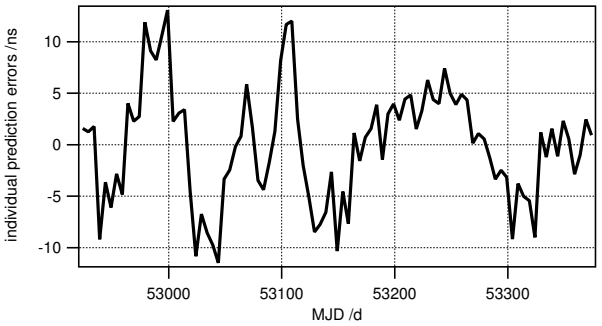


Figure 5: Individual prediction errors obtained by applying the optimized DGSF-1 prediction algorithm to the steps removed residual for a prediction interval of 50 d.

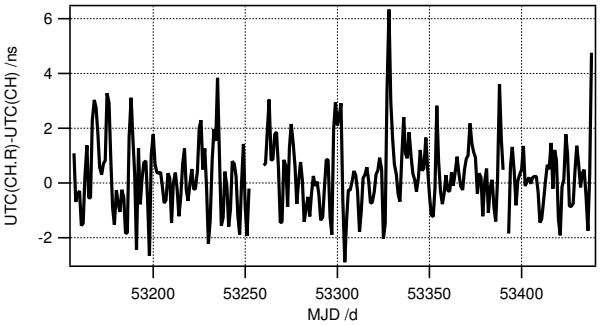


Figure 6: UTC(CH.R)-UTC(CH).

is shown on Figure 6. Thanks to the predictive steering control algorithm the experimental RMS deviation of UTC(CH.R)-UTC(CH) is only 1.3 ns. Simulations reported in [4] predicted a RMS value of 1.9 ns. Hence experimental performance is even better than expected. The visible spikes are due to the transient following each steering of UTC(CH). Just after a rate correction of UTC(CH) the prediction of the maser vs UTC(CH) is biased on the first day. The bias vanishes on the next day.

7 Conclusion

After estimation and removal of the spontaneous rate steps, the time scale of our hydrogen maser is predictable to 5.5 ns RMS over a prediction interval of 50 d. The drift and steps removed residual has an Allan deviation of 2×10^{-15} over a 50 d averaging interval. The observation that the hydrogen maser can have such an outstanding long-term stability without an ACT demonstrates once more that the design of an ACT system that actually removes the frequency drift of the hydrogen maser without degrading its intrinsic long-term frequency stability is a very demanding challenge. This paper shows that in some applications the drift can be compensated by software instead of relying

on an ACT. Unfortunately the presence of spontaneous rate steps makes prediction more difficult. It would be very interesting to study the physical origin of the rate steps. The short-term predictability of the hydrogen maser time scale is also very good as illustrated by the performance of UTC(CH.R).

References

[1] L. G. Bernier, "Experimentation at METAS with a Simple Steering Algorithm based on Linear Prediction," in *Proc. IVth International Time Scale Algorithms Symposium*, BIPM, Sèvres, March 18-29, 2002.

[2] L. G. Bernier, "Use of the Allan Deviation and Linear Prediction for the Determination of the Uncertainty on Time Calibrations Against Predicted Time scales," in *IEEE Trans. on Instrum. and Meas.*, vol. 52, no. 2, April 2003, pp. 483-486.

[3] L. G. Bernier, "Application of the GSF-1 Algorithm to the Near-Optimal Time scale Prediction of the Hydrogen Maser," in *Proc. PTTI 2003*, San-Diego, December 2003, pp.221-236.

[4] L. G. Bernier, G. Dudle, "Practical Performance of the UTC(CH.R) Real Time Realization of UTC(CH) and Prospects for Improvement," in *Proc. EFTF 2004*, University of Surrey, Guilford, U.K., April 5-7 2004.

TIME SCALE PREDICTION USING INFORMATION FROM EXTERNAL SOURCES

A. Bauch[(1)], A. Sen Gupta[(2, 3)], E. Staliuniene[(3)]

(1) Physikalisch-Technische Bundesanstalt, Braunschweig, Germany
(2) National Physical Laboratory of India, New Delhi, India
(3) work performed under PTB Contract

ABSTRACT

We studied the prediction of UTC(k) time scales with respect to UTC based on data published in Circular T and on results of time comparisons among the UTC(k)-scales obtained during the period for which the prediction was required. Even if the predictability of one selected scale UTC(i) was insufficient for the envisaged application, combination with data from three European institutes allowed to improve its predictability considerably. The require-ments assumed were those established for the future Galileo System Time GST, the time scales involved were those of IEN, NPL, PTB, and VSL. The additional time comparison data used were those which were generated routinely, based on TWSTFT and GPS CV.

1. INTRODUCTION

Galileo System Time (GST) shall become the time reference for all operations internal to the future European satellite navigation system Galileo and shall also be useful as time reference for external users who will access GST through the Galileo signal in space. GST may find wide applications if it becomes a stable and thus predictable time scale with a reasonably small offset from International Atomic Time (TAI) and thus – modulo an integer number of seconds – from Universal Coordinated Time (UTC). Some details on GST will be given in the following section. For now, one needs to know that GST shall be realized in collaboration of the Galileo Ground Segment and European timing institutes under the coordination of the Galileo Time Service Provider (GTSP). The GTSP establishment is one of the ongoing activities pursued by the Galileo Joint Undertaking.

In this paper we report on a study which may be helpful to develop a simple and reliable algorithm for the realization of GST. Our key idea is that the future GTSP shall rely only on the standard product which each European timing institute "k" realizes and distributes, namely its local realization of UTC, named UTC(k). UTC(k) is published as UTC – UTC(k) in the Circular T of the Bureau International des Poids et Mesures (BIPM). Our work has been focused on the required prediction of the offset between UTC and GST. The proposed prediction algorithm includes the following steps which we will detail further below:

1) Prediction of UTC – GST based on data published in Circular T,
2) Prediction of UTC – UTC(k) based on data published in Circular T,
3) Measurement of GST – UTC(k) in quasi real time,
4) Prediction of UTC – GSTf using UTC(k) as an intermediate time scale,
5) Assessment of the quality of the prediction UTC – UTC(k) based on historical data and determination of a statistical weight in the combination of inputs in step (4).

Since Circular T is published monthly, the steps 1), 2), and 5) are repeated in the same rhythm, whereas the steps 3) and 4) can in principle be done on a day-to-day basis, provided that the measurement data are made available accordingly. Our algorithm is adaptive to the quality of data provided from month to month, and it is open to the inclusion or suspension of institutes "k", whenever a new institute agrees to provide data timely in the required format.

This paper is organized as follows. After a description of GST itself we explain our selection of data for this study in Section 3. In Sections 4 and 5 we describe the prediction algorithm, the time comparison data which we used, and the results of GST prediction, respectively. A discussion of the results and indications of short-comings of the current study and of further work to be done are given in the concluding section.

2. GALILEO SYSTEM TIME

Here we give a short description of the requirements and the scheme of realization of GST. Thereby we want to make the selection of input data and the type of analysis understandable to the reader. GST will be a time scale without leap seconds with a well defined and published offset from UTC, which, however, is not relevant in the present context. A concept of shared responsibility regarding its realization has been adopted:

- The Galileo Operating Company will be responsible for systems, services, and product necessary for the Galileo navigation mission;

- The Galileo Time Service Provider (GTSP) will be responsible for the additional facilities necessary to provide a dedicated time service via Galileo.

The quality of the time service was laid down in the following requirements (expressed in our own words):

1. GST shall be steered to TAI with a maximum offset of 50 ns for 95% of the time (over any yearly time interval),
2. The difference GST − TAI shall be known with an uncertainty of 28 ns (95% confidence).

The latter requirement needs some interpretation. Due to the nature of the realization of TAI by the BIPM Time Section, TAI is available only in deferred time with a maximum delay of 45 days. It is technically trivial to ensure that the uncertainty of *final* GST - TAI values which could, e. g., be published in the BIPM Circular T is much below the required value. The problem lies in the need to predict the difference for a full month based on data from the past month. In this context, the above requirement specifies the acceptable deviation between the predicted time difference and the truth becoming obvious some weeks later.

At least in the beginning of the system operations, GST will be realized as a physical signal in a so-called Precise Timing Facility (PTF) [1] which will become part of the Galileo Ground Segment. GST will be based on an ensemble of atomic clocks maintained at the PTF, and will be routinely compared to several UTC(k)-time scales of European timing institutes "k" using Two-Way Satellite Time and Frequency Transfer (TWSTFT) and GNSS Common View (CV) time transfer. Here GNSS (Global Navigation Satellite System) stands for the US Global Positioning System as long as the Galileo satellite constellation is not complete. These time transfer links will establish the relation between PTF and the international time community. Data collected in one institute will be forwarded to BIPM so that GST – UTC can be calculated and published.

Further than that, the collected data shall be used in near real-time to improve the accuracy of the prediction of UTC – GST during the one-month period when definitive values are not yet available. The latter process was simulated in our study using real data collected in the past.

Operation and evaluation of the time links, determination of the UTC - GST time offset, prediction of this offset for the current epoch, and advising the PTF on the necessary frequency steering of GST are the tasks of GTSP [2, 3]. The number and selection of laboratories contributing in this process, and the kind of data they should provide, is still in debate. In our study we assumed that the number is small and, regarding the kind of data, we chose the simplest approach: The future GTSP shall rely only on the standard product which each European timing institute "k" realizes and distributes, namely its local realization of UTC, named UTC(k).

3. SELECTION OF FICTIVE GST AND REFERENCE TIME SCALES

As the data base for our study we wished to have regular time scale comparison data using TWSTFT and GPS between a few institutes over a period of about two years. These institutes should be listed in the Circular T so that the differences UTC − UTC(k) during the same period are known. One UTC(k) had to be given the role of "GST" in our study, since GST is not realized yet. It is named GSTf in the following (f stands for fictive). As said before, we anticipate that GST will be reported in the BIPM Circular T as is GPS time today.

The time scales UTC(k) are typically generated in such a way that a local frequency standard (caesium clock or hydrogen maser) is steered in frequency so that the time difference UTC − UTC(k) remains reasonably small, ±100 ns being a typical goal in these days. Note, that the magnitude of this offset itself has no impact on the usability of a certain time scale in our study. The frequency steering applied in the laboratories needs to be known in order to be able to predict UTC − UTC(k) for the period following one for which data were published in Circular T. With these data available, we could perform the two steps

1) Prediction of UTC – GSTf based on data published in Circular T,
2) Prediction of UTC – UTC(k) based on data published in Circular T.

It turned out that "good" TWSTFT data, with few interruptions, were available between four European institutes, IEN, NPL, PTB, and VSL for the required duration of almost two years for our simulation study. TWSTFT data were also available for the same period between these institutes and USNO whose UTC scale of exceptional stability could serve as a very stable source of excellent predictability, as will be shown later in Section 5.6. As IEN has the particular role of providing the experimental system time E-GST in the Galileo System Test Bed GSTB-V1 [4, 5] whose active phase ended with the end of 2004, UTC(IEN) seemed to be a natural choice for the representation of GSTf in our study. As will be shown subsequently, GSTf was not always fully compliant with the requirements for GST mentioned in Section 2. We shall show that prediction of UTC − GSTf based on TWSTFT and GPS data connecting GSTf to the other institutes can provide the required predictability. In our study we did not address the aspect of generating frequency steering steps, but this is a rather easily added complexity to the prediction algorithm.

4. DESCRIPTION OF THE TIME SCALE PREDICTION PROGRAM AND THE DATA

4. 1 Time Scale Prediction Program

The time scale prediction algorithm, briefly described in the following, is the same as in our earlier study [6]. The basic data used in the prediction are the time differences ΔT = UTC - UTC(k) in the Circular T of the BIPM Time Section, "k" representing one of the contributing timing institutes, including IEN. The differences ΔT for a given month are made available typically around the 15th of the following month. Thus the requirement for prediction is as long as 45 days later than the last published value. As the Circular T data are given at 5-day intervals for one month we typically have 7 data points. Then there are 2 points - namely, the 5th and the 10th of the following month, which can be considered as dead time. Having received the Circular T we are required to predict the ΔT for the period of one month, starting from the 15th. The programme predicts values at 5-day intervals, a total of 7 values for one month, so that it is convenient to compare the predicted numbers with the actual ones in the Circular T for the following month. This procedure is named 7-2-7, namely, 7 past data points, 2 dead time points, and 7 predicted points in the future. Depending on the noise type present in the UTC(k) scale typically three kinds of prediction strategy can be adopted for the 7 future points. These are:

1. Last Value Prediction (LV) - Here all the predicted 7 future points are set equal to the last value for the previous month as given in the Circular T. This is considered optimum in the case of white frequency noise and zero mean frequency offset.

2. Linear Prediction (Lin) - In this case the values of the 7 future points are predicted to be a linear extrapolation of ΔT starting from the last value of the previous month, and with a slope determined by the first and the last value for the previous month.

3. Linear Least-Squares Fit (Linsq) - This is the same as in the Lin case except that the slope is determined by the least square fit to all the 7 points of the previous month. Lin and Linsq are considered closer to optimum when there is a significant frequency offset between UTC and UTC(k).

The prediction interval and the past interval containing the data were chosen as of equal duration. In our earlier study we showed that the use of 5 or 9 points (instead of 7) for the linear prediction was not in general favorable. The algorithm also takes into account intentional rate changes introduced in the UTC(k), by altering the slope of the predicted 7 points. The information on the steering performed is available from the BIPM ftp time server as Notes to the UTC - UTC(k) data. However, we have obtained the complete steering information from the participating institutes, as it was noticed that the list reported by BIPM was not always complete.

4. 2 Short description of TWSTFT and GPS data

The stations involved performed TWSTFT in co-ordination by the CCTF Working Group on TWSTFT. At each of the stations, named A and B for now, a characteristic PRN spread spectrum signal, synchronous with the local 1 PPS time references $T(\mathrm{Ref}_{A,B})$ is generated on a 70 MHz intermediate frequency. It is up-converted to the RF region (Ku-band, up link at 14 GHz) and transmitted to a geo-stationary telecom-munication satellite. This was the IS-706 until September 2003, and the IS-903, both of Intelsat, afterwards. At the satellite, the signal is translated to the downlink frequency (12 GHz) and sent back to earth. The time of arrival (TOA) of the PRN coded signal from station A (B) is determined at station B (A) by cross-correlation of the received and the down-converted PRN signal with a local replica of the same PRN code, synchronous with the local reference clock. TOA measurements were made each second during sessions of two minutes per station pair. The time differences exhibit a $1\,\sigma$ - standard deviation of typically 400 ps to 700 ps. At the end of a session, a quadratic fit is done to the data and the fitted TOA midpoint values (including the propagation delays) are exchanged. This allows calculation of the quantity of interest, $T(\mathrm{Ref}_A) - T(\mathrm{Ref}_B)$, as explained by Kirchner [7] and in an ITU-R Recommendation [8]. During most of the time covered by our study, TWSTFT sessions were performed nominally once per day or, more recently, four times per day. If appropriate, daily mean values were calculated for our study. The data sets contain very few gaps in the sense that no data were available for a day with MJD ending in 4 and 9. The gaps could have been closed by interpolation, but this was not done. To give an example, there are no TWSTFT data available for GSTf prediction at Modified Julian Dates (MJD) 52864 to 52874, so that such data points are missing in Figure 3.

The GPS data used were generated from C/A code L1 receivers at IEN and PTB. The equipment at PTB was upgraded during the study period, so that during the last 12 months about 80 GPS common view observations were recorded per day whereas this number was only 45 in the initial period. All individual observations were post-processed by BIPM using precise operational satellite ephemerides and ionosphere maps provided by the International GPS Service for Geodynamics (IGS) in the same way as all single frequency GPS links are treated by BIPM. In the future, one could use GPS data generated from dual-frequency geodetic receivers which are processed according to the TAI P3 scheme [9]. Such data could be available on a day-to-day basis as required for the application in mind, but have a similar quality as the post-processed data used here.

5. SIMULATION OF PREDICTION AND RESULTS

The period of simulation chosen was constrained to 550 days, between May, 2003 and October, 2004, by the availability of TWSTFT data from all the chosen institutes. The data available to us had been evaluated with PTB as the pivotal institute, i.e., the data were of the form $[UTC(PTB) - UTC(k)]^{TW}$. However, for the present study we needed to get the links to IEN. This was achieved with a simple transformation as follows:

$$[UTC(k) - GSTf]^{TW} = [UTC(PTB) - GSTf]^{TW} - [UTC(PTB) - UTC(k)]^{TW} \quad (1)$$

This transformation adds some extra noise which is, however, negligible compared to the instability of the time scales involved.

5. 1 Selecting the Right Prediction Strategy

The prediction algorithm, as described in Section 4.1, was applied to the Circular T data over the period of the study. The different prediction strategies adopted, namely, LV, Lin, and Linsq, lead to different results. This is illustrated in Fig.1 where we show the published UTC − GSTf data (connected dots) along with the predicted values with different strategies. The prediction errors are shown in Fig. 2 as the difference between the published UTC − GSTf data and the predicted values for the three different cases. Using these series of error values we could work out the standard deviation σ, the mean deviation, and the maximum and minimum deviations which we consider as useful statistical measures to enable us to select the right prediction strategy.

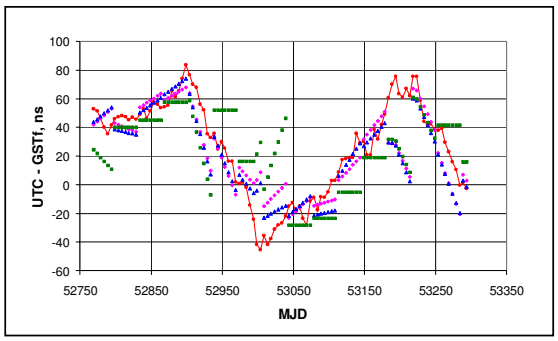


Fig. 1 Predictions of UTC - GSTf (● plus line) using 7-2-7 Lin (symbol ♦), Linsq (symbol), and LV strategies (symbol ■).

In Table 1 we show the prediction error statistics for each of the time scales for all the prediction strategies, LV, Lin, and Linsq, over the entire chosen period. LV had been found the optimum prediction strategy for UTC(PTB) in [6], but here LV gave higher values of σ for all three other time scales. The results have very little difference for the Lin or Linsq predictions. Based on these observations, we decided to uniformly adopt the 7-2-7 Lin prediction strategy for the present work.

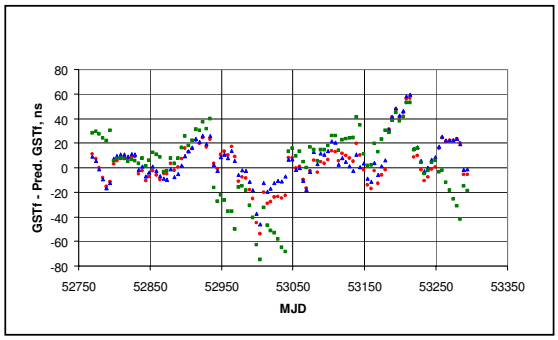


Fig. 2 Errors of prediction of UTC - GSTf, using Lin (symbol ●), Linsq (symbol), and LV strategies (symbol ■).

Table 1 Statistics of the errors with different prediction strategies

7-2-7	σ, ns	mean, ns	max diff, ns	min diff, ns
PTB LV	8.7	-0.3	24.6	-17.2
PTB Lin	12.7	3.0	37.1	-20.0
PTB Linsq	12.5	1.6	28.8	-25.3
NPL LV	15.8	0.7	33.5	-63.2
NPL Lin	8.2	5.3	22.2	-1.7
NPL Linsq	8.0	5.8	21.0	-2.9
VSL LV	28.2	-1.0	66.3	-68.2
VSL Lin	19.9	3.5	73.2	-33.3
VSL Linsq	20.5	4.0	79.6	-24.2
GSTf LV	27.9	4.4	53.0	-74.8
GSTf Lin	18.6	3.3	56.3	-53.9
GSTf Linsq	17.9	5.4	59.3	-46.2

5. 2 Simulated Prediction of GSTf including TWSTFT data

The prediction results for GSTf have already been shown in Figs. 1 and 2 and the statistics of the errors given in Table 1. Now we include steps 3) and 4), as given in the Introduction:

3) Measurement of GST − UTC(k) in quasi real time,
4) Combination of steps (2) and (3) to predict UTC − GST in real time.

The prediction of GSTf is now done via the UTC(k)-scale as an intermediate scale by using the time link between IEN and k using the following relation:

$$\text{Predicted}[UTC - GSTf] = \text{Predicted}[UTC - UTC(k)] + [UTC(k) - GSTf]^{TW \text{ or } GPS}. \quad (2)$$

Fig. 3 shows the results of these individual predictions of UTC − GSTf using TWSTFT data. The statistics of the errors are shown in Table 2. Significant scatter of

the individual predictions are visible in Fig.3 and the error statistics given in Table 2 contains still unacceptably high value of the standard deviations σ.

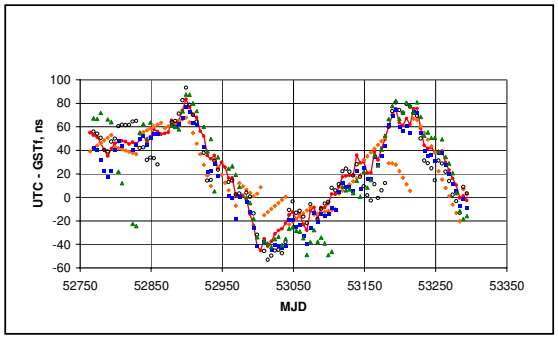


Fig. 3 Prediction of UTC - GSTf (● plus line) via UTC(PTB) (symbol O), UTC(NPL) (symbol ■), UTC(VSL) (symbol) and itself (symbol ♦) using TWSTFT data.

Table 2 Statistics of errors of GSTf predicted via TWSTFT to three timing institutes

	σ, ns	mean, ns	Max diff, ns	min diff, ns
GSTf Lin	18.6	3.3	56.3	-53.9
Predicted via PTB	12.9	3.3	37.1	-19.4
Predicted via NPL	8.6	6.7	22.3	-2.3
Predicted via VSL	18.3	3.8	70.0	-31.1

5. 3 Average Prediction

A simple way to improve the prediction accuracy is to take an average of the individual predictions. In case of the absence of predicted data points via one or more UTC labs (mainly due to the absence of TWSTFT data) we have taken the average of the remaining ones. The average prediction is shown in Fig. 4 where a smoothing effect is immediately apparent. There is one point at MJD= 53004 which is clearly way off. This happens to be a day when no TWSTFT data were available and so the average is equal to the direct prediction of GSTf.

The error statistics, the first line in Table 3, also reflects this single outlier and we get σ = 9.4 ns and mean = 3.9 ns. Using the data of the GPS link to PTB for this particular day removes this outlier, giving rise to line 2 in Table 3.

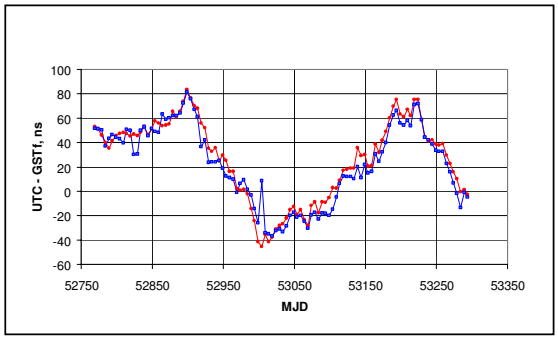


Fig. 4 Average prediction of UTC - GSTf (● plus line) via UTC(PTB), UTC(NPL), UTC(VSL) and itself using TWSTFT links (■ plus lines).

Table 3 Statistics of error of GSTf average predictions

	σ, ns	mean, ns	max diff, ns	min diff, ns
Average Prediction	9.4	3.5	19.5	-53.9
Aver. Prediction: Outlier removed	7.8	4.1	19.5	-15.5
Weighted t Aver. Prediction	8.9	4.3	19.5	-53.9
Weighted Aver. Prediction, outlier removed	7.4	4.6	19.5	-12.7

5. 4 Weighted Average Prediction

A further improvement can be expected when the individual predictions are combined with individual statistical weights, mentioned earlier as step 5):

5) Assessment of the quality of the prediction UTC – UTC(k) based on historical data and determination of a statistical weight in the combination of inputs in step (4).

It has been observed that the prediction accuracy is inversely related to the stability of the UTC scale used. To a first order we can define the weights W_k of UTC(k) as

$$W_k = \left(1/\sigma_k^2\right) \Big/ \sum_k \left(1/\sigma_k^2\right), \qquad (3)$$

where the variances σ_k are the instabilities of the different time scales as given in Table 2. We used these to determine a weighted average prediction of GSTf. This is shown in Fig. 5 and the corresponding statistics is given in lines 3 and 4 of Table 3. There is not a remarkable improvement compared to the case of simple average since the σ_k are not greatly different and also because of the number of participants in the average being rather few. In line 4 of Table 3 we give the result corrected for the one outlier at MJD 53004.

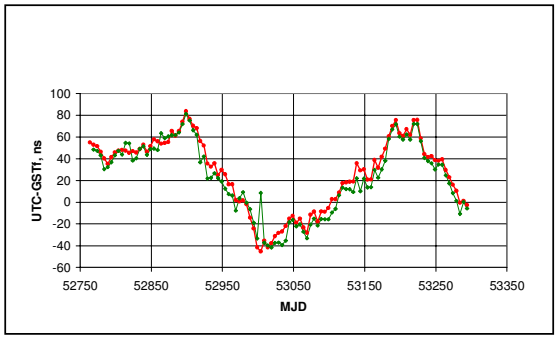


Fig. 5 Weighted average predictions of UTC - GSTf (● plus line) via UTC(PTB), UTC(NPL), UTC(VSL) and itself using TWSTFT links (symbol ♦ plus line).

5. 5 Prediction Using the GPS link

We have carried out the prediction simulations in the previous subsections using the TWSTFT time links. It is important to consider time links using GPS (or Galileo in the future) as an alternative. We have not repeated the entire simulation exercise by replacing the TWSTFT data with GPS data. Instead, we have just done the prediction of GSTf via PTB using the GPS link, as a representative case. The results are shown in Fig. 6 and the statistics are given in Table 4. These are very similar to the results shown in Table 2 using the TWSTFT link. It is clear then that GPS could be used as an alternate link to TWSTFT, or indeed as a supplementary link, to provide redundancy.

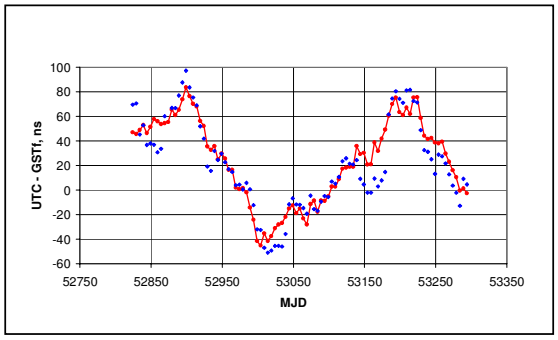


Fig. 6 Prediction of UTC - GSTf (● plus line) via PTB using GPS data (symbol +).

Table 4 Statistics of errors of GSTf predicted via PTB using GPS data

7-2-7	σ, ns	mean, ns	max diff, ns	min diff, ns
Predicted via PTB	13.2	3.4	34.5	-24.9

5. 6 Prediction of GSTf via USNO

UTC(USNO) is known as a stable time scale which could be used as an ultimate reference for prediction with high accuracy. TWSTFT data between USNO and PTB was available for our period of simulation. Using (1) we could then get the relation between UTC(USNO) and GSTf using TWSTFT. The GSTf predicted via USNO using the same scheme as before, namely, 7-2-7 Lin strategy, is shown in Fig.7. The statistics for the errors for this case, given in Table 5, are clearly indicative of the excellent predictability of UTC(USNO) and serve as a benchmark for what can be obtained once a superior UTC(k) time scale (or more than one), e. g. based on advanced primary clocks, would be available in European timing institutes.

Table 5 Statistics of errors of GSTf predicted via USNO using TWSTFT link

7-2-7 Lin	σ, ns	mean, ns	max diff, ns	min diff, ns
Predicted via USNO	4.6	0.3	7.8	-14.7

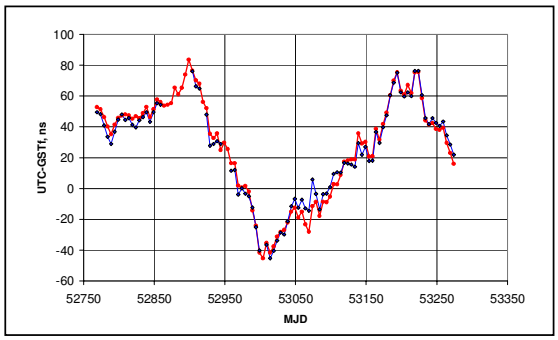


Fig. 7 Prediction of UTC - GSTf (red dots plus line) via UTC(USNO) using TWSTFT data (symbol ♦ plus line).

DISCUSSION

Among the tasks of a national timing institute “k” is the realization of its local UTC(k) time scale and its steering towards the international UTC. Additionally, it may be requested to predict the time difference UTC – UTC(k) for the period when this information is not yet available due to the production of UTC in deferred time. Usually, those tasks are performed autonomously, i. e., based on the data provided in the BIPM Circular T and internal records of the performance of the clocks involved in the generation of UTC(k). The prediction of UTC − GST where GST is the future Galileo System Time would constitute a similar activity. Assuming that the GST stability is not perfect due to limitations in the equipment used for its generation it may be wise to draw upon available UTC(k) time scales for improving the predictability. We have shown that indeed the prediction based on data available in real time in addition to the historical Circular T data may be helpful.

Our prediction algorithm requires only data which are readily available:

1. UTC – GST as to be published in Circular T,
2. UTC – UTC(k) published in Circular T,
3. Measurement of GST – UTC(k) in quasi real time,

so that the co-ordinating effort to obtain these data is minimized. We have shown in Section 5.3 and 5.4 that the access to the three time scales produced by NPL, PTB, and VSL would already be sufficient to fulfill the predictability requirement. In the meantime the institutes SYRTE in France (Observatoire de Paris), and METAS in Switzerland are participating in routine TWSTFT comparisons, both equipped with high quality atomic frequency standards. As we have shown in section 5.5 that also GPS common view data would be possibly sufficient for the purpose, it becomes clear that several more European institutes could in principle provide the required information based on GPS data only and would still increase the reliability of the dedicated timing service to be provided by Galileo.

In details, our study could have been refined in several points. One could select the optimum time scale prediction method for each participating time scale separately so that optimum use would be made of the information at hand. This selection could be repeated periodically, and the statistical weights assigned would have to be adapted accordingly.

We produced a time difference prediction only for one out of five days as this allowed immediate comparison with Circular T data. Of course, in real operations one point per day would be required.

We have still dismissed the question of how to limit the offset UTC − GSTf to well within ±50 ns. Actually the required frequency steering needs to be taken into account when making the predictions. It is considered a separate task to define an algorithm for this steering since a compromise must be found between small deviation from zero and increase of frequency instability of GST due to aggressive steering.

Last not least, we would like to point again to the fact that an ultimately stable reference could allow an almost perfect prediction, as is indicated by the results given in Section 5.6. It is very likely that a time scale based on an active hydrogen maser steered in frequency towards a cold-atom primary frequency standard could have a similar performance as UTC(USNO). Already today several European institutes (IEN, METAS, NPL, SYRTE, PTB) are close to having the required equipment operational which would be needed to generate such a time scale.

ACKNOWLEDGEMENT

We gratefully acknowledge the support offered by IEN, NPL, and VSL in providing the required UTC(k) steering corrections, and the consent to participate in this study. Gerard Petit provided the GPS data used. Our study evolved from discussions with Alexandre Moudrak and Johannes Furthner of DLR (Deutsches Zentrum für Luft und Raumfahrt, Institute for Communications and Navigation, Oberpfaffenhofen) about the best practice of generating GST. The work was partially funded by Galileo Joint Undertaking (GJU), Bruxelles, in the frame of a contract given jointly to IEN, NPL, and PTB.

REFERENCES

[1] S. Bedrich, A. Bauch , J. Laverty, A. Moudrak, W. Schäfer, Design of the Galileo Precise Time Facility, Proc. 18th European Frequency and Time Forum, 2004, paper 109.

[2] J. Laverty, “The Galileo Time Service Provider: A progress report”, to be presented at EFTF 2005, paper 84.

[3] A. Moudrak, “Galileo Time Service Provider - The focal point for European timing expertise”, to be presented at EFTF 2005, paper 46.

[4] F. Cordara et al., The Galileo System test Bed V1: First Experimental Results, Proc. 18th European Frequency and Time Forum, paper 111.

[5] F. Cordara, Galileo system test Bed V1: Results on the Experimental Galileo System Time, EFTF 2005, paper 42.

[6] A. Bauch, D. Piester, E. Staliuniene, A new realization strategy for the time scale UTC(PTB), Proc. 2004 IEEE International Ultrasonics, Ferroelectrics, and Frequency Control 50th Anniversary Joint Conference (paper P1FC_I_7), pp. 518-523.

[7] D. Kirchner, Two-way Satellite Time and Frequency Transfer (TWSTFT): Principle, Implementation, and Current Performance, Review of Radio Sciences 1996-1999, Oxford University Press, 1999, pp. 27-44.

[8] Recommendation ITU-R TF.1153-1 “The operational use of two-way satellite time and frequency transfer employing PN time codes”, ITU, Radiocommunication Study Group, Geneva, last update 2003.

[9] P. Defraigne and G. Petit, "Time transfer to TAI using geodetic receivers", *Metrologia* **40** (2003), pp. 184-188.

EVALUATION OF INFLUENCE DUE TO THE OPERATIONAL DEAD TIME OF AN ATOMIC FOUNTAIN FREQUENCY STANDARD AT NMIJ

S. Yanagimachi, T. Ikegami, and T. Kurosu

National Metrology Institute of Japan (NMIJ)/AIST, AIST Tsukuba Central 3, Ibaraki 305-8563, Japan

Keywords: hydrogen maser, atomic fountain, dead time

Abstract

National Metrology Institute of Japan has two commercial hydrogen masers (HM), one of which is used to link an atomic fountain frequency standard to International Atomic Time. The other is used to monitor a frequency stability of the HMs. If we operate the atomic fountain frequency standard for a day of five days scheduled by BIPM, the uncertainty due to dead time is estimated to be 2.1×10^{-15}.

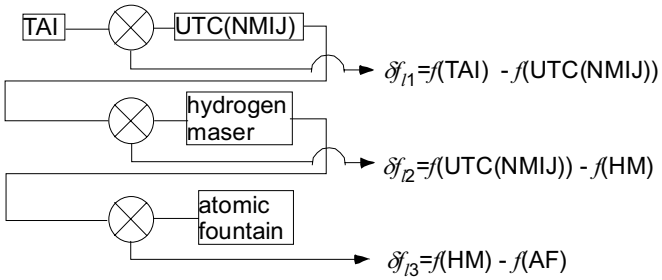
1 Introduction

For the purpose of keeping the International Atomic Time(TAI) with smaller deviation from the SI second, it is necessary that calibration data of TAI using primary frequency standards are reported to BIPM. Presently, at National Metrology Institute of Japan (NMIJ), the primary frequency standard employing an atomic fountain method show considerably small uncertainty, which reaches low level of the order of 10^{-15} [1]. When we compare the frequency of the atomic fountain frequency standard (AFFS) and that of TAI, we have to estimate the uncertainty due to the operational dead time of AFFS, since it is not easy to operate it continuously for five days at the moment because of its complexity. Currently, National Institute of Standard and Technology (NIST) and Bureau National de Metrologie-Systems de Reference Temps-Espace (BNM-SYRTE) published their own method to evaluate the uncertainty caused by AFFS's dead time [2,3]. In the present work, referring to these two reports, we estimate the statistical uncertainty due to the operational dead time of AFFS at NMIJ.

2 Link between TAI and AFFS

The time and frequency link to TAI consists of three comparisons in NMIJ, as shown in Fig. 1. The first link is a remote comparison using GPS and TWSTFT between TAI and UTC(NMIJ) which is kept by Cs atomic clocks. This is managed by BIPM and the comparison results are reported every five days regularly. As a result, one can obtain the frequency difference over a time interval of 5 days between TAI and UTC(NMIJ), δf_{l1}. The uncertainty due to this link is originated from the noise of the link itself, which is estimated to be 1×10^{-15}/month by BIPM. The second link is between UTC(NMIJ) and a hydrogen maser (HM) which is used as a flywheel. This comparison is carried out continuously by phase comparison, which provides the frequency difference over a time interval of 5 days between UTC(NMIJ) and HM, δf_{l2}. The noise level of this comparison is 1.5×10^{-15}/day and negligible. The third link gives the frequency difference over a time interval of 5 days between HM and AFFS, δf_{l3}. If the AFFS is operated continuously in the scheduled five days, the noise level of this link is also negligible like the second link. Finally, one can obtain the frequency difference over a time interval of 5days between TAI and AFFS, δf, as

$$\delta f = \delta f_{l1} + \delta f_{l2} + \delta f_{l3} . \quad (1)$$



AFFS; atomic fountain frequency standard HM; hydrogen maser
UTC; coordinated universal time
TAI; international atomic time

Figure 1. Block diagram of link from the atomic fountain frequency standard to International Atomic Time at NMIJ

3 Hydrogen masers

NMIJ uses two commercial HMs (RH401A; Anritsu Corp.) not only as a flywheel oscillator of AFFS but also as a reference oscillator for evaluation of the frequency stability and accuracy of AFFS. The better frequency stability HM has, the more effective the accuracy evaluation is, and moreover, the smaller uncertainty due to the operational dead time is. From these reasons, it is important that HMs show as good performance as possible. For this purpose, two HMs are set up at a room where the temperature and humidity is controlled precisely (Fig. 2). The controlled temperature is 23 °C ± 0.5 °C (RMS), and the controlled humidity is 51 % ± 1 %(RMS). The 5 MHz and 10 MHz signals from HMs is



Figure 2. Two commercial hydrogen masers

provided through a distribution amplifier (HPDA-15RM/B; Spectra Dynamics Inc.).The phases of signals from AFFS and two HMs are compared by using two vector volt meters (Hewlett-Packard; 8508A). These results are stored in a personal computer. The phase comparison between two HMs is carried out continuously in order to observe the possible frequency change and measure the frequency stability. Furthermore, the time deviations during two HMs and UTC(NMIJ) are compared through 1 pps signals by the time interval counter (Hewlett-Packard; 5370B) once a day, the results of which are reported to BIPM once a month. Figure 3 shows the frequency difference over a time interval of 1 day between two HMs, which is obtained from the data of the time deviations for approximately two month (2004/8/11-10/7, MJD 53228-53285). The observed frequency fluctuation in peak-to-peak is approximately 1×10^{-14}, which is preferable. On the basis of the observed data, the Allan standard deviation $\sigma_y(\tau_{day})$ is well described as a function of $\tau_{day}^{-1/2}$ that is characteristic of white frequency noise, and thus the following relation

$$\sigma_y\left(\tau_{day}\right) = 2.3\times10^{-15}\times\frac{1}{\sqrt{\tau_{day}}} \qquad (2)$$

is obtained, as is shown in Fig. 4, where τ_{day} is expressed in 1 day unit.

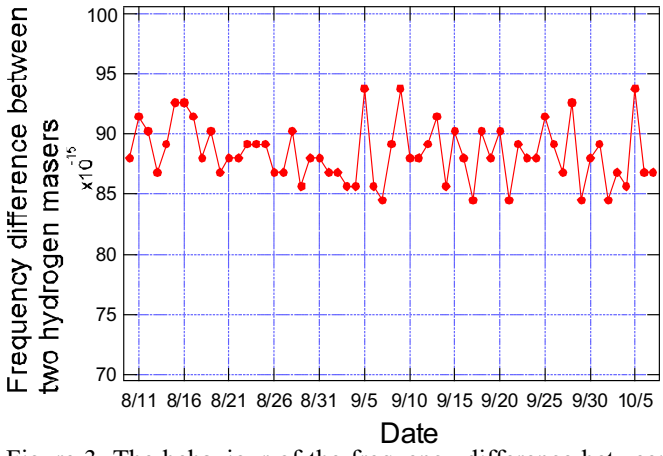


Figure 3. The behaviour of the frequency difference between two HMs during two months.

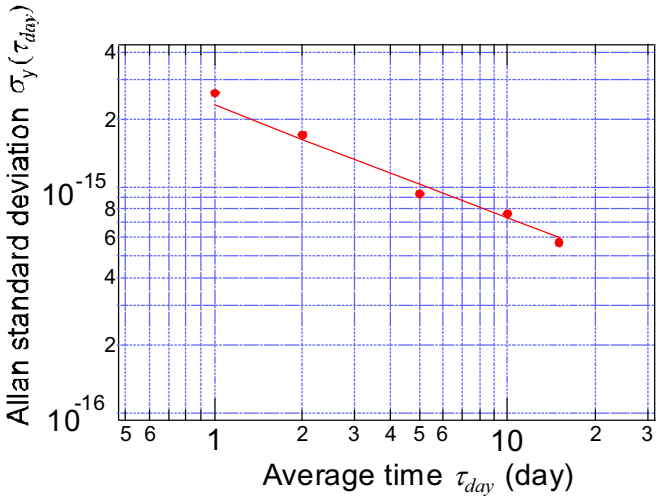


Figure 4. Allan standard deviation obtained from comparison data between two HMs.

4 Uncertainty due to the operational dead time of an atomic fountain frequency standard

In this section, we first describe the statistical expression to estimate the uncertainty due to the operational dead time of AFFS. Next, we apply it for the observed data shown in Fig. 3 and estimate the uncertainty due to the dead time. For this purpose a few definitions and assumptions are introduced in the following.

I. $f_i\left(\mathrm{TAI}\right)$ and $f_i\left(\mathrm{HM}\right)$ represent the frequency over a time interval of one day in TAI and HM respectively, where subscription *i* that ranges between 1 and 5 means the *i*th day in the scheduled 5 days.

II. $f\left(\mathrm{AFFS}\right)$ is defined as the frequency over a time interval of one day in AFFS. It is assumed that this frequency does not show a statistical noise, and is constant for five days. The statistical noise is included by Type A uncertainty of AFFS.

III. It is assumed that the first and second links are so tight that the frequency difference, δf_{l1}, δf_{l2}, are given without

uncertainty due to the dead time. The uncertainty of the first link is investigated and announced by BIPM. The uncertainty of the second link is clear, as is already described in Sec. 2. Therefore, the frequency difference over a time interval of 5 days between TAI and HM, $\delta f_{l12,}$ is obtained from the following equation,

$$\delta f_{l12} = \delta f_{l1} + \delta f_{l2} = \frac{1}{5}\sum_{i=1}^{5}\left\{ f_i\left(\mathrm{TAI}\right) - f_i\left(\mathrm{HM}\right)\right\}. \quad (3)$$

When AFFS is operated for five days continuously, the frequency difference over a time interval of 5 days, δf_{l3}, is expressed by

$$\delta f_{l3} = \frac{1}{5}\sum_{i=1}^{5}\left\{ f_i\left(\mathrm{HM}\right) - f\left(\mathrm{AFFS}\right)\right\}. \quad (4)$$

From Eq. (3) and (4), the frequency difference over a time interval of 5 days between TAI and AFFS, δf, without the dead time is expressed as

$$\delta f = \delta f_{l12} + \delta f_{l3} = \frac{1}{5}\sum_{i=1}^{5}\left\{ f_i\left(\mathrm{TAI}\right) - f\left(\mathrm{AFFS}\right)\right\} \quad (5)$$

This is idealistic case, and then δf is independent of the frequencies of HMs.

On the other hand, when the AFFS is operated on the 1st day of five days, the frequency difference between HM and AFFS, δf_{l3}', is given by

$$\begin{aligned}\delta f_{l3}' &= f_1\left(\mathrm{HM}\right) - f\left(\mathrm{AFFS}\right) \\ &= \frac{1}{5}\sum_{i=1}^{5}\left\{ f_1\left(\mathrm{HM}\right) - f\left(\mathrm{AFFS}\right)\right\}.\end{aligned} \quad (6)$$

By using δf_{l3}', the frequency difference with the dead time, δf', is derived as follows,

$$\begin{aligned}\delta f' &= \delta f_{l12} + \delta f_{l3}' \\ &= \frac{1}{5}\sum_{i=1}^{5}\left\{ f_i\left(\mathrm{TAI}\right) - f_i\left(\mathrm{HM}\right) + f_1\left(\mathrm{HM}\right) - f\left(\mathrm{AFFS}\right)\right\}.\end{aligned} \quad (7)$$

The frequency difference with the dead time, δf', is rewritten by using δf in Eq. (5) as

$$\delta f' = \delta f - \frac{1}{5}\sum_{i=1}^{5}\left\{ f_i\left(\mathrm{HM}\right) - f_1\left(\mathrm{HM}\right)\right\}. \quad (8)$$

The second term in Eq. (8) gives the error in the frequency difference with the dead time δf'. This error is originated from the fact that the frequency over a time interval of one day for i=1 does not correspond to the frequency over a time interval of 5 days in HM. As a result, this error is dependent on the frequency stability of HM that is used as flywheel, and gives the uncertainty due to the dead time in the third link. Equation (8) was derived on the basis of the assumption that only 1st day is live time: live time means the time interval during which AFFS is compared with HM. When AFFS is operated only on the 2nd, 3rd, 4th or 5th day, Eq. (8) is easily modified for appropriate equations.

Here, we estimate the frequency fluctuation of the HM, which results in the error of the frequency difference with the dead time, δf', by using the frequency difference between two independent HMs. One of example for the estimation is shown in Tab. 1, which shows the frequency difference between two HMs from 53249 to 53278(MJD). The first column represents index of the ith day of the scheduled five days. The forth column shows the calculated frequency difference from that of the first day (i=1). The sixth column shows the calculated value followed by the second term in Eq. (8). When we calculated RMS of 6 samples shown in sixth column, we obtained 2.0×10^{-15}.

By using this method, we calculated the RMS of the second term in Eq. (8), assuming that AFFS is operated only on the 1st, 2nd, 3rd, 4th or 5th day of five days, where all of the data of frequency difference between two HMs (MJD 53228-53285) shown in Fig. 3 is used. The calculated result is 2.1×10^{-15}, which shows uncertainty due to the operational dead time, when AFFS is operated for a day of the five days.

We can confirm a validity of this method by comparing to the method used in NIST, which is derived by T. E. Parker et, al [2]. From Ref. [2], once the frequency stability of the flywheel $\sigma_{y}(T)$ is given, the uncertainty $\sigma_d(T,\tau)$ due to the operational dead time is determined by the following equation,

$$\sigma_d\left(T,\tau\right) = \sqrt{\left(\frac{T}{\tau} - 1\right)}\sigma_y\left(T\right). \quad (9)$$

In the Eq. (9), T is measurement interval, and τ is live time in the third link. These notations are the same as those used in Ref. [2]. When we use $\sigma_{y}(T) = 2.3\times10^{-15}\times T^{-1/2}$ of Eq. (2), T=5 days, τ=1day, $\sigma_d(T=5,\tau=1)$ is calculated to be 2.1×10^{-15}, from Eq. (9), which is in good agreement with our estimated value. Figure 5 shows the uncertainty $\sigma_d(T,\tau=1)$ as a function of measurement interval T, together with the value obtained from the statistical method described here. The values from two methods have good correspondence. This is because the frequency fluctuations of HMs in this experiment are approximately expressed by the white frequency noise. When the flicker-frequency or random-walk frequency are dominant for 1 < T < from 10 to 20 days, we have to use other equations of $\sigma_d(T,\tau)$ specialized for such types of noise.

5 Conclusions

We reported the present status of the time and frequency link from our AFFS to TAI. Currently, we can estimate the uncertainty due to the operational dead time of AFFS by observing the frequency fluctuation between two commercial HMs whose noises show white frequency noise. The uncertainty due to the dead time of AFFS is estimated to be 2.1×10^{-15}, when AFFS is operated for one day of the scheduled five days.

i	MJD	DATE	Frequency difference between two HMs	f_i(HM)-f_1(HM)	$\Sigma\{f_i(\mathrm{HM})-f_1(\mathrm{HM})\}/5$
1	53249	2004/9/1	8.68E-14	0.00E+00	9.26E-16
2	53250	2004/9/2	8.68E-14	0.00E+00	
3	53251	2004/9/3	8.56E-14	-1.16E-15	
4	53252	2004/9/4	8.56E-14	-1.16E-15	
5	53253	2004/9/5	9.38E-14	6.94E-15	
1	53254	2004/9/6	8.56E-14	0.00E+00	2.5E-15
2	53255	2004/9/7	8.45E-14	-1.16E-15	
3	53256	2004/9/8	8.91E-14	3.47E-15	
4	53257	2004/9/9	9.38E-14	8.10E-15	
5	53258	2004/9/10	8.80E-14	2.31E-15	
1	53259	2004/9/11	8.80E-14	0.00E+00	9.E-16
2	53260	2004/9/12	8.91E-14	1.16E-15	
3	53261	2004/9/13	9.14E-14	3.47E-15	
4	53262	2004/9/14	8.56E-14	-2.31E-15	
5	53263	2004/9/15	9.03E-14	2.31E-15	
1	53264	2004/9/16	8.80E-14	0.00E+00	2.E-16
2	53265	2004/9/17	8.45E-14	-3.47E-15	
3	53266	2004/9/18	9.03E-14	2.31E-15	
4	53267	2004/9/19	8.80E-14	0.00E+00	
5	53268	2004/9/20	9.03E-14	2.31E-15	
1	53269	2004/9/21	8.45E-14	0.00E+00	3.7E-15
2	53270	2004/9/22	8.91E-14	4.63E-15	
3	53271	2004/9/23	8.80E-14	3.47E-15	
4	53272	2004/9/24	8.80E-14	3.47E-15	
5	53273	2004/9/25	9.14E-14	6.94E-15	
1	53274	2004/9/26	8.91E-14	0.00E+00	-9.E-16
2	53275	2004/9/27	8.68E-14	-2.31E-15	
3	53276	2004/9/28	9.26E-14	3.47E-15	
4	53277	2004/9/29	8.45E-14	-4.63E-15	
5	53278	2004/9/30	8.80E-14	-1.16E-15	

$\sqrt{[\Sigma\{f_i(\mathrm{HM})-f_1(\mathrm{HM})\}/5]^2} = 2.0\times10^{-15}$

Table 1. Example for statistical estimation of the uncertainty due to the dead time of the third link

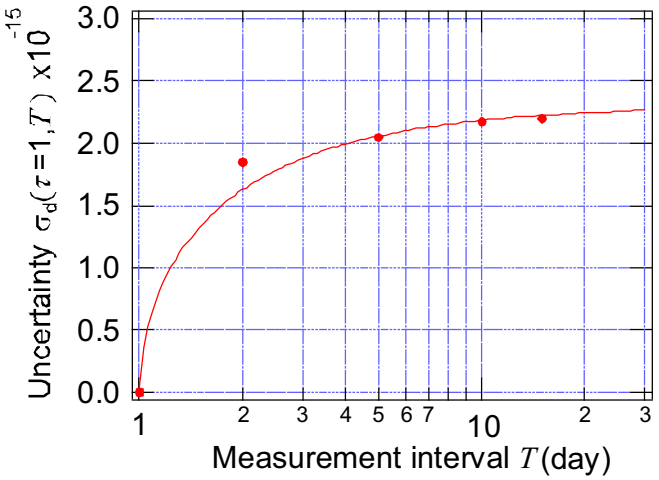


Figure 5. Uncertainty from Eq. (8) (full line); Uncertainty from statistical method (plot)

References

[1] T. Kurosu, Y. Fukuyama, K. Abe, S. Yanagimachi and Y. Koga, "Evaluation of the Cs atomic fountain frequency standard at NMIJ/AIST", in *Proc. 2003 IEEE International Frequency Control Symposium*. pp. 123-126, (2003).

[2] T.E. Parker, D. A. Howe and M. Weiss, "Accurate frequency comparison at the 1x10^{-15} level", in *Proc. 1998 IEEE International Frequency Control Symposium*, pp. 265-272, (1998).

[3] SYRTE-FO2_53109-53149[7 June 2004], Reports of evaluation of Primary Frequency Standard, available at http://www.bipm.org/en/scientific/tai/time_ftp_data.html

DEVELOPMENT OF TIME MANAGEMENT SYSTEM FOR SATELLITE POSITIONING SYSTEM USING QUASI-ZENITH SATELLITES

Y. Takahashi*, J. Amagai*, M. Fujieda*, S. Yokota*, K. Kimura*, H. Ito*, S. Hama*, T. Morikawa*, I. Kawano, S. Kogure****

*National Institute of Information and Communications Technology (NICT)
4-2-1, Nukuikita-machi, Koganei, Tokyo 184-8795 Japan,
E-mail: yasu@nict.go.jp
** Japan Aerospace Exploration Agency (JAXA)
2-1-1 Sengen, Tsukuba-shi, Ibaraki, 305-8505 Japan

Keywords: Quasi-zenith satellite (QZS), satellite positioning System, Atomic clock, Time management System, Time transfer, TWSTFT

Abstract

The first Japanese Quasi-Zenith satellite (QZS) will be launched FY 2008. The QZS System (QZSS) is developed as an integrated satellite service system for communications, broadcasting, and positioning for mobile users in throughout specified regions in Japan from a high angle of elevation. The purpose of the satellite positioning system using QZS is to complement and augment the GPS. The national institutes concerned have been developing the positioning system since 2003 and intend to carry out experiments and research three years after the launch.

National Institute of Information and Communications Technology (NICT) is currently developing a precise time management system of the QZSS, such as the space-borne hydrogen maser (SHM) atomic clock, the space-borne time transfer subsystem, and same ground facilities for the QZSS. As the orbit-determination accuracy of the satellite with this system is required to be about 1m, it is necessary to synchronize atomic clocks both satellites and ground stations in the sub ns range. Time management stations for the QZSS will be kept by the UTC(NICT). The concept design for the time management system for the QZSS was has been completed and we are developing engineering models for this system.

1. Introduction

The Japanese government and a Japanese company (Advanced Space Business Corporation; ASBC) are planning to launch three QZSs around 2008 at inclined. The QZSS is developed as an integrated satellite service system for communications, broadcasting, and positioning for mobile users in throughout specified regions of Japan from a high angle of elevation. The purposes of the satellite positioning system using QZS is to complement and augment the GPS [1]. The national institutes concerned have been developing the positioning system since 2003 and intend to carry out experiments and research three years after the launch.

The NICT and the JAXA already developed the on-board high-precision time comparison equipment [2] for Engineering Test Satellite VIII (ETS-8) which is planned to be launched in FY 2005. The NICT has also conducted Two-Way Satellite Time and Frequency Transfer (TWSTFT) using a geostationary satellite [3] and developed a new MODEM for it. The time management system for QZSS is based on those experiences. The NICT is mainly in charge of developing a timing system for the satellite positioning system using QZS, such as the on-board hydrogen maser atomic clock and the precise time management system of the QZSS. This paper introduces time management system for satellite positioning system

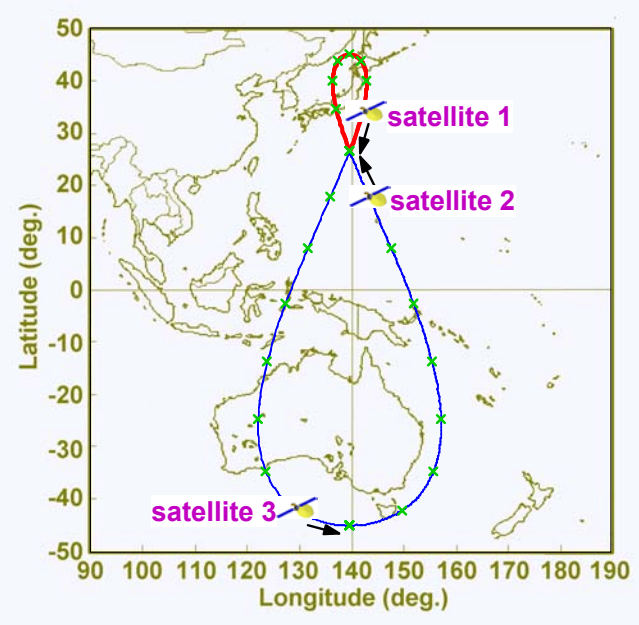


Figure 1. Example of the footprint of the three QZSs. (asymmetrical figure "8")
eccentricity = 0.1, inclination = 45 degree

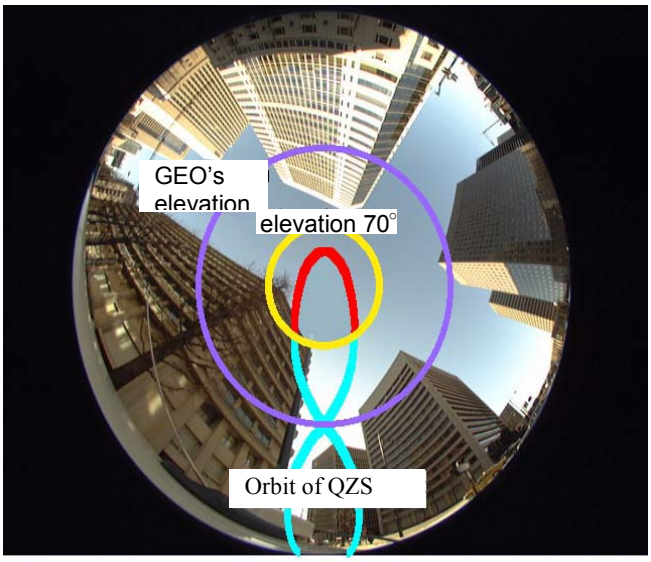


Figure 2. Sample Photograph of orbit of QZS viewed from ground in Tokyo

using the QZSS.

2. QZSS

The idea for a satellite using quasi-zenith orbit was proposed by NICT in 1972 [4]. This is an elliptical orbit with an inclination of 45 degrees to provide a high elevation to users located in mid-latitude regions like a Japan. The project to study the use of satellite using quasi-zenith orbit has been proposed by the Federation of Economic Organizations.

The Japanese government and a Japanese company ASBC began to study the QZSS from FY2003, and was planning to launch the first QZS around 2008 to achieve inclined quasi-zenith orbit. Figure 1 shows an example of the footprint of the satellites [5]. In that case, the minimum elevation reaches 78 degrees in Tokyo, for example. It is, therefore, significant merit to both communications and positioning for users located in mid-latitude. The QZSS is developed as an integrated satellite service system for communications, broadcasting, and positioning for mobile users in specified regions of Japan from a high angle of elevation. The purpose of the satellite positioning system using QZS is to complement and augment the GPS. The experiments of the satellite positioning system using QZS will be conducted by the institutes concerned for three years.

Figure 2 is a sample photograph of the orbit of the QZS viewed from the ground in Tokyo. At least one of two or more satellites located within the outer circle of buildings in the figure can service roads to an elevation of 70 degrees or more.

3. Positioning System

Figure 3 is an overview of the positioning system of t he QZSS. As the on-board system, ASBC is in charge of manufacturing the bus system and the communications service part of the on-board equipment, while JAXA and NICT are in charge of

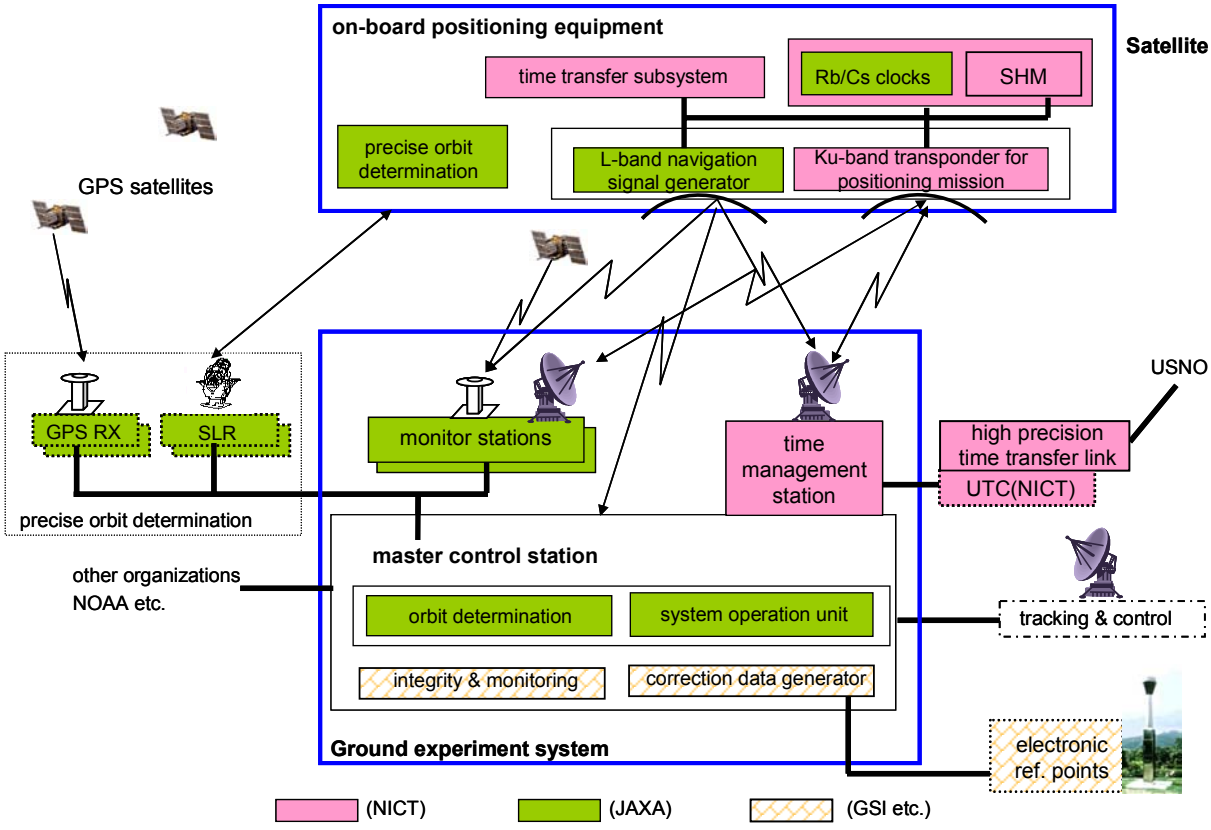


Figure 3. Overview of positioning system of QZSS

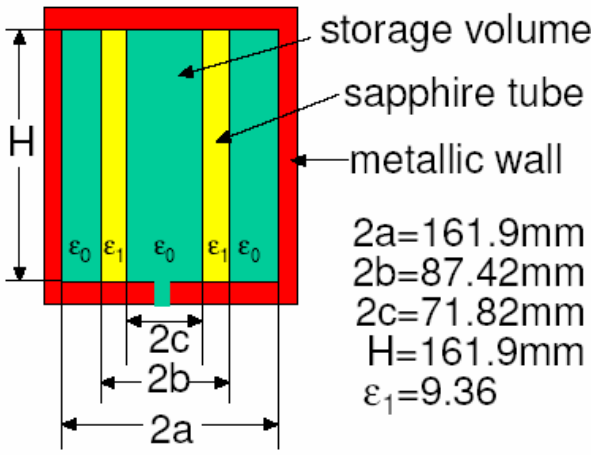


Figure 4. Structure of the sapphire loaded cavity.

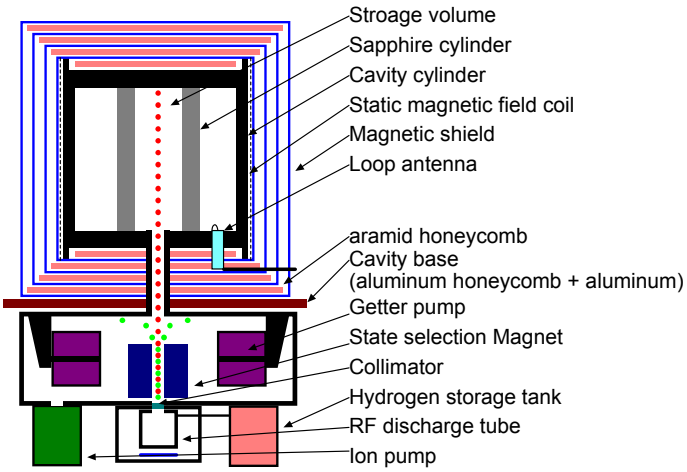


Figure 5. BBM structure of SHM.

manufacturing the positioning system.

As the ground system, ASBC is in charge of manufacturing the feeder link station and satellite control station. JAXA, NICT, and the other institutes are in charge of manufacturing the master control stations (MCS) and the Time Management Stations (TMS) for the main and sub-stations, and monitoring stations that use the positioning system. Each organization is in charge of other systems as well.

4. Space-borne Hydrogen Maser

A Bread Board Model of space-borne hydrogen maser (SHM) for the satellite positioning system using Quasi-Zenith Satellite System has been developed. Analysis and experiments for technical problems such as miniaturization, achievement of long life, mechanical vibration proof, and space environment characteristics have been performed and technical data necessary for EM development have been obtained.

The most signicant factor determining the dimensions of the hydrogen maser is size of the cavity resonator. If the cavity can be reduced in size, the vacuum enclosure and the magnetic shields that surround the cavity may in turn be made smaller, and the entire hydrogen maser can be made to be small and lightweight. In this point of view, we have developed a sapphire-loaded cavity, which is shown in Figure 4. It is 161.9mm in diameter and 161.9mm high, and its volume is reduced by 1/7 compared to a conventional maser cavity that is about 300mm in diameter and 300mm high.

In order to miniaturize the BBM, the metallic cylinder of the cavity is also used as the vacuum enclosure in the new design. The new design minimizes the use of outgassing materials in vacuum, which results in improved lifetime of SHM. With this new design, the weight of BBM, shown in Figure 5, is reduced from 72kg to 52.6kg, compared to prototype model.

5 Time Management System

5.1 Structure of Time Management System

Figure 6 has an overview of the proposed time management system, which is within the bold line. The positioning system of QZS has:

- Three on-board atomic clocks (two Cs and/or Rb and a hydrogen maser),
- L-band transmitters for positioning signals (L1, L2, L5, and an original signal compatible with modernization GPS),
- Ku-band transmitters and receivers for time transfer, and
- A precise time comparison unit (TCU).

Ground facilities for the positioning system using QZS are composed of:

- Two MCS to control the whole positioning system,
- Monitoring stations to determine the satellite orbit (approximately ten stations both nationally and internationally), and
- Two TMS for time management.

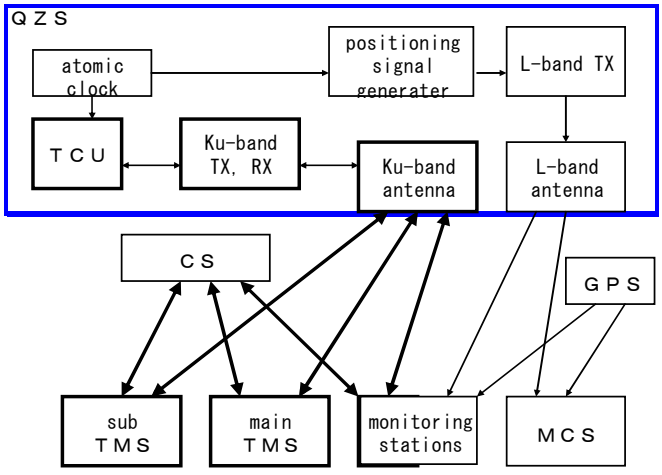


Figure 6. Overview of proposed time management system.

5.2 Differences and Interoperability with GPS

A characteristic difference between the QZSS and the GPS is its orbit, which is elliptical with an eccentricity of about 0.1. Its altitude is also different, which means the rotation period and the velocity around the earth also differ. A lower altitude and a larger velocity produce delay due to the relativistic effect. The average difference can be compensated by shifting the frequency of the on-board clock beforehand. However, the eccentricity causes large delay at the perigee and small delay at the apogee. The change in delay experienced by a QZS in one day is much larger than that of GPS. The effect to the positioning caused by the delay change is expected to be compensated by receivers.

The reference clock in the TMS will be maintained precisely and continuously by referring to UTC (NICT), which is linked to UTC (USNO), the basis of GPS time, by using TWSTFT method. Consequently, the QZSS time is expected to be maintained to the GPS time with a precision of about 3 ns rms over any 24-hour period. Providing precise time comparison with an independent method helps precise orbit determination.

NICT discussed the concept of timing compatibility for GPS and QZSS with USNO and GPS JPO. QZSS is to adopt the same method of counting seconds as GPS. QZSS will therefore have the same numbering for seconds as GPS (without adding a leap second). The GPS to QZSS time offset will be provided to users by coefficients that have been included in the QZSS navigation message to support compatibility between GPS and QZSS. Each QZS clock will be steered from the TMS.

5.3 Necessary precision of time keeping

As the accuracy of determining the orbit of the QZS is required to be about one meter, it is necessary to synchronize the on-board atomic clocks and ground stations within 1ns. Consequently, a TWSTFT method is used for these time comparison. Time management stations for the QZSS will be kept through UTC (NICT). The TMS compares many atomic clocks and these are:

- Time comparison between on-board atomic clocks and those of the TMS,
- Time comparison between monitoring-station atomic clocks and the TMS, and
- Time comparison between main-TMS and the sub-TMS atomic clocks.

The on-board Time Transfer Subsystem (TTS) is installed on the QZS, the main functions of which are:

- Time comparison between the satellite and ground stations,
- Time comparison between the satellite's on-board clocks, and
- Transponder for time comparison between ground stations

6. Functions of On-Board Equipment

Figure 7 is a block diagram of the on-board equipment used for the positioning system. The TTS's range is within the chain line, and is used for highly accurate time transfer between ground stations, and between ground stations and the satellite. It will have the following functions and equipment, which can be unified to achieve its other purposes. The TTS has

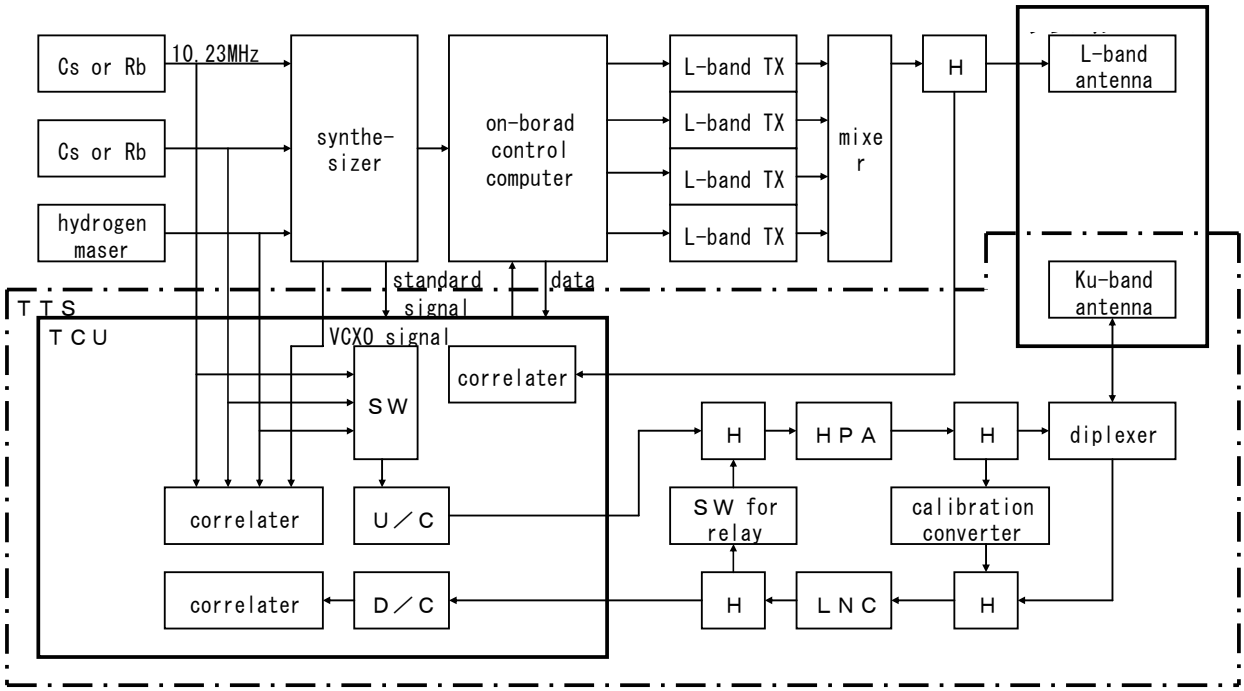


Figure 7. Block diagram of on-board equipment for positioning system

functions to precisely compare the time. This is achieved as follows:

- Time comparison based on a standard signal from the on-board atomic clock,
- Time comparison between the ground station and satellite: Transmit and receive the time transfer signal through the Ku-band transmitting and receiving equipment and the Ku-band antenna, and compare the ground station signal at the TCU,
- Time transfer between ground stations: Receive, convert the frequency, and transmit the time transfer signal from the ground stations,
- Phase difference measurement function of the three atomic clocks and VCXO signal: Compare the four signals at the TCU.
- Delay time measurement function of Ku-band transmitting and receiving equipment: Convert the frequency of the transmission signal, return to the TCU through the Ku-band receiving equipment, and measure its signal at the TCU,
- Delay time measurement function of L-band positioning signals: Measure the signals from the L-band transmission equipment at the TCU,
- Data transmission and reception: Superimpose the telemetry for the positioning system on the time transfer signal, and transmit at the Ku-band transmitting equipment, and
- Data interface function: Send the measurement results of the TCU to the on-board control computer.

Moreover, the L-band positioning signal, which JAXA is in charge of, selects one of three atomic clocks and uses the VCXO output in the synthesizer as an on-board standard signal for positioning equipment. The L-band positioning signal is scheduled to transmit L1, L2, L5, and an interchangeable original signal for modernization GPS.

7. Method of Time Comparison

The positioning system should be able to measure the difference between all these clocks. This system will establish atomic clocks in the QZS, main TMS, sub TMS, and all monitoring stations. Therefore, the TMS compares the time with all the other stations, and the time differences are obtained for these. The method by which time their times are compared is described in what follows.

7.1 Time comparison between satellite and ground stations

The comparison of the TMS of the clock and the time on the satellite need to be highly accurate because we plan to do the positioning within an accuracy of about 1 m. The time comparison needs to be done in sub-ns. Therefore, we planned on using a method of TWSTFT when the comparison signals for time were sent to one another.

7.2 Time comparison between ground stations

The time comparison between ground stations occurs between the main TMS and the sub TMS, and between the TMS and monitoring station. The main TMS will be located at NICT's Koganei Headquarters, because the main TMS can use UTC(NICT) for standard signal. We plan to place the sub TMS on Okinawa as the QZS can be seen around the clock. Unfortunately, the satellite cannot be seen from Tokyo for the entire 24 hours.

A time comparison between the main TMS and the sub TMS is needed, at the very least, because the sub TMS controls in time that the main TMS cannot be seen for the QZS. The QZS does not necessarily need to be used, and time between ground stations can be compared with the existing geostationary Communications Satellite (CS). We plan to use the CS because it enables easy operation and simplifies construction of the earth station. However, there is the possibility of placing a monitoring station where the CS cannot be used. In this case, the QZS will be used.

7.3 Time comparison between on-board clocks

There will be three on-board atomic clocks, and the standard signal to generate the positioning signal will use VCXO output based on these clocks. We have no good policies to compare time between the four signals and ground clock, because it consumes too many satellite resources. The time comparison between the ground station and satellite is assumed to be done based on VCXO output and as many as four on-board clock signals based on VCXO output at the TCU.

8. Conclusions

The research and development for this positioning system began in the 2003 fiscal year. It was necessary to develop a plan to launch a satellite within a relatively short period of five years. Consequently, we needed to efficiently undertake this project in cooperation with many institutes. The TTS, which is NICT's installation equipment, underwent its conceptual design last year. It is scheduled for completion of its basic design in the current fiscal year when manufacturing of the engineering model will begin.

Part of this work was done with the support of Ministry of Internal Affairs and Communications, Japan.

References

[1] I. Kawano, et. al. "Japanese Experimental GPS Augmentation Using Quasi-Zenith Satellite System (QZSS)", ION GNSS 2004, California, 2004.

[2] Y. Takahashi, “Development of Time Comparison Equipment for ETS-VIII Satellite”, 2004 Conference on Precision Electromagnetic Measurements, London, 232-233, 2004.
[3] M. Imae, et. al. “Two-Way Satellite Time and Frequency Transfer Networks in Pacific Rim Region”, IEEE Trans. Instrumentation and Measurement, 50(2), 559-562, 2001.
[4] K. Takahashi, “Suitable Orbits for Various Satellite Missions”, Review of RRL, 18(97), pp.345-353, 1972 (in Japanese).
[5] K. Kimura, et al. “Optimum constellations using elliptical inclined synchronous orbits for mobile and fixed satellite communications”, 50th IAC, Amsterdam, IAF-99-A.2, 1999.

TWSTFT PLANS IN THE PACIFIC-RIM REGION AND EUROPE LINK

Hideo MAENO[1], Miho FUJIEDA[1], Masanori AIDA[1], Lam Quoc Tung[1], Ryo TABUCHI[1]
Fumimaru NAKAGAWA[1], Tomonari SUZUYAMA[2] and Yukio Takahashi[1]

[1]*National Institute of Information and Communications Technology*
4-2-1 Nukuikita-machi, Koganei, Tokyo 184-8795 JAPAN
maeno@nict.go.jp, TEL +81 42 327 6096, FAX. +81 42 327 6689
[2]*National Institute of Advanced Industrial and Technology*
1-1-1 umezono, Tsukuba, Ibaragi 305-8586 JAPAN

ABSTRACT

The National Institute of Information and Communications Technology (NICT) has been constructing the precise time transfer network using Two-Way Time and Frequency Transfer (TWSTFT) technique in collaboration with the major time and frequency institutes in the Pacific Rim region.

The TWSTFT techniques have been used as the primary time transfer link for the TAI calculation of the links between NICT-NMIJ, NICT-NTSC, NICT-TL, and NICT-AUS (NML) during the period between January 2002 and September 2002. In addition to above experiments, time-transfer has been regularly performed also between NICT-USNO (Vandenberg) twice a week. This network is being expanded to include KRISS and SPRING. We will implement dense or continuous time transfer using TWSTFT method in this region.

NICT is preparing a compact transportable TWSTFT station for calibration of TWSTFT in the Pacific Rim region. The transportable station will be available soon. We will conduct the domestic evaluation experiments this year and start calibration trip to other institutes next year. Furthermore, we are planning to extend the TWSTFT link between NICT-PTB via PAS-4 satellite covering Asia/India-Europe. The antenna installation is scheduled in this June. This plan using multi-channel modem is positioned as a very significant step to connect the Pacific-Rim region and Europe.

We will present the current status and future plan of the TWSTFT network in the Pacific Rim region and Europe link.

1. INTRODUCTION

The National Institute of Information and Communications Technology (NICT) developed the new Two-Way Satellite Time and Frequency Transfer (TWSTFT) modem, and we call it "NICT modem". It can be operated automatically with the remote by multi-channels mode. We have conducted test observations using the NICT modems in cooperation with many cooperators. This modem has been already installed in NICT, NMIJ, NTSC, TL, KRISS, SPRING, NMI and NICT's two LF stations. We are continuing an evaluation experiment of its performances using actual satellite link and this is in an end stage.

2. TWSTFT STATUS IN THE PACIFIC-RIM AND EUROPE

We have conducted the routine and test observations using Atlantis and NICT modem. The network of TWSTFT is shown in figure1. Figure1 is shown Network of TWSTFT in Asian Pacific Rim and Europe stations. The used modems are two types. One is the Atlantis modem which is used for the routine observations of TWSTFT in the Asia and Pacific Rim. The participated stations are NICT and NMIJ in Japan, NTSC in China, TL in Taiwan, NMI in Australia and Vandenberg (USNO) in USA. Another is the NICT modem which is developed by NICT. This modem is available for multi-point simultaneous time transfers. The test observations of the modem are conducted among NICT, NMIJ, NTSC, TL, NMI, KRISS and SPRING. Now we have

checked the performance and the quality of data, and we have updated the system in present. We will install the NICT modem at PTB in Germany until next June and we will start the observations between NICT, KRISS and PTB in Europe.

The two satellites, JCSAT and PAS-8 are used in the both observations, and the other satellite PAS-4 will be used in the future for the test experiments between NICT, KRISS and PTB.

The routine observations are conducted on two days (Tuesday and Friday) every week, and three times of 5 minutes for each day by Atlantis Modem. The NICT modem routine observations are started on two days (Tuesday and Friday) every week, and four times of 5 minutes for each day from 1st February. The test observations are conducted for seven days every week. The NICT modem is operated in automatic mode by the remote control from NICT. Table 1 shows the modems and satellites in this region.

NICT has three sites for the TWSTFT in Japan. One is placed at the headquarters of NICT and others are placed at the two LF broadcasting stations. The location of these three stations is plotted in figure 2. The LF stations have Cs clocks with normal performance tube, and we use UTC(NICT) for the reference signal at the headquarters of NICT. These stations use PAS-8.

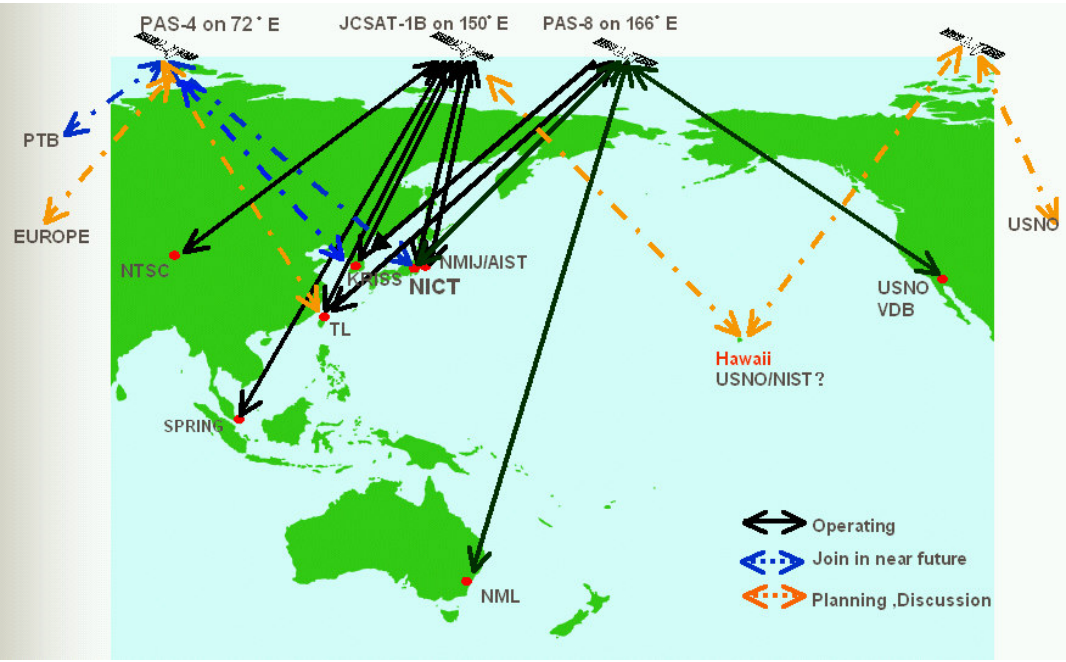


Figure 1 Network of TWSTFT in Asian and Pacific Rim

Table 1 The routine and test observation

station	Routine (Atlantis Modem)	Routine and Test (NICT modem)	Satellite
NICT(Japan)	o	o	JCSAT-1B and PAS-8
NMIJ(Japan)	o	o	JCSAT-1B
NMI(Australia)	o	o	PAS-8
NTSC(China)	o	o	JCSAT-1B
TL(Taiwan)	o	o	JCSAT-1B and PAS-8
KRISS(Korea)	×	o	JCSAT-1B and PAS-8
SPRING(Shingapole)	×	o	JCSAT-1B
VDB(USNO, USA)	o	×	PAS-8

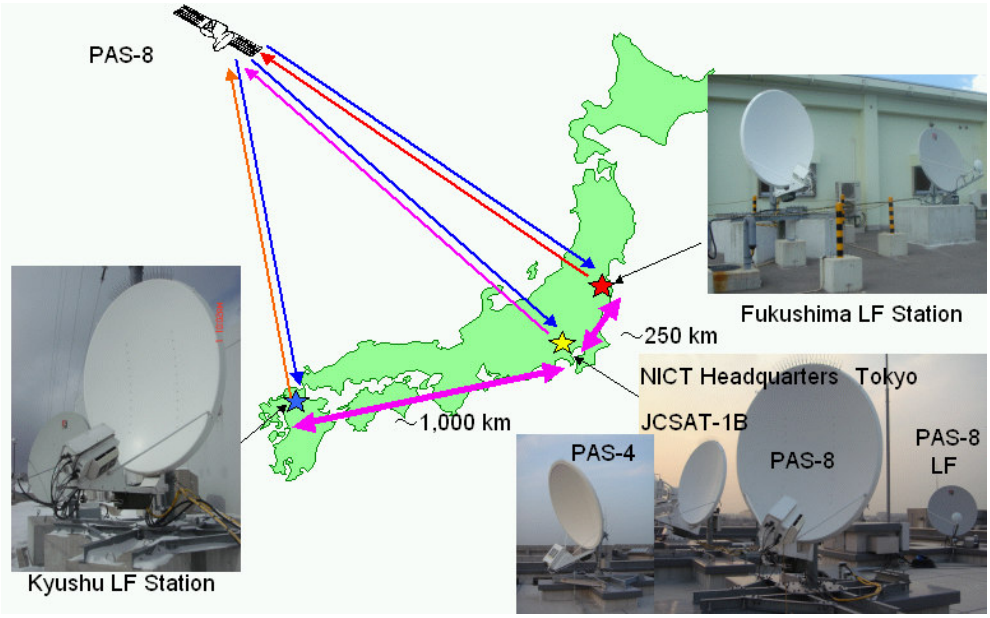


Figure 2 domestic LF stations network

3. BASIC SPECIFICATIONS OF THE NICT MODEM

The basic specification of the NICT modem's transmission part, receiving part, and data exchanging part are summarized and shown in figure 3, table 1, 2, and 3 respectively. As the channel 2 of the transmission part and channel 8 of the receiving part are used for the delay variation measurement described, so the maximum number of the stations which can participate at the same time is eight.

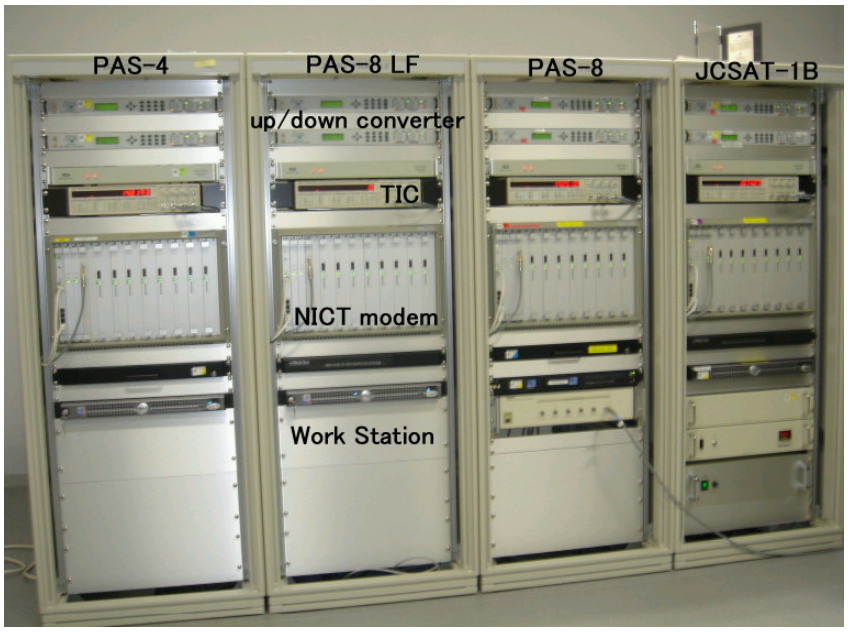


Figure 3 Overview of the TWSTFT systems

4. TEST OBSERVATION RESULTS OF NICT MODEM AND COMPARED WITH ATLANTIS MODEM AND GPS CV

4.1 NICT modem

The NICT modem was developed by NICT and we have conducted the test experiments using the NICT modem in cooperation with many institutes. We continue to update the system. We fixed some problems of firmware and hardware of the system. Therefore, we started the routine observation from 1st February.

Recently, we have uploaded the NICT observation results on our Website to report them to BIPM. In the near future, we will calculate offset values for these data to complete the ITU standard format.

4.2 Improvement

1) The data acquisition program was fixed in May, 2004. Besides the constant offset, the results of the NICT modem have been consistent with those by the Atlantis and GPS CV since May, 2004.

2) SPRING's LNA has broken. We are going to replace it soon in this year.

4.3 Data results

We show the stability of the TWSTFT in three months using NICT modem.

1) NICT-TL

Figure 4 shows test observation results of NICT modem and compared with Atlantis modem on NICT-TL by JCSAT link. The program of data acquisition was fixed in May, 2004. The relative results of the NICT modem are consistent with those of the Atlantis and GPS CV from May, 2004. There was a phase jump in the NICT results via JCSAT at the beginning of July. The cause is still under examination.

2) NICT-NML

Figure 5 shows similar results with the NICT NML by PAS-8 link.

The data were well correlated between NICT modem data compared with Atlantis modem and GPS CV.

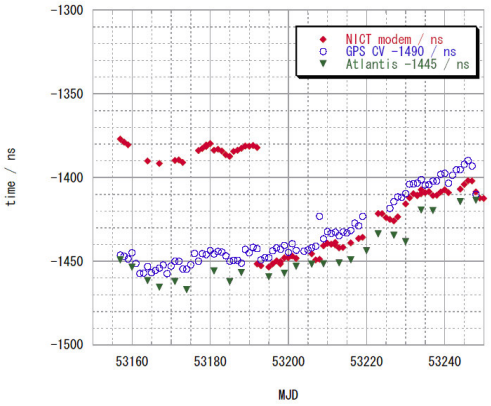


Figure 4 NICT-TL by JCSAT link

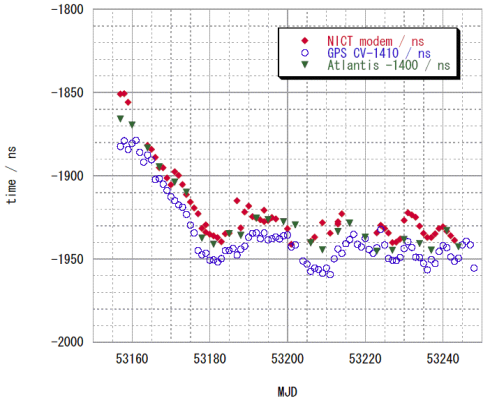


Figure 5 NICT-NML by PAS-8 link

Table 1. Basic Specifications of Transmission part

Number of channel	2 one for time transfer one for calibration
Center frequency	70.000MHz ± 500 kHz changeable by 625Hz step
PN generator	12 stages FSR 4095 bits/cycle
Drive frequency	2.0475 MHz synchronized to External 10 MHz signal
Period of PN code	2.000 ms
Output level	-20 dBm € -40 dBm
Data modulation	500 bps (measured data from 8 Rx channels)

Table 2. Basic Specifications of Receiving part

Number of channel	8 seven for time transfer one for calibration
Center frequency	70.000MHz ± 500 kHz
PN generator	12 stages FSR 4095 bits/cycle
Drive frequency	2.0475 MHz
Period of PN code	2.000 ms
Minimum operational C/No	+45 dB€ Hz
PN code sync. time	<20 sec. (after frequency search)

Table 3. Basic Specifications of Exchanging Data among participating stations

DATA rate	500 bps
Contents	
Preamble	24 bits
Epoch of measurement	
MJD	20 bits (5 digits in BCD code)
Total second	20 bits (5 digits in BCD code)
Tx ID	4 bits
Measured data	x 8
Rx ID	4 bits
Measured time interval data	48 bits (12 digits in BCD code)
CRC	16 bits

5. FUTURE PLAN

We will make an establishment of the calibration and transportable portable system. The domestic calibration observations will be started soon and after that we will perform the calibration observations with the collaboration institutes from next year. NICT, KRISS and PTB (Germany) are constructing the preparations for PAS-4 link by the NICT modem in cooperation. We hope that experiment on the PAS-4 link would be done by May or June, 2005.

6. CONCLUSION

The TWSTFT is very important to compare the standard time and frequency precisely and to contribute it to the international time and frequency standard, such as UTC and TAI. We continue to perform the TWSTFT observations in the Asian and Pacific Rim and to develop the TWSTFT techniques. We have obtained highly precise NICT modem data and these data are reported to BIPM. The data play an important role of the maintenance for the time and frequency standard and of the human life information.

STEERING A TIME SCALE BY THE APPLICATION OF THE THEORY OF CONTROL

Marcello Farina[♯], Lorenzo Galleani[†], Patrizia Tavella[*]

[♯] Dipartimento di Elettronica e Informazione, Politecnico di Milano, Via Ponzio 34, 20133 Milan, Italy
farina@elet.polimi.it
[†] Dipartimento di Elettronica, Politecnico di Torino, C.so Duca degli Abruzzi, 24, Torino, Italy
[*] Istituto Elettrotecnico Nazionale (IEN) "G. Ferraris", Torino, Italy

ABSTRACT

The goal of this paper is to investigate clock steering methods in the framework of control theory. The current steering techniques have been developed using different approaches. Our aim is to apply control theory methods to the analysis of the most common steering strategies.
The first issue to be studied is the evaluation of the clock models, and the laws that govern the control signal and the controlled one. A correct modelling of the control electronic device (actuator), which is typically a microstepper, allows to obtain a suitable model of the controlled process.
The obtained results are presented, they are very encouraging and suggest further investigation of these problems.

1. INTRODUCTION

The aim of this paper is to "translate" the problem of steering a time scale into a classical control problem. This means that we have to characterize all the metrological quantities involved in steering in a formal language that is more familiar to control engineers.

The first step is to analyze, under the "Control Theory" point of view, the properties, the validity, the performances and the theoretical limits of the most common clock steering methods. We observed that, looking through the "Control Theory lens", each of the steering methods currently applied can be seen as an application of different classical or more advanced control methods.

Many steering methodologies have been proposed so far to solve control problems ([1],[2],[3]), each of them with different static and dynamic performances, robustness, noise rejection. Hence there is a wide choice of methods to meet the desired metrological specifications.

We will discuss the basic steps required for the design of a control system (see, e.g. [4]), then we will examine the currently used steering techniques to identify the different control solutions and in some cases, such as the optimal control technique proposed by Koppang ([5],[2],[6]), Control Theory is directly and explicitly involved.

2. DESIGNING A CONTROL SYSTEM

The first step of a Control System design process is to analyze the *process* from a very practical point of view. We call a *process* any problem that can be *modeled* as a mathematical system and that can be controlled to meet a given set of requirements. In this step, the aim is to deduce, in an intuitive way, the nature of the process itself, and the *signals and properties* of the process. The value of some of these signals or properties can be directly influenced by a device called *actuator*. These signals are the *inputs* ($u(t)$) of the system. Moreover, we can find signals or properties that related to the state of the process and that can be directly measured by a device called *transducer*. These signals or properties are called the *outputs* of the process ($y(t)$). Then we have the *state variables* ($x(t)$), that are used to describe the inner state of the process. These can be both measurable and non-measurable quantities.

After identifying all the features that characterize the process under analysis, the second step is to create a *mathematical model* of such process. In this step, the aim is to define the dynamic or algebraic relationships between input, state variable output values, and the dependency of these variables on time. These relationships are generally expressed with algebraic and differential equations. Also the effect of the transducer and the actuator can be modeled through mathematical equations. The result can be either an *algebraic* or a *dynamic system* of equations, with either *time-dependent* or *time-independent* correlations between variables. The equations can be either *linear* or *non-linear*. If we take noise into account the model is *stochastic*, and deterministic otherwise. Moreover, if we use data that are continuous in time the system will be said to be a *continuous-time-model*. Otherwise, if the available data are discrete in time (i.e. taken every day at midnight, or every second), the system is called a *discrete-time-model*, and will involve finite-difference equations rather than differential equations.

In general, a linear time-independent (LTI) stochastic system is expressed by

$$\begin{cases} \dot{\mathbf{x}}(t) \equiv \frac{d}{dt}\mathbf{x}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{B}\,\mathbf{u}(t) + \mathbf{w}(t) \\ \mathbf{z}(t) = \mathbf{C}\,\mathbf{x}(t) + \mathbf{D}\,\mathbf{u}(t) + \mathbf{e}(t) \end{cases} \quad (1)$$

where the vector $\mathbf{x}(t) = [x_1(t), ..., x_n(t)]'$ represents the time dependent values of the n state variables, the vector $\mathbf{z}(t) = [z_1(t), ..., z_r(t)]'$ represents the time dependent values of the r outputs, the vector $\mathbf{u}(t) = [u_1(t), ..., u_m(t)]'$ represents the time dependent values of the m inputs. Moreover, the vector $\mathbf{w}(t) = [w_1(t), ..., w_n(t)]'$ represents the time-dependent values of the n stochastic variables modeling the *process noise* and the vector $\mathbf{e}(t) = [e_1(t), ..., e_r(t)]'$ represents the time-dependent values of the r stochastic variables modeling the *measurement noise*. Finally, the matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ express the dynamic and algebraic relations between these variables. Deterministic models are expressed by the same formulation where the vectors $\mathbf{w}(t)$ and $\mathbf{e}(t)$ are set to zero. Superscript $'$ denotes transposition.

The next step consists on the definition of the performance required to the Controlled System, that can be formulated by mathematical formula as performance indicators. We can use dynamic and static indicators, or indicators that are related to the control methodologies that will allow to design the proper regulator for the process (such as cost functions for optimization methods). In particular cases, such as the problem that we study in this paper, performance indicators already exist. For example, in steering problems, a common performance indicator for the output clock signal is the *Allan Variance* [7].

The next step of a Control System design process is the choice of the *Control Methodology* that we can apply to the specific problem. This choice is in part related to the kind of performance indicators that the specifications impose. For example, if they are expressed in terms of *bandwidth of the system*, or dynamic system response to standard exogenous signals, it is easier to apply a classical frequency domain control design methodology (i.e. with Bode diagrams). If there is instead a quadratic cost function to minimize, the natural choice is to apply optimal (or sub-optimal) methods (like, for example, Optimal Control [8] or Model Predictive Control [9]). Otherwise, sometimes choices are imposed by noise or perturbations on the system (uncertainties on the system model, for example). In the second case it is better to use a robust control method (like $\mathbb{RH}_\infty$ control [4]). We observe that many control methods have been proposed, each of them having different advantages depending on the problem they are applied to.

The last step of the Control System Design process is the design of a controller that can meet the required performances.

One can evaluate objectively the Control System's performances in terms, for example, of universal indicators, that allow to compare different control systems. We will analyze some well known steering methods to show in practice how to go through the above steps and design a control system.

3. THE CLOCK PROCESS

It is very important to start by defining the problem of steering as a control problem, and to define what the main features of the *process* are, mainly in terms of its inputs, outputs, and so on.

The simplest problem that we aim to solve is, in general, the problem of controlling the signal of a single clock, given a reference signal. In particular, in this paper, we will analyze two different cases. The first problem is a problem of steering of a Cesium clock to a reference time-scale with superior performances than the Cesium itself (in terms of Allan Deviation, for all time intervals τ). The second case represents the steering of a particular clock to another one. The reference clock, in this case, does not necessarily have always better performances than the steered clock. For example, if our aim is to steer a Hydrogen Maser to a Cesium Clock, we have to take into account that, for small τ, the behavior of the H Maser is better (in terms of Allan Variance) than the behavior of the Cesium Clock. On the contrary, for larger τ, the behavior of the Cs Clock is better. The aim for the H Maser is to keep the good performances for small τ and to gain the excellent Cesium-like behavior for larger τ.

The quantity that we can measure about the process is the phase difference between the two clock signals. This quantity represents the *output* of the *process*. The *input* signal is imposed by an external device. Depending on the particular device that we take into account (microphase stepper, frequency stepper, ...), the input signal will influence the output in different ways. The relation between input and output will be modeled by the system equations, which will differ according to the particular device that one uses. A simple scheme of the process is given in Figure 1.

4. MATHEMATICAL MODELS

We briefly show an overview of the clock models that can be found in literature ([10] and [11]). Moreover, we also discuss about the scope, the function and the possible mathematical models of *actuators*, and compare them to the ones that already exist in literature.

4.1. Modeling the absolute clock signal

The continuous-time LTI differential system describing the model of the absolute phase deviation of a clock is given by

$$\dot{\mathbf{x}}(t) = \mathbf{A}\,\mathbf{x}(t) + \mathbf{w}(t) \qquad (2)$$

where:

$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}, A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{w}(t) = \begin{bmatrix} w_1(t) \\ w_2(t) \\ w_3(t) \end{bmatrix}$$

For a 3-state model see [10], [11]. The components $x_1(t)$, $x_2(t)$, $x_3(t)$ represent respectively the

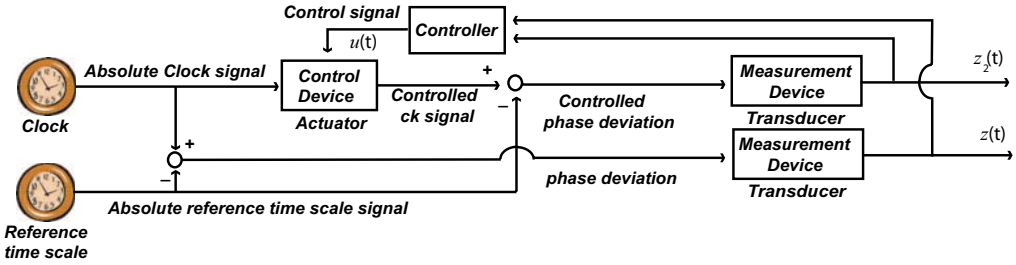


Figure 1: The controlled process

absolute phase deviation, a component of the frequency deviation and a component of the frequency drift of the clock signal. The stochastic processes $w_1(t)$, $w_2(t)$, $w_3(t)$ are independent white Gaussian noises with zero mean and variances σ_1^2, σ_2^2, σ_3^2 driving the White Frequency Noise (WFN), Random Walk FN and Random Run, respectively. Note that $\mathbf{w(t)}$ is also indicated as $\frac{d\mathbf{W}(t)}{dt}$, where $\mathbf{W}(t)$ is a vector of Wiener processes.

The discrete-time stochastic finite difference equations corresponding to the system in (2) are

$$\mathbf{x}(k+1) = \Phi_\tau \mathbf{x}(k) + \mathbf{w}^*(k) \tag{3}$$

with

$$\Phi_\tau = \begin{bmatrix} 1 & \tau & \tau^2/2 \\ 0 & 1 & \tau \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{w}^*(t) = \begin{bmatrix} w_1^*(t) \\ w_2^*(t) \\ w_3^*(t) \end{bmatrix}$$

The noises in $\mathbf{w}^*(t)$ are still white and Gaussian, but now exhibit correlation. They have zero mean and covariance matrix Σ.

$$\Sigma = \begin{bmatrix} \sigma_1^2\tau + \frac{\sigma_2^2}{2}\tau^3 + \frac{\sigma_3^2}{20}\tau^5 & \frac{\sigma_2^2}{2}\tau^2 + \frac{\sigma_3^2}{8}\tau^4 & \frac{\sigma_3^2}{6}\tau^3 \\ \frac{\sigma_2^2}{2}\tau^2 + \frac{\sigma_3^2}{8}\tau^4 & \sigma_2^2\tau + \frac{\sigma_3^2}{3}\tau^3 & \frac{\sigma_3^2}{2}\tau^2 \\ \frac{\sigma_3^2}{6}\tau^3 & \frac{\sigma_3^2}{2}\tau^2 & \sigma_3^2\tau \end{bmatrix} \tag{4}$$

4.2. Modeling the phase difference between two "free" clocks

If we have two different clocks (or time scales), supposed to be independent, we can call $x_1^i(t), x_2^i(t), x_3^i(t)$ $(i = 1, 2)$ the absolute phase deviation, the frequency deviation and the frequency drift of the $i-th$ clock. If we call $z(t)$ the output, that is the phase difference between the two clocks, we have:

$$z(t) = x_1^1(t) - x_1^2(t) + e(t) \tag{5}$$

where $e(t)$ is a stochastic process that represents the *measurement noise*, which we suppose to be a white noise with zero mean and variance σ_e^2 that represents the effect of the *transducer* on the output. As for the model, we can write:

$$\begin{cases} \begin{bmatrix} \dot{\mathbf{x}}^1(t) \\ \dot{\mathbf{x}}^2(t) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{A} \end{bmatrix} \begin{bmatrix} \mathbf{x}^1(t) \\ \mathbf{x}^2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{w}^1(t) \\ \mathbf{w}^2(t) \end{bmatrix} \\ z(t) = [1\ 0\ 0\ -1\ 0\ 0] \begin{bmatrix} \mathbf{x}^1(t) \\ \mathbf{x}^2(t) \end{bmatrix} + e(t) \end{cases} \tag{6}$$

Analogously, we can write the model in a different way. If the stochastic processes acting on the two clocks signals are independent, then we can write the system of clock differences as:

$$\begin{cases} \dot{\mathbf{x}}^z(t) = \mathbf{A}\mathbf{x}^z(t) + \mathbf{w}^z(t) \\ z(t) = \mathbf{C}\mathbf{x}^z(t) + e(t) \end{cases} \tag{7}$$

where:

$$\mathbf{x}^z(t) = \begin{bmatrix} x_1^z(t) \\ x_2^z(t) \\ x_3^z(t) \end{bmatrix} = \begin{bmatrix} x_1^1(t) - x_1^2(t) \\ x_2^1(t) - x_2^2(t) \\ x_3^1(t) - x_3^2(t) \end{bmatrix}$$

$$\mathbf{w}^z(t) = \begin{bmatrix} w_1^z(t) \\ w_2^z(t) \\ w_3^z(t) \end{bmatrix} = \begin{bmatrix} w_1^1(t) - w_1^2(t) \\ w_2^1(t) - w_2^2(t) \\ w_3^1(t) - w_3^2(t) \end{bmatrix}$$

$$\mathbf{C} = [1\ 0\ 0]$$

where $x_1^z(t)$, $x_2^z(t)$, $x_3^z(t)$ represent the phase deviation, frequency deviation and drift of the difference between the clocks signals, $w_1^z(t)$, $w_2^z(t)$, $w_3^z(t)$ are white Gaussian noises with zero mean and variances $\sigma_{1,ck1}^2 + \sigma_{1,ck2}^2$, $\sigma_{2,ck1}^2 + \sigma_{2,ck2}^2$, $\sigma_{3,ck1}^2 + \sigma_{3,ck2}^2$ respectively, and covariance matrix analogous to (4) and Φ_τ is as in 3.

We can write the discrete-time model through the discretization of system (7):

$$\begin{cases} \mathbf{x}^z(k+1) = \Phi_\tau \mathbf{x}^z(k) + \mathbf{w}^{z*}(k) \\ z(k) = \mathbf{C}\mathbf{x}^z(k) + e^*(k) \end{cases} \tag{8}$$

4.3. Modeling the clocks and the actuator

In this section we first show the model of a common actuator, the *phase micro-stepper* (PMS). We evaluate how this device influences the output of the system. Then we show the model of a different actuator described by Koppang ([5], [2], [6]) and then discuss the one implicitly used in the BANG-BANG steering ([3]).

In continuous-time, the PMS has the effect of adding a (normalized) frequency step Δf at time t to the signal of the controlled clock 1. So $u(t) = \Delta f$ at time t. In the discrete-time system, the control signal will be constant during the sampling interval τ, so $u(t) = u(k)$, $k\tau < t < (k+1)\tau$.

Given the free uncontrolled clock frequency deviation from the reference, $y_f(t)$, and the controlled clock frequency deviation from the reference $y_c(t)$, the effect of the PMS is $y_c(t) = y_f(t) + u(t)$. Let's be $x_1^{z,u}(t)$ the state variable representing the phase deviation of the controlled clock from the reference, then $y_f(t) = \dot{x}_1^z(t)$ and $y_c(t) = \dot{x}_1^{z,u}(t)$. So:

$$\dot{x}_1^{z,u}(t) = \dot{x}_1^z(t) + u(t)$$

The system (7), in case of a controlled clock becomes:

$$\begin{cases} \dot{\mathbf{x}}^{z,u}(t) &= \mathbf{A}\,\mathbf{x}^{z,u}(t) + \mathbf{B}\,u(t) + \mathbf{w}^z(t) \\ z_2(t) &= \mathbf{C}\,\mathbf{x}^{z,u}(t) + e_2(t) \end{cases} \tag{9}$$

where $e_2(t)$ represents a white Gaussian measurement noise with zero mean and variance σ_{ec}^2 and:

$$\mathbf{x}^{z,u}(t) = \begin{bmatrix} x_1^{z,u}(t) \\ x_2^{z,u}(t) \\ x_3^{z,u}(t) \end{bmatrix} \quad , \mathbf{B} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Note that in this particular case $x_2^{z,u}(t)$ and $x_3^{z,u}(t)$ are not controllable, and so $x_2^{z,u}(t) = x_2^z(t)$ and $x_3^{z,u}(t) = x_3^z(t)$. Finally, for the discrete-time model we have, from (8):

$$\begin{cases} \mathbf{x}^{z,u}(k+1) &= \Phi_\tau\,\mathbf{x}^{z,u}(k) + \mathbf{B}\,\tau\,u(k) + \mathbf{w}^{z*}(k) \\ z_2(k) &= \mathbf{C}\,\mathbf{x}^{z,u}(k) + e_2^*(k) \end{cases} \tag{10}$$

System (10) is similar to the model that is used in the Swiss implementation ([1], [12]).

Regarding system (10), we observe that Koppang ([2], [6]) formalized the problem through the equations

$$\begin{cases} \mathbf{x}^{z,u}(k+1) &= \Phi_\tau\,\mathbf{x}^z(k) + \mathbf{B}_K\,u(k) + \mathbf{w}^{z*}(k) \\ z_2(k) &= \mathbf{C}\,\mathbf{x}^{z,u}(k) + e_2^*(k) \end{cases} \tag{11}$$

where

$$\mathbf{B}_K = \begin{bmatrix} \tau \\ 1 \\ 0 \end{bmatrix}$$

In addition, in [2] and [6] a 2nd order model is considered, which means that $w_3^z(t)$ is equal to zero, or $\sigma_{3,ck1} = \sigma_{3,ck2} = 0$.

The actuator modeled in the BANG-BANG (see [3]) algorithm influences the clock reading by adding a step to the phase "acceleration" of the signal. This can be mathematically translated into system (9), where now the input matrix $\mathbf{B}'$ is:

$$\mathbf{B}' = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

This, in terms of a discrete-time model similar to (11), makes the discrete-time input matrix $\mathbf{B}_{BB}$ such that:

$$\mathbf{B}_{BB} = \int_0^\tau \Phi_{\tau-v}\,\mathbf{B}'\,dv = \begin{bmatrix} \tau^2/2 \\ \tau \\ 0 \end{bmatrix} \tag{12}$$

5. PERFORMANCE INDICES

Intuitively, the aims of control are mainly two:

. Deterministic rate elimination

. Noise reduction

A powerful performance indicator of noise affecting a time scale signal is the *Allan Deviation*. The Allan Deviation does not indicate thought whether a constant rate, i.e. a frequency offset, is present or not. This fact can easily be evaluated, for example, by calculating the index:

$$J = \frac{1}{N}\sum_{k=1}^{N}\left[z(k) - \left(\frac{1}{N}\sum_{i=1}^{N} z(i)\right)\right]^2 \tag{13}$$

where N is the number of samples of the process. J corresponds to the classical variance of the measured phase difference. $J_{free\,running}$ is obtained with the measures $z(t)$ as in (8) while $J_{controlled}$ has the measures $Z_2(t)$ as in (10). To be effective in rate elimination a steering method must have

$$J_{controlled} < J_{free\,running}$$

In general, we can compare n steering methods by comparing their indices J. For example, if we have two methods and $J_1 << J_2$, we conclude that method 1 is more effective than method 2 in rate elimination.

Other criteria can be formulated to evaluate the performances of steering strategies, such as criteria on robustness, or on dynamic performance of the control systems (i.e. *assessment time* and so on).

6. ANALYSIS OF THE MOST COMMON STEERING METHODS

6.1. "Signal Based" Open loop control approach

We refer to the most classical control approach, which is adopted in many laboratories, the *open loop* approach. The open loop is used to eliminate the frequency offset (or the deterministic frequency drift) of a clock signal.

We start with the autonomous discrete-time system (8). If our aim is to eliminate a constant rate, we calculate an average frequency offset by way of

$$\bar{y}_f(k) = \frac{z(k) - z(k-N_y)}{N_y\,\tau} \tag{14}$$

By means of the estimates $\bar{y}_f(k)$ the frequency at instant $(k+1)$ is predicted as $\hat{y}_f(k+1)$ by a least square approach or using the last value. Being the controlled system expressed by (10), we impose a control law such that:

$$u(k) = -\hat{y}_f(k+1) \tag{15}$$

In some papers, we have an indication of how the parameter N_y is quantified. For example one can take $N_y = 20$ ([1], [12]) with $\tau_0 = 1\,day$, being τ_0 the sampling time.
If we instead deal with a signal $z(t)$ characterized by a deterministic frequency drift, as in the case of an H Maser, we want to control the clock signal in order to eliminate the effect of both the frequency offset and drift. We thus calculate a series of average frequency offsets (14) at each epoch k. Then by observing the frequency estimates over a period $N_d\tau$, also the frequency drift $\bar{d}_f$ is estimated by a suitable technique and the frequency for the next instant is predicted as $\hat{y}_{f,d}(k+1)$. We then apply a steering:

$$u(k) = -\hat{y}_{f,d}(k+1) \tag{16}$$

This approach was used, for example, in the steering of the experimental Galileo System Time [13].

6.2. Optimal Control

Referring to the model expressed by the system (11), the control signal $u(k)$ is such that the following cost function J_{Opt} is minimized:

$$J_{Opt} = \mathbb{E}\left[\sum_{k=1}^{\infty}\left(\mathbf{x}^{z,u}(k)^T W_Q \mathbf{x}^{z,u}(k) + u(k) W_R u(k)\right)\right] \tag{17}$$

This functional is obtained by Optimal Control, a well-established control technique that is ideal for such optimization problems. In Optimal Control we introduce a *gain* vector $\mathbf{G}$, such that:

$$u(k) = -\mathbf{G}\,\mathbf{x}^{z,u}(k)$$

Very different results could be consequent to different choices of the weighting matrices W_Q and W_R. Among the matrices that we will take into account, some are proposed by Koppang in [2], [6], and some are different, with the aim of evaluating the effect of changes in the "control effort" matrix W_R. Namely we use the following set of matrices.

$$\begin{cases}
(1) & W_Q = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} \\
 & W_R = 10^5 \\
(2) & W_Q = \begin{bmatrix} 0.001 & 0 \\ 0 & 0.001 \end{bmatrix} \\
 & W_R = 1 \\
(3) & W_Q = \begin{bmatrix} 10^{-4} & 0 \\ 0 & 10^9 \end{bmatrix} \\
 & W_R = 10^9 \\
(4) & W_Q = \begin{bmatrix} 10^{-4} & 0 \\ 0 & 2\,10^6 \end{bmatrix} \\
 & W_R = 1
\end{cases} \tag{18}$$

6.3. Bang Bang - Sliding Mode Control

The control law known as Bang Bang [3] is used to steer GPS time to UTC(USNO). We interpret the Bang Bang in the frame of Control Theory. Let's suppose that, at time t_0, we can approximate the phase deviation of the signal with the linear function

$$x_1(t) = x_1(t_0) + x_2(t_0)\,t$$

We then impose a steering signal on the phase acceleration as:

$$x_1^u(t) = x_1(t_0) + x_2(t_0)\,t + u(t_0)\,\frac{t^2}{2} \tag{19}$$

First we suppose to apply a control signal at t_0 opposite to the phase rate x_2 as:

$$u^0(t_0) = -K_{steering}\,sign(x_2(t_0)) \tag{20}$$

then we look for the time $\bar{t}$ at which the function (19) has a stationary point:

$$\frac{d\,x_1(t)}{dt}|_{t=\bar{t}} = x_2(t_0) - K_{steering}\,sign(x_2(t_0))\,\bar{t} = 0$$

By substituting $\bar{t}$ and (20) in (19) we obtain $x_1(\bar{t})$. We define the quantity $s(t_0)$ as

$$s(t_0) = x_1(\bar{t}) = x_1(t_0) + \frac{x_2(t_0)^2\,sign(x_2(t_0))}{2\,K_{steering}}$$

At time t_0, we apply a final control signal given by

$$u(t_0) = -K_{steering}\,sign(s(t_0))$$

If we take into account model (11) with the actuator (12), the non-linear control law applied at time k becomes:

$$u(k) = -K_{steering}\,sign(s(k)) \tag{21}$$

where:

$$s(k) = x_1^{z,u}(k) + \frac{x_2^{z,u}(k)^2\,sign(x_2^{z,u}(k))}{2\,K_{steering}}$$

In Control Theory this methodology is known as *sliding mode control* [14], and s is the *switching surface*. For GPS case ([3] and [6]) the value of $K_{steering}$ is $10^{-19}s/s^2$.

7. SIMULATIONS AND PERFORMANCE COMPARISON

We tested the different strategies by simulation. We use simulated data that allow us to compare absolute clock signal results. The performances of the steered clocks are evaluated in terms of the performance indices previously described, such as the Allan Deviation and the index J in (13). We present the results of a single simulation only, but these results are representative of a large class of simulated cases. We have studied two application cases that are described in the following. The sampling time τ is one day.

7.1. Cesium clock steered versus very stable reference time scale

This case may represent a Cesium master clock (typically representing $UTC(k)$) steered to the Universal Coordinated Time (UTC). Indicating by Cs the clock signal to be steered and ref the reference time scale, we consider the Cesium clock affected by WFN and Random Walk Frequency Noise (RWFN) whose standard deviations are given by typical Cs values:

$$\begin{cases} \sigma_1^{Cs} = 2.592\,10^{-3}\,\mu s/(day)^{\frac{1}{2}} \\ \sigma_2^{Cs} = 1.496510^{-4}\,\mu s/(day)^{\frac{3}{2}} \\ \sigma_3^{Cs} = 0\,\mu s/(day)^{\frac{5}{2}} \end{cases} \qquad (22)$$

which correspond [11] to the Allan deviation reported in Figure 3. The standard deviations of analogous noises of the reference clock are ten times lower. The simulated free absolute clock signals are shown in figure 2. We now

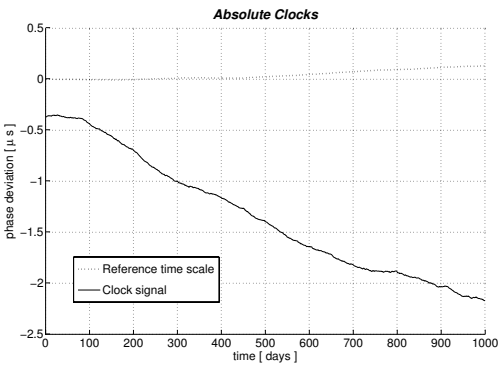


Figure 2: Absolute phase deviations of the free running clocks: case 1

apply the three different control techniques.

a) *Signal Based.* The clock signal is supposed to be characterized by constant frequency offset. We use the control expressed by (15), with $N_y = 20$.

b) *Optimal Control.* We analyze the effect of the different sets of weighting functions W_Q and W_R seen for Optimal Control. In Figure 3 we can see the stability of the absolute controlled signals in the different cases. We choose controller 2 for the following analysis, because it shows the best performance in terms of Allan deviation.

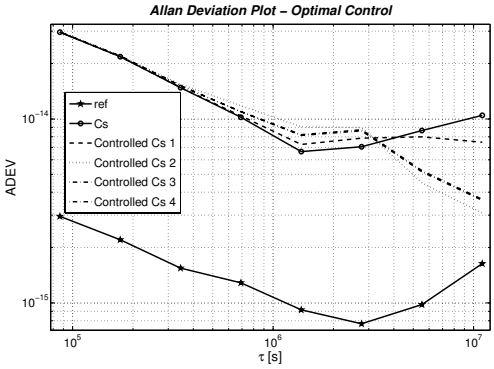


Figure 3: Allan deviation of the controlled (with Optimal control with different weighting matrices) absolute Cs phase deviations

c) *Bang Bang*: here we apply the laws discussed in section 6.3 with $K_{steer} = 10^{-19} s/s^s$.

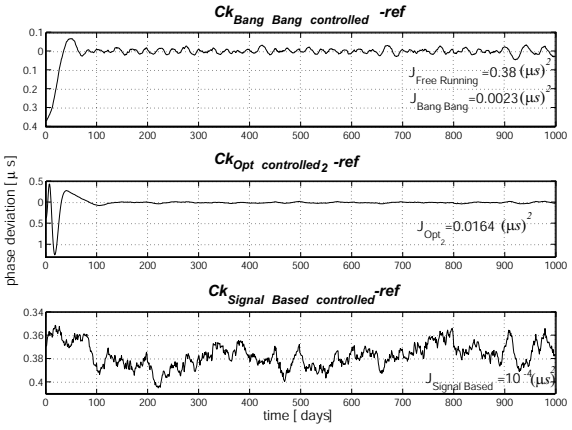


Figure 4: Phase deviation of the steered clock signals with respect to the reference clock(Steered with the three different steering methods) and estimates of the J index

In Figure 3 and Figure 4 results on the performances of the different control methods are shown. Despite its simplicity, the open loop (Signal based) method appears to be the most performing, both from the point of view of the Allan Deviation and the rate elimination estimated by the index J.

7.2. H Maser VS Cesium clock time scale

In this second case, we will simulate the signal of an H Maser clock and a reference Cesium clock. The H Maser

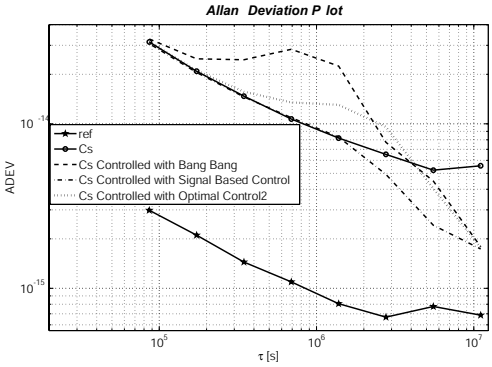


Figure 5: Allan deviation of the absolute phase deviations of the steered clock (Steered with the three different steering methods) and of the reference time scale

noise variances are given by

$$\begin{cases} \sigma_1^{Hm} &= 8.6410^{-5}\,\mu s/(day)^{\frac{1}{2}} \\ \sigma_2^{Hm} &= 2.9930\,10^{-5}\,\mu s/(day)^{\frac{3}{2}} \\ \sigma_3^{Hm} &= 10^{-19}\,\mu s/(day)^{\frac{5}{2}} \end{cases}$$

corresponding [13] to the Allan deviation of Figure 7. The reference clock is given by a Cesium clock (22). The simulated data are shown in Figure 6.

We consider the discussed control methods:

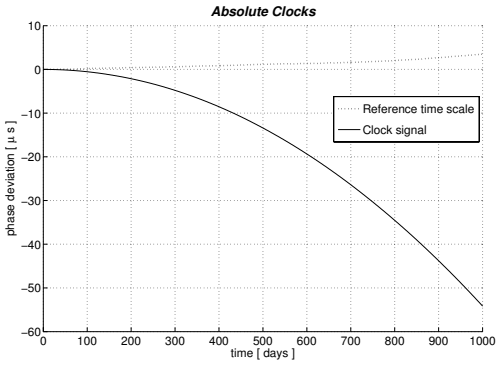


Figure 6: Absolute phase deviations of the free running clocks: Case 2

a) *Signal Based*. The clock signal is supposed to be affected by frequency drift. We thus use the control expressed by (16), with $N_y = 1$ and $N_d = 20$.

b) *Optimal Control*. We first evaluate the different weighting functions. In Figure 7 we represent also the ADEV of the simulated H maser with drift removed to show its pure stochastic noise. The control aims to reduce the frequency drift and the resulting ADEV of the controlled signal is actually reduced as far as the frequency drift effect is concerned. The effect of the different Optimal Controls are different. We retain weighting matrices 1 and 3 in the following analyses. As we will see later, choosing the weighting matrices 1 is very good for noise reduction, but it is not optimal to minimize the index J. The control parameters would need a better tuning to obtain the best capabilities from this system.

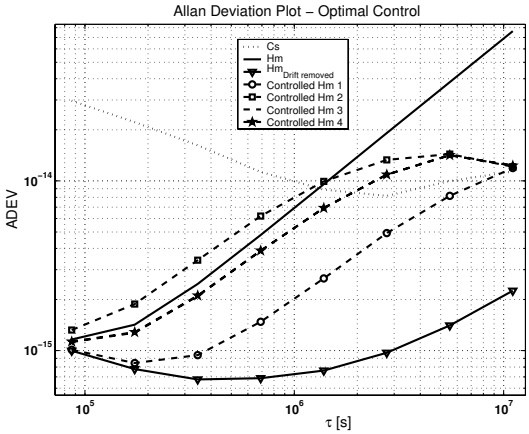


Figure 7: Allan deviation of the controlled(with Optimal control with different weighting matrices) H Maser phase deviations

c) *Bang Bang*. Same case of the previous section.

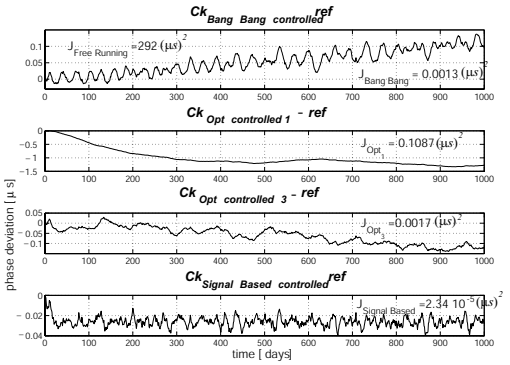


Figure 8: Phase deviation of the steered clock signals with respect to the reference clock (Steered with the three different steering methods) and estimates of the J index

The results on the performances of the different control methods are shown in Figure 8 and Figure 9. The open loop (Signal based) method seems to perform well. The Optimal Control is not the best in eliminating the rate and frequency drift despite its good performances in terms of Allan Variance. Note that, for Bang Bang control, in

this case, the results show that the control signal $u(k) = -K_{steer} sign(s(k))$ is not "strong" enough to correct the clock signal when its slope is bigger than a critical value. For the Bang Bang, as well as for the Optimal control and

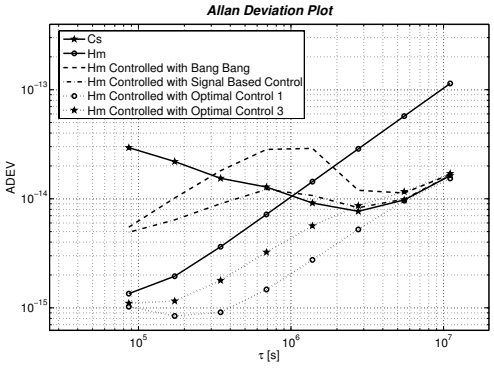


Figure 9: Allan deviation of the absolute phase deviations of the steered clock (Steered with the three different steering methods) and of the reference time scale

Signal based, an optimal tuning of the parameter should be examined to optimize the results.

8. CONCLUSIONS

We have analyzed the most common clock signal steering methods and we have written them in the framework of Control Theory. We have shown that the problem of steering a time scale to a reference can be formally interpreted as a classical control problem, and thus solved using consolidated control methods. We described three well known steering approaches: the "Signal Based approach", which is one of the most classical approaches, also studied for an application to the Galileo System Time; the "Bang Bang" method, currently used for the steering of GPS to UTC(USNO); the "Optimal Control approach", considered at USNO. These steering methods are analyzed, and their effects are evaluated by simulation. The performances of the methods are measured using the Allan Deviation and the variance index J. The optimal tuning of the different control parameters was not tested in this work and that could clearly improve the obtainable performances. Here we wanted to concentrate on the interpretation of the steering strategies showing that they can be described in the frame of the Control Theory whose potentiality seems very high for improving clock control. The obtained results are very interesting and suggest new approaches for clock steering strategies.

9. ACKNOWLEDGEMENTS

The authors are greatly indebted to Prof. S. Bittanti of Politecnico di Milano for helpful discussions, for the encouragement and support to this work.

10. REFERENCES

[1] L. G. Bernier and G. Dudle, *Practical Performance of the UTC(CH.R) Real Time Realization of UTC(CH) and Prospects for Improvement*, Proc. European Time and Frequency Forum, 2004, 288–295

[2] P. A. Koppang and R. Leland, *Linear Quadratic Stochastic Control of Atomic Hydrogen Masers*, IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 1999, Vol. 46, 517–522

[3] L. S. Huser and S. T. Hutsell *1998 GPS Time Transfer Performance*, 30th Annual Precise Time and Time Interval (PTTI) Meeting, 1988, 125–130

[4] J. Doyle, and B. Francis and A. Tannenbaum, *Feedback Control Theory*, 1990, MacMillan Publishing

[5] L. A. Breakiron and P. A. Koppang, *Frequency Steering of Hydrogen Masers*, 1996 IEEE International Frequency Control Symposium, 1113–1122

[6] P. A. Koppang and D. N. Matsakis, *Control of precise Clocks*, IV TSAS-25, 1999, 1–4

[7] D. Allan *Time and Frequency (Time- Domain) Characterization, Estimation and Prediction of Precision Clocks and Oscillators*, Proc. IEEE Transactions on Ultrasonic, Ferroelectrics, and Frequency Control, 1987, Vol. 40, 647–654.

[8] B. D. O. Anderson and J. B. Moore, *Linear Optimal Control*, 1971, Prentice-Hall, New Jersey

[9] G. C. Goodwin, M. M. Seron and J. A. De Doná, *Constrained Control and Estimation*, 2004, Springer

[10] L. Galleani, L. Sacerdote, P. Tavella and C. Zucca *A mathematical model for the atomic clock error*, Metrologia, 2003, Vol.3, 257–264

[11] C. Zucca and P. Tavella *The Clock Model and Its Relationship with the Allan and Related Variances*, IEEE transactions on ultrasonics, ferroelectrics, and frequency control, 2005, Vol.52, 288–295

[12] L. G. Bernier *Experimentation at METAS with a Simple Steering Algorithm based on Linear Prediction*, Proc. IVth International Time Scale Algorithms Symposium, 2002, 257–264

[13] F. Cordara, R. Costa, L. Lorini, D. Orgiazzi, V. Pettiti, I. Sesia, P. Tavella, P. Elia, M. Mascarello, M. Falcone, J. Hahn, *Experimental Galileo System Time (E-GST): One Year of Real Experiment*, Proceedings of the 36th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 2004

[14] V. Utkin, Guldner and Jingxin Shi, *Sliding Mode Control in Electromechanical Systems*, 1999, Taylor & Francis

Typical combined uncertainty evaluation on the Ku band TWSTFT link

P. Merck and J. Achkar

Observatoire de Paris
SYRTE
61, avenue de l'Observatoire F-75014 PARIS, France
philippe.merck@obspm.fr ; joseph.achkar@obspm.fr

Keywords: TWSTFT, communications satellite, comparison, uncertainty budget.

ABSTRACT

This paper presents a typical evaluation of both statistical uncertainty U_A and systematic uncertainty U_B on the Two-Way Satellite Time and Frequency Transfer (TWSTFT) link, based on the differentiation of the time equation. Recently, calibration of the differential delay stations using a portable station has been achieved and the results are taken into account. Typical values of U_A = 0,53 ns and U_B = 0,96 ns are found, giving a combined uncertainty U = 1,10 ns (1σ) calculated on the two-way two-minute measurement data. As an example, clock differences between BNM-SYRTE and PTB over the year 2004 are reported including uncertainties, using both Ku band TWSTFT and GPS Single Frequency Common View techniques. Type A and Type B uncertainty results on TWSTFT links may be considered in the uncertainty budget of International Atomic Time (TAI).

1. INTRODUCTION

TWSTFT technique is an alternative method to GPS time transfer used to compare atomic clocks disseminated in different metrology institutes in the world (Europe, United States and the Asia Pacific region), through geostationary communications satellites using microwave links between the ground clocks and the satellite transponders [1], contributing to TAI calculated by the Bureau International des Poids et Mesures (BIPM). The principle scheme of this technique is described in Figure-1.

TWSTFT technique permits to eliminate the first-order effects due to path reciprocity of signals, contributing to an improvement of accuracy when comparing two remote time scales noted *UTC(1)* and *UTC(2)*. The differentiation δ of the time equation, with the assumption that all the parameters describing it are uncorrelated, is given by equation (1), according to [2] and [3].

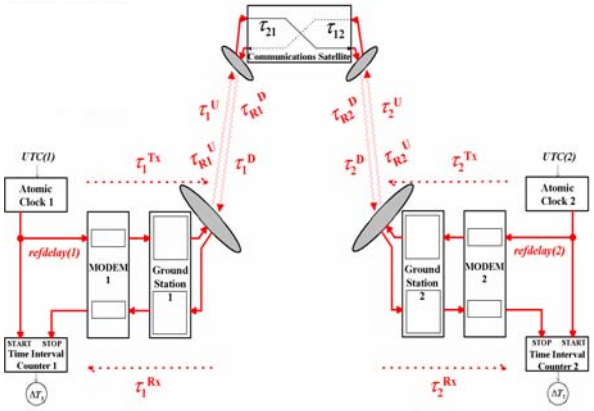


Figure-1: principle scheme of TWSTFT technique

$$\delta(UTC(1)-UTC(2)) = \sqrt{\frac{1}{4}\left(\delta\Delta T_1^{\,2}+\delta\Delta T_2^{\,2}+\delta\Delta\tau_{TR}^{\,2}+\delta\Delta\tau_{UD}^{\,2}+\delta\Delta\tau_{sat}^{\,2}+\delta\Delta\tau_{R}^{\,2}\right)+\delta refdelay(1)^2+\delta refdelay(2)^2} \quad (1)$$

with (i = 1,2 ; j = 1,2 ; $i \neq j$):

- ΔT_i: time difference measured with a Time Interval Counter (TIC) at station i,

- *refdelay(i)*: delay between *1pps UTC(i)* and *1pps TX(i)* signals at station i,

- $\Delta\tau_{TR}$ = *CALR(1)-CALR(2)*: delay difference inside earth stations, with *CALR(i)* = $\tau_i^{TX} - \tau_i^{RX}$,

- $\Delta\tau_{UD} = \Delta\tau_{UD12} - \Delta\tau_{UD21}$: delay difference due to the atmospheric path for up and down links, with $\Delta\tau_{UDij} = \tau_i^U + \tau_j^D$,

- $\Delta\tau_{sat} = \tau_{12} - \tau_{21}$: delay difference of the satellite transponder,

- $\Delta\tau_R = \Delta\tau_{R12} - \Delta\tau_{R21}$: delay difference due to earth rotation, with $\Delta\tau_{Rij} = \tau_{Ri}^U + \tau_{Rj}^D$.

From this relationship, the development of an uncertainty budget which takes into account Type A and Type B uncertainties [3] of all the parameters described above is carried out in this paper.

2. EVALUATION OF STATISTICAL UNCERTAINTY U_A

Type A uncertainty on the two-way two-minute clock difference *UTC(1)-UTC(2)* can be evaluated by the mean of statistical methods. The following contributions are considered:

- ($A_{1,1}$, $A_{2,1}$) : contributions due to noise in the measurement two-way system including: atomic clocks, earth stations, satellite transponder, atmospheric paths and the earth rotation (Figure-1).

$A_{1,1}$ and $A_{2,1}$ are estimated by considering the *DRMS(i)* value recorded in the ITU data file [2] of each station i (i = 1,2), which represents statistical uncertainty on each 1s measurement. Therefore $A_{1,1}$ and $A_{2,1}$ uncertainties on each two-minute sample are calculated using (2), according to [3]:

$$A_{i,1} = \frac{1}{2}\frac{DRMS(i)}{\sqrt{N_1}}, \quad N_1 : number\ of\ samples \tag{2}$$

with the assumption that White-noise Phase Modulation (WPM) is observed for averaging times of two minutes [4].

- ($A_{1,2}$, $A_{2,2}$) : contributions due to the measurement of *refdelay(i)* at each station i (i = 1,2). These two contributions can be estimated considering the *RSIG(i)* value recorded in the ITU data file of each station *i*, which represents statistical uncertainty on each individual measurement. Then $A_{1,2}$ and $A_{2,2}$ uncertainties on each *refdelay(i)* are calculated using (3), as WPM is observed:

$$A_{i,2} = \frac{RSIG(i)}{\sqrt{N_2}}, \quad N_2 : number\ of\ samples \tag{3}$$

- ($A_{1,3}$, $A_{2,3}$) : contributions due to measurement of internal delays inside each earth station including transmission and reception paths. These contributions are calculated using a calibration device, in relative mode with a Portable Station (PS) or in absolute mode with a satellite simulator. The first one is considered in this paper, therefore these contributions are defined by (4):

$$A_{i,3} = \frac{1}{2}\delta\left(CALR(i) - CALR(PS)\right)_A \tag{4}$$

3. EVALUATION OF SYSTEMATIC UNCERTAINTY U_B

Type B uncertainty on the two-way two-minute clock difference *UTC(1)-UTC(2)* is estimated by other means (other than statistical methods). The following contributions are considered:

- ($B_{1,1}$, $B_{2,1}$) : contributions due to the measurement of ΔT_i two-way two-minute data by using the Time Interval Counter (TIC) of each station i (i = 1,2) :

$$B_{i,1} = \frac{1}{2}\delta\Delta Ti_B \tag{5}$$

In the case of using the internal TIC of a SATRE modem, these contributions are equal to 0,2 ns (1σ) for a 2,5 MChip/s pseudo-noise (PN) transmission code sequence.

- ($B_{1,2}$, $B_{2,2}$) : contributions due to the measurement of *refdelay(i)* using a TIC at each station i (i = 1,2). By using an SR620 TIC, the contributions are equal to 0,335 ns (1σ).

- ($B_{1,3}$, $B_{2,3}$) : contributions due to the calibration of time difference delays of earth stations. They are considered when a portable station is used in co-location with earth stations [5].

These contributions are calculated using (6), according to (1).

$$\begin{aligned} B_{i,3} &= \frac{1}{2}\delta\left(CALR(i) - CALR(PS)\right)_B \\ &= \frac{1}{2}\sqrt{\delta\Delta T_{PS_B}{}^2 + 4\delta refdelay\left(PS\right)_B{}^2 + \delta\Delta T_{i_B}{}^2 + 4\delta refdelay\left(i\right)_B{}^2} \end{aligned} \tag{6}$$

- B_4 : contribution due to ionosphere path, with the assumption that uplink and downlink frequencies are equal for each station (e.g. European links) :

$$B_4 = \frac{1}{2}\delta\Delta\tau_{UD_B} = \left(\frac{40{,}3\,TEC}{c\sqrt{2}}\right)\sqrt{\left(\frac{2}{f_u{}^3}\right)^2\delta f_u{}^2 + \left(\frac{2}{f_d{}^3}\right)^2\delta f_d{}^2} \tag{7}$$

with:

TEC : total electronic content
c : light speed in the ionosphere
f_u : uplink frequency
f_d : downlink frequency

This contribution is negligible within the acceptable frequency range in operation, for uplink and downlink frequencies.

- B_5 : contribution due to delay difference inside the satellite transponder, which is equal to zero for European links as the same transponder is used for up and down links.

$$B_6 = \frac{1}{2}\delta\Delta\tau_{RB} = \frac{1}{2}\sqrt{\begin{aligned}&\left[\cos[LA(1)]\sin[LO(1)-LO(s)]-\cos[LA(2)]\sin[LO(2)-LO(s)]\right]^2\left[\left(\frac{2Rr}{c^2}\right)^2\delta\Omega^2+\left(\frac{4\Omega Rr}{c^3}\right)^2\delta c^2+\left(\frac{2\Omega r}{c^2}\right)^2\delta R^2+\left(\frac{2\Omega R}{c^2}\right)^2\delta r^2\right]\\&+\left(\frac{2\Omega Rr}{c^2}\right)^2\left[\begin{aligned}&\left[-\cos[LA(1)]\cos[LO(1)-LO(s)]+\cos[LA(2)]\cos[LO(2)-LO(s)]\right]^2\delta LO(s)^2\\&+\left[\left[\sin[LA(1)]\sin[LO(1)-LO(s)]\right]^2\delta LA(1)^2+\left[\sin[LA(2)]\sin[LO(2)-LO(s)]\right]^2\delta LA(2)^2\right]\\&+\left[\left[\cos[LA(1)]\cos[LO(1)-LO(s)]\right]^2\delta LO(1)^2+\left[\cos[LA(2)]\cos[LO(2)-LO(s)]\right]^2\delta LO(2)^2\right]\end{aligned}\right]\end{aligned}} \quad (8)$$

- B_6 : contribution due to Sagnac effect, according to (8) :

with (i = 1,2)

LA(i) : latitude of the station *i*
LO(i) : longitude of the station *i*
LO(s) : longitude of the satellite s
c : light speed in the atmosphere
r : radius of the earth
R : distance between satellite to center of earth
Ω : angular speed of the earth

The most important source of error in (8) is due to the uncertainty on the geostationary satellite position (*δLO(s)* parameter) which is typically on the order of 0,5 ° (e.g. orbital position of INTELSAT satellites are given with a precision of 0,5 °).

4. TWSTFT UNCERTAINTY BUDGET

The uncertainty budget on the PTB-OP time transfer link calibrated with a portable station [5] is given as a typical example in Table-1. It takes into account the whole contributions described above. Total Type A, Type B and combined uncertainties are calculated using (9), (10) and (11) respectively.

$$U_A = \sqrt{A_{1,1}^2 + A_{2,1}^2 + A_{1,2}^2 + A_{2,2}^2 + A_{1,3}^2 + A_{2,3}^2} \quad (9)$$

$$U_B = \sqrt{\begin{aligned}&B_{1,1}^2 + B_{2,1}^2 + B_{1,2}^2 + B_{2,2}^2 + B_{1,3}^2 + B_{2,3}^2\\&+ B_4^2 + B_5^2 + B_6^2\end{aligned}} \quad (10)$$

$$U = \sqrt{U_A^2 + U_B^2} \quad (11)$$

Uncertainty contribution	*Uncertainty value (ns)*
$A_{1,1}$	0,033
$A_{2,1}$	0,016
$A_{1,2}$	0,003
$A_{2,2}$	0,002
$A_{1,3}$	0,307
$A_{2,3}$	0,430
Type A uncertainty (ns)	**0,53**
$B_{1,1}$	0,200
$B_{2,1}$	0,200
$B_{1,2}$	0,335
$B_{2,2}$	0,335
$B_{1,3}$	0,540
$B_{2,3}$	0,540
B_4	0,000
B_5	0,000
B_6	0,180
Type B uncertainty (ns)	**0,96**
Combined uncertainty (ns)	**1,10**

Table-1: typical uncertainty budget on the PTB – OP TWSTFT link

($A_{1,1}$, $A_{2,1}$) and ($A_{1,2}$, $A_{2,2}$) uncertainties were estimated respectively by the computation of a mean *DRMS(i)* and a mean *RSIG(i)* for each station *i* over the year 2004 period, assuming that N_1 = 120 samples and N_2 = 100 samples at each station.

$A_{1,3}$ and $A_{2,3}$ are estimated by the standard deviation on each individual measurement obtained on both calibration results *CALR(PTB)-CALR(PS)* and *CALR(OP)-CALR(PS)* [6]. Indeed, as only 13 and 28 measurements were carried out respectively at BNM-SYRTE and PTB (Figure-2), it is not sufficient enough to determine what kind of noise is predominant in the data and then it is not possible to compute a standard deviation on mean calibration results.

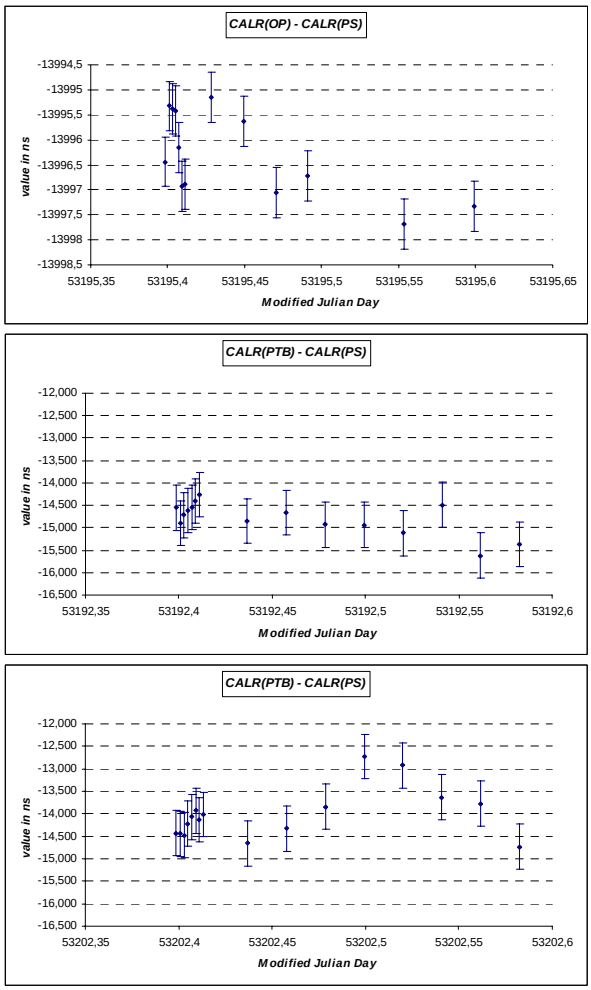


Figure-2: relative calibration results at PTB and OP [6]

Accuracy of 1,10 ns on the clock differences *UTC(PTB)-UTC(OP)* using the Ku band TWSTFT link through INTELSAT 903 is obtained. Type A and Type B uncertainty values obtained are consistent with the evaluation made by the BIPM (which is respectively 0,5 ns for Type A uncertainty and 1 ns for Type B uncertainty) [7].

5. APPLICATION: CLOCK DIFFERENCES EXPECTED

In order to compare the uncertainty budget of the TWSTFT link against other time transfer techniques, like the GPS, an application is given in this paper with remote clock differences of BNM-SYRTE and PTB, including uncertainties, using the two available techniques: Ku band TWSTFT through INTELSAT 903 and GPS Single Frequency Common View.

For the GPS time transfer, PTB-OP data (five-day means) including related uncertainties are considered from the BIPM Circular T [7], with Type A and Type B uncertainties of 2,5 ns and 5,0 ns respectively.

For the two-way time transfer, the evaluation of type A uncertainty on each TWSTFT point averaged on 5 continuous days is not easy to obtain. The example given below develops this particular point.

The time deviation $\sigma_x(\tau)$ calculated on the quantity *UTC(PTB)-UTC(OP)* obtained by TWSTFT over the year 2004, is plotted in Figure-3:

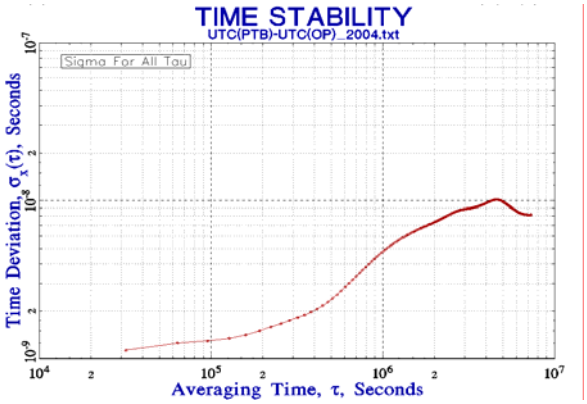


Figure-3: time deviation on *UTC(PTB)-UTC(OP)* obtained by TWSTFT over the year 2004

For a τ = 4d, $\sigma_x(\tau) \approx$ 1,8 ns.

Moreover, the Allan standard deviation of $8{,}2.10^{-15}$ and $6{,}0.10^{-15}$ are expected for averaging times of 4 days on *UTC(PTB)* based on Caesium Laboratory standard and *UTC(OP)* based on high performance commercial Caesium standard, respectively [8]. By using equations (11) and (12) when considering White-noise Frequency Modulation WFM (slope in $\tau^{+1/2}$ on 4d according to Figure-3) [9],

$$\frac{MDEV^2}{ADEV^2} = 0{,}5 \qquad (11)$$

$$TDEV = \frac{\tau.MDEV}{\sqrt{3}} \qquad (12)$$

the calculated $\sigma_x(\tau)$ due to the clocks noise is about 1,5 ns compared to the value given on the plot of Figure-3.

So, for a τ = 5d, the noise of the TWSTFT link, not appearing in Figure-3, is either White-noise Phase Modulation WPM (slope in $\tau^{-1/2}$) or Flicker-noise Phase Modulation FPM (slope in τ^0). For the both cases, the value of 0,53 ns on the type A calculated in considering two-way two minute measurement data seems to be conservative for data averaged on 5d.

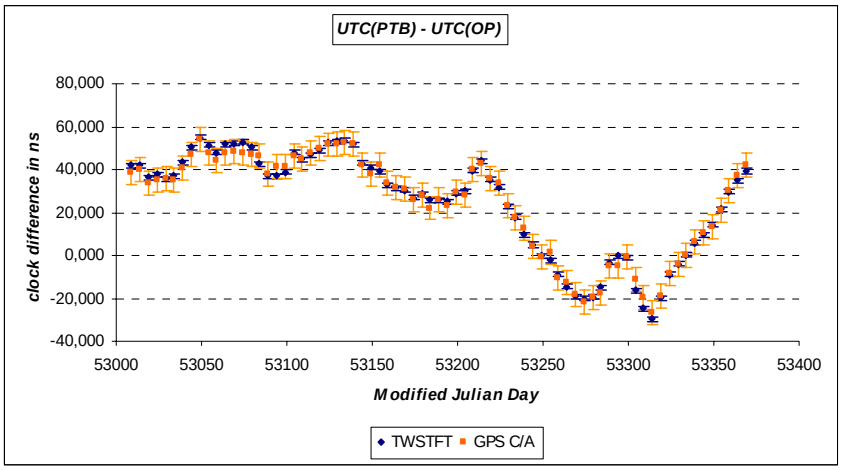


Figure-4: *UTC(PTB)-UTC(OP)* over the year 2004 using TWSTFT and GPS

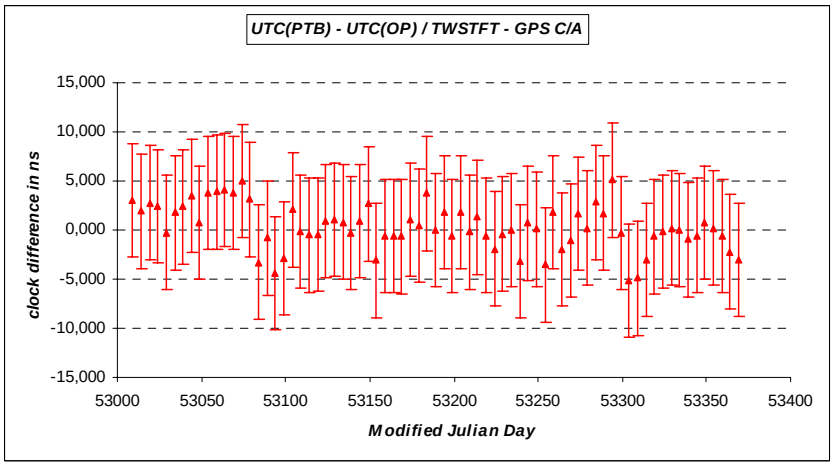


Figure-5: difference between TWSTFT and GPS C/A code techniques over the year 2004

Two-way two-minutes clock differences data (5-days averaged) calculated according to GPS data dates and related differences between the two techniques are reported in Figure-4 and Figure-5, respectively.

In the results plotted above, contributions of non reciprocal residual effects described in the TWSTFT time equation and reported in table-2 [5], have been considered:

non reciprocal residual effect	*value (ns)*
$\frac{1}{2}\,\Delta\tau_{TR}$	-6990,950
$\frac{1}{2}\,\Delta\tau_{UD}$	0
$\frac{1}{2}\,\Delta\tau_{sat}$	0
$\frac{1}{2}\,\Delta\tau_{R}$	8,180
$\frac{1}{2}\,\Delta\tau$	**-6982,770**

Table-2: contributions of non reciprocal residual effects on the PTB-OP TWSTFT link

A good agreement is observed between the two techniques (Figure-5), giving a mean value of 0,32 ns, widely within the combined uncertainty of the non favorite case as 5,6 ns for the GPS. This particular result indicates that no important effect was forgotten in the evaluation of the uncertainty budget on the two-way time transfer link. However, the level of GPS C/A uncertainties limits the capabilities to detect some effects on the order of nanosecond or below when using the Ku band TWSTFT link.

CONCLUSION

A typical evaluation of Type A and Type B uncertainties on the Ku band TWSTFT link was carried out by taking into account all the parameters describing the two-way time equation. Accuracy of 1,10 ns was obtained on the clocks comparison between PTB and

BNM-SYRTE. This very promising result shows that the nanosecond accuracy level can be reached with this time transfer link. This work will be completed in considering additional contributions of uncertainties due mainly to high frequency reflections effects in cables and a possible variation of internal delays of the stations during the calibration campaign, which are the most important source of errors in the uncertainty budget. Moreover, calibration results were assumed to be constant values over one year. In reality, some effects like temperature, humidity and pressure can affect differential delay stations in time and the monitoring of these internal delays will be carried out at BNM-SYRTE using a satellite simulator actually in development, in order to take into account systematic effects in the uncertainty budget as seasonal variation.

ACKNOWLEDGEMENT

The authors thank the CCTF Working Group on TWSTFT for their support. Thanks also for the useful discussions done with Dirk Piester (PTB) on TWSTFT aspects and with P. Wolf and P. Uhrich on GPS time transfer aspects.

REFERENCES

[1] D. Kirchner. “Two-Way Time Transfer Via Communication Satellites”, Proc. of the IEEE, vol. 79 (7), pp. 983-989, July 1991.

[2] ITU-R TF.1153-2 Rec. The operational use of two-way satellite time and frequency transfer employing PN codes, 2003.

[3] ISO/TAG 4/WG 3. Guide to the expression of uncertainty in measurement, June 1992.

[4] D Kirchner, H Ressler, R Robnik. Recent work in the field of two-way satellite time transfer carried out at the TUG. 11th EFTF 1997, Neuchâtel, pp. 205-208.

[5] D Piester, R Hlavac, J Achkar, G de Jong, B Blanzano, H Ressler, P Merck, J Becker, O Koudelka. Calibration of four European TWSTFT Ku band earth stations with a portable station through Intelsat 903. To be presented at EFTF 2005.

[6] O Koudelka, H Ressler, B Blanzano. Two-Way Time Transfer Calibration campaign. Institute of Applied Systems Technology report, 2004.

[7] BIPM. Circular T, 2004, available on http://www.bipm.fr/en/scientific/tai/time_ftp.html.

[8] C Thomas. Impact of new clock technologies on the stability and accuracy of the International Atomic Time TAI. 1996 IEEE International Frequency Control Symposium, pp. 1123-1130.

[9] D B Sullivan, D W Allan, D A Howe, F L Walls. Characterization of clocks and oscillators. NIST Technical Note 1337, 1990.

A NEW 25 MHZ STANDARD FREQUENCY AND TIME TRANSMITTER FOR LOCAL SERVICE

K. Kalliomäki[1], T. Mansten[1], I. Iisakka[1]

[1] MIKES, Electricity Group, Otakaari 7B, FI-02150, Espoo, Finland

Keywords: standard frequency, radio clock

Abstract

In Finland, TV synchronization signal has been used to distribute standard frequency and time. Recent digitalization of the TV links has degraded the accuracy of the method and made it unfit for the calibration of high quality oscillators. Alternative ways to distribute standard frequency and time in Helsinki area were studied. As a result MIKES developed a new 25 MHz standard frequency transmitter and a suitable phase lock receiver for it. The carrier of the transmitter is based on an atomic clock and it is amplitude modulated by a time code.

First tests with the transmitter-receiver combination indicate that Allan FDEV is $1 \cdot 10^{-11}/\tau$ where τ is between 1 and 100s. This is good enough for most oscillator calibrations. As a result of the project, a freely available standard frequency and time distribution service will be available in the near future.

1 Introduction

MIKES (earlier VTT Automation) has taken care of Finnish official time and frequency since 1976. Finnish broadcasting company (YLE) disseminated as a partner time signal using audio channels, and frequency with the aid of TV synchronisation signal. The relative accuracy of the frequency obtained was better than $1 \cdot 10^{-11}$.

The digitalisation of TV links caused random delays at every link station and the accuracy of the frequency is now 100 times worse than earlier when analogue links were used. Calibration of high quality oscillators needs better, easily available reference. That is why MIKES has designed its own transmitter and a suitable receiver.

2 Frequency selection

After studying the tables of allocated frequencies for standard frequency transmitters, we selected 25 MHz (λ = 12 m) for several reasons. First, when using a high frequency the antenna size is convenient. Secondly, 25 MHz lies at short wave band, which means that one can use cheap, commonly available receivers and transmitter components. Earlier (1977) we tried to use 250 MHz, but with poor success due to lack of suitable receivers. Third reason was that we could utilise 27 MHz citizen band components.

3 25 MHz Transmitter

Transmitter (see Fig. 1) is now intermittently on air for test cycles to find a suitable format for time signal modulation. IRIG-B and DCF-77 code are considered. The latter is better for aural time signal reception because of 1 Hz (second tick) bit rate and easily recognisable minute mark.

Transmitter power is around 100 W and antenna gain near 10 dB (vertical ¾ λ dipole). Antenna height during test cycles is 20 m above sea level but ~~it~~ the antenna will be placed to the height of 40 m by the end of this year.

The 25 MHz carrier comes from a synthesised HP signal generator that is locked to our atomic clock (Cs2). The relative accuracy of the carrier frequency is better than $1 \cdot 10^{-13}$. One of the time codes mentioned above amplitude modulates the carrier. The two codes include full time code consisting of day, hour, minute and second data.



Fig. 1: The 25 MHz transmitter

4 Coverage area of the signal

The coverage area is estimated to be the region inside the Kehä III (the Helsinki area belt highway) when using outdoor antenna with preamplifier. At first tests signal level at Soukka, 12 km westward from MIKES seemed to be 0,5 mV

when using a good outdoor antenna. The signal was traced at Lohja, too, when using a sensitive receiving system.

Because of good conductivity of seawater, stable ground wave signal may be usable at Tallinn, Estonia, offering a tool for "international" frequency comparison. AS Metrosert in Tallinn, an organisation like MIKES, is interested in testing this possibility.

5 25 MHz standard frequency receiver

As a thesis work, a receiver for the signal mentioned above has been developed. It consists of standard components like a conventional short wave receiver; an oven controlled 10 MHz crystal oscillator (OCXO) and IRIG demodulator. Associated electronics take care of phase locking etc.

Figure 2 shows the structure of the standard frequency receiver. Both the received signal and the OCXO reference signal, phase modulated at 2,5 kHz, pass the short wave receiver. Thus delays in receiver are mainly eliminated. Phase difference is measured by applying a 2,5 kHz phase detection circuitry.

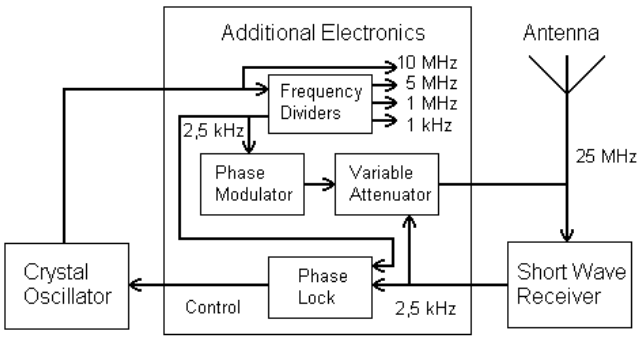


Fig. 2: Simplified receiver block diagram

Figure 3 shows the FDEV (Allan frequency deviation) of the 10 MHz output of the receiver. Test site with an indoor antenna was in a laboratory, within 50 meter from transmitter antenna. Thus the signal level was quite high, around 1 mV, but the received signal was subject to interference from the laboratory. A special phase comparator, capable of resolving $1 \cdot 10^{-13}$ just in one second, was used along with a Hydrogen maser (HM1) as a stable reference.

The uppermost curve shows the stand-alone stability (no lock) of the OCXO and the two lowest curves the stability in locked condition. One second frequency stability is better than $1 \cdot 10^{-11}$ (10 ps/s) . The stability is inversely proportional to the integration time and the reference stability $1 \cdot 10^{-13}$ is gained on ten minutes.

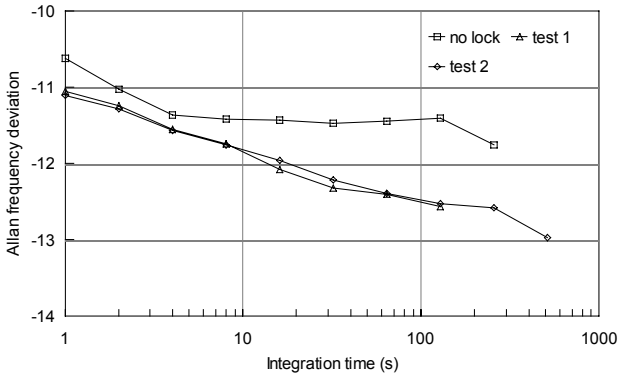


Fig. 3: Allan frequency deviation of the 10 MHz output of the receiver

Conclusions

The short-term frequency stability of the constructed receiver, at least near the transmitter site, is about 100 times better than the earlier TV-frequency standard before year 2001, when stable analog TV-links still were used. Thus the constructed TX/RX system fulfils our goals. Approximately 100 TV-frequency standard units, produced by Jutel Oy, are still in use in Finland due to the comprehensive coverage of the TV signal. Because the coverage area of the new system is the capital area only, we estimate that some 30 new receiver units are needed.

GPS is a strong competitor, but governed by US military authorities. In EU we can't trust that this service is always available, at least free of charge (maybe GALILEO changes this). Therefore, an independent service seemed to be a good alternative for GPS. If the field is strong enough, the standard frequency is available even without a special receiver, e.g. by using a spectrum analyzer for frequency comparison.

THE EFFECT OF COUNTER'S TRIGGER LEVEL ON TIMING BELOW NANOSECOND

K. Kalliomäki [1], J. Mannermaa [2] T. Mansten [1]

[1] MIKES, Electricity Group, Otakaari 7B, FI-02150, Espoo, Finland
[2] Nokia Corporation Technology Platform, FIN-33100, Tre, Finland

Keywords: Cable delay, dispersion

Abstract

In Time Metrology clocks produce ticks, whose rise times vary from 2 ns to 20 ns, but accuracy requirements approach 1 ps. Therefore, the trigger level of a time interval counter becomes critical at mutual comparisons. In addition, the bandwidth of coaxial cables is limited and the propagation delay increases significantly on frequencies below 30 MHz. Thus the shape of the timing pulse may change dramatically, if long cables are used.

To find the right trigger level several factors have to be taken into account, like varying pulse shapes, frequency dependent attenuation and propagation delay in cables, pulse amplitude and trigger level tolerances, mismatching of impedances etc. The simulation show that in case of well designed instrumentation 50% point is natural selection for trigger level. But if there are any suspicions of pulse shape distortion due to limited bandwidth or of amplitude levels, one has to lower trigger level, even down to 10 % of nominal pulse amplitude. To fulfil most stringent requirements, optical fibers or e.g. 100 MHz carrier as a vernier in connection with conventional tick must be used.

1 Introduction

Some 40 years ago one µs accuracy was good enough in time comparisons, but now we need even one ps. But due to the limited bandwidth of standard coaxial cables the second pulse has not become steeper; rise time is typically between 2 and 20 ns. The rise time of timing pulse may be thousand times higher than the required accuracy.

In the past history trigger level of time interval counter was not critical but now it is very important, key question in mutual comparisons.

In principle the rise time of second pulses should be 0.1 ns or less to guarantee trouble-free selection of the trigger level. But commonly used coaxial cables are not good enough. Firstly, the attenuation in coaxial cables is too high rounding pulses, because the spectrum of pulse extends to 10 GHz. Secondly, delay dispersion below 30 MHz changes the shape of the pulse.

Quite soon we have to use optical components to avoid problems with conventional RF-connectors, cables and joints.

This problem came up as an important theme in EUROMET TF-Technical Committee Annual Meeting at NPL, April 2004.

Due to complicated situation with several factors affecting to the optimum trigger level, authors used computer simulation to study the problem. In addition, phase delay and attenuation was measured in typical cables.

Following factors have been considered; varying pulse shapes, frequency dependent attenuation in cables, pulse amplitude and trigger level tolerances, mismatching of impedances.

It seems that we have to live with standard cables, connectors and available second pulses quite long. Then optical fibers replace coaxial cables in critical applications. Temporarily we can combine standard tick as a coarse measure and e.g. 100 MHz carrier phase as a vernier.

2 Mission impossible, beginning of the pulse

In one application, Loran-C, the receiver calculates the beginning of the timing burst to distinguish between different propagation modes like ground- and skywave. In this case however, (1) the shape of the transmitted pulse is carefully standardized and even (2) pre-distorted to match receiver filter. In addition, timing burst contains 100 kHz cycles (“vernier”); receiver measures the amplitudes (3) of successive cycles and extrapolates (4) the beginning. By this way the receiver can omit the precursors of burst, namely signal components outside -3dB limits of the receiver RF-filter which propagate fast through the filter. In all filters signal group delay has a maximal value at band pass and near zero value at band reject.

To find the beginning of pulse is out of question in case of conventional timing pulse, the second tick. Pulse shape is (1) not standardized and it's not (2) pre-distorted to match low pass properties of coaxial cables, or other possible filters. Only one (3) comparator is set to conventional trigger level. There are no suitable "time markers" (4) along the slope to check the result like in Loran-C.

Fig. 1 shows what happens to steep pulses after 2 ns delay in a lossy cable (-6 dB at 160 MHz). The beginning of the pulse is very indistinct, signal slope starts from zero. Precursors, i.e. Fourier components above band limit give rise to unclear start of the pulse. They propagate faster than basic signal, which has the expected delay. In other words, the pulse shapes change. It is impossible to determine the beginning of the pulse without knowing the exact shape of it.

It is possible, by using a suitable filter to predistort the second pulse, i.e. to cut frequency components outside the – 3 dB limit of propagation media. Then the pulse shape is preserved quite well but still we only with difficulty can find the beginning of the pulse.

As a conclusion, try 50 % trigger level as a first approximation

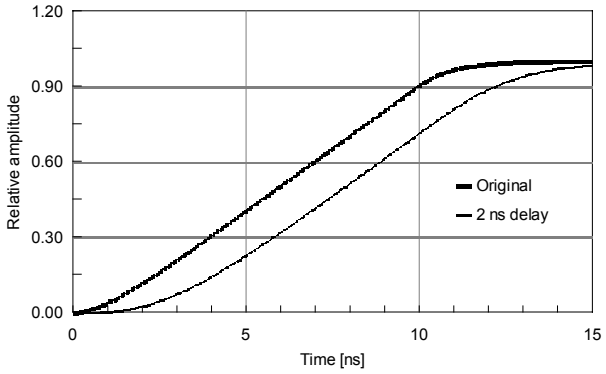


Fig. 1: Original and delayed timing pulse

3 Bandwidth of the cable good enough compared to the tick

If the rise time of the tick is around 10 ns or higher and cables are relatively short and of high quality (low attenuation at 100 MHz), Fig. 2 shows what happens when cables of different lengths (L, 2·L and 3·L) are used. Pulse shape does not change essentially.

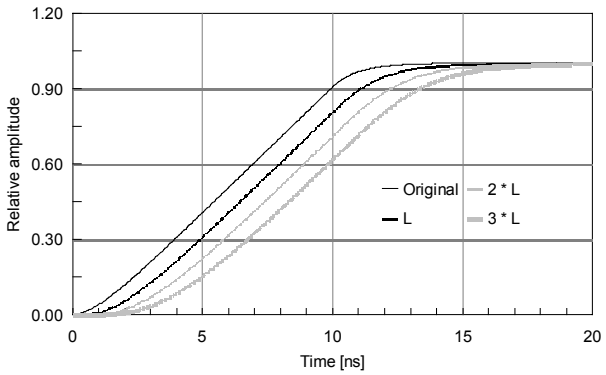


Fig. 2: Changes of pulse shapes with different lengths of cable

Next we have to consider slopes of those pulses, with comparator (its trigger level) in mind. After some calculation we see that the slopes (δu/δt) of all pulses in Fig. 2 are quite the same between 10% and 80% of maximum amplitude.

The height of the pulse depends on several factors like mismatch, attenuation in cable and specifications of pulse source. Possible variation (δu) in voltage is directly proportional to amplitude, assuming that the rise time is constant. Because the slope is relatively flat, the time error δt increases proportional to the selected trigger level.

It easy to see, that e.g. at 80 % trigger level the amplitude variation (δu) is four times bigger than at 20 % trigger level point. Timing error (δt) increases correspondingly. Therefore, one has to select a low trigger level if one compares timing pulses from different sources where the above mentioned amplitude problems may be present.

In fixed installations problems hardly exist assuming that ambient conditions are stable. In this case 50 % point is good selection.

4 Bandwidth of the cable insufficient compared to the tick

In case of a long cable (at least 10 m) and relatively sharp pulse (rise time less than 10 ns), the cable has remarkable effect on pulse shape. Fig. 3 depicts what happens to a pulse shape at different lengths of "poor" cables. Pulse shape changes continuously. Cable works as a filter limiting high frequency components, decreasing rise time and "rounding" the general pulse shape. What is the correct trigger level if one has to measure the delay in cable?

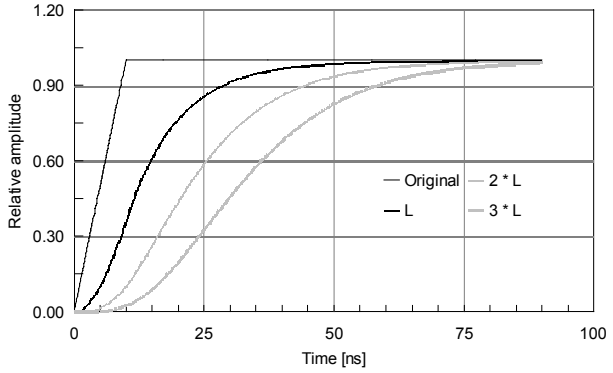


Fig. 3: Cable too long for sharp timing pulse

The basic problem is that the original pulse is too sharp to the transmitting media. One has to "match" the pulse to the media (cable). In principle it is possible to "predistort" an excessively sharp pulse e.g. by using a long (>>10 m) cable as a filter. Then the delay in unknown cable is measured by using this rounded pulse.

The slope (δu/δt) is highest between 30..40 % of maximum amplitude but slope does not change very much between 20 % and 80 %. Therefore, if the above-mentioned amplitude uncertainties are true, one has to use low (e.g. 30 %) trigger level. But in case of true cable delay measurement, 50 % trigger level is recommended in connection with a pulse matched to the frequency response of the cable. Sharp pulse is not the best.

5 Systematic delay error due to delay dispersion

Long cables always change the shape of the timing pulse, high frequency components are attenuated and some velocity dispersion takes place. Fig. 4 shows the measured velocity coefficient of an RG223 coaxial cable. Below 30 MHz the inductance of the center conductor increases due to the increasing skin depth, and both the propagation delay and the characteristic impedance change. Above 30 MHz the velocity coefficient is quite stable.

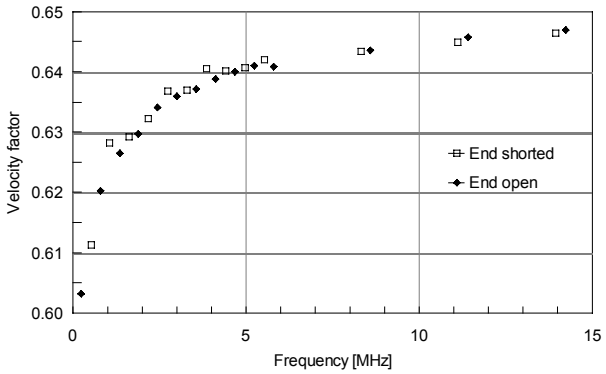


Fig. 4: Measured velocity factor of a long cable (RG223)

6 Recommendations for cable delay measurements

Don't use too sharp pulse. High frequency attenuation and velocity dispersion change the pulse shape. It could be useful to "match" the input pulse by using long cable as a "filter". Check amplitude values at both ends of the cable and then use 50 % trigger level.

Remember time interval counter errors. Reversing the channels (A -> B and B -> A) could be useful. Accurate amplitude measurement may be a problem, check your value by using a >100 % trigger level; counting must stop. The inherent nonlinearity of the analog interpolator of a time interval counter may change the reading even by the amount of guaranteed resolution. In order to lower the counter error, it could be useful to lock the signal source and the reference oscillator of the counter to a common reference and add phase shifter to the reference input of the counter.

If accurate delay is the key question, it could be reasonable to use second pulse as a coarse marker and e.g. 100 MHz sine wave as a vernier. The vernier frequency must be higher than 30 MHz to avoid delay dispersion.

COMPUTER SPICE SIMULATION OF OSCILLATION STABILITY AREA FOR DUAL-MODE CRYSTAL OSCILLATOR

A. Kosykh, A. Gubarev, S. Zavjalov, I. Khomenko,
Omsk State Technical University
Russia, 644050, Omsk, Mira Avenue 11, intecs@mail.ru

1. INTRODUCTION

Now the certain interest for using a dual-mode oscillator in highly stable TCXO and OCXO is observed. Excitation of an additional resonator oscillation mode (mode B in SC-cut resonator) as the sensor of temperature allows reducing considerably static and dynamic errors of temperature measurement in comparison with external thermosensor using. Such XO (crystal oscillator) are capable to provide frequency stability of 10^{-8} and above, simultaneously possessing small weight and dimensions and fast setting time. The main obstacle in a way of two-frequency XO using is low stability of a dual-frequency self-oscillation regime.

It is especially difficult to ensure two-frequency excitation of resonator in oscillator circuits with one active element (single loop oscillator). Interaction of asynchronous oscillations on a common nonlinearity in most cases leads to suppressing all other oscillations by more active one. In these circuits the two-frequency oscillations is possible only in the presence of active element, having specific volt-ampere characteristic at operating point neighborhood and auto-biasing network. But even at presence of necessary nonlinearity, the oscillator can support a two-frequency oscillation regime only at the certain oscillation amplitudes. It means that the two-frequency oscillator will drive not all resonators at two mode, but only resonators with the certain ratio of mode dynamic resistances.

The necessity to take into account a high orders nonlinear effects extremely complicates the analytical analysis of two-mode oscillators. Research and optimization of such oscillators carry out mainly by experimental methods. Therefore, the developer may estimate an ability of the oscillator circuit to work as two-mode only for the particular cases for available resonators. The oscillator circuit can work with some resonator and don't work with others, but this doesn't give enough information about stability area boundary. Development of the methods of quartz oscillator computer simulation [2-4] enables to receive this information by numerical experiment.

For dual-mode oscillator analysis, the method, based on calculation of parameters of a nonlinear active part of self-oscillatory system by means of widespread simulation programs, with the subsequent analysis of entire self-oscillatory system by analytical methods is used (Fig. 1) [3]. Such approach allows bypassing restrictions of SPICE simulation programs for high-Q circuits simulation (extremely long transients). There, the SPICE simulation is used only for calculation of negative resistance of one-port drive circuit macromodel.

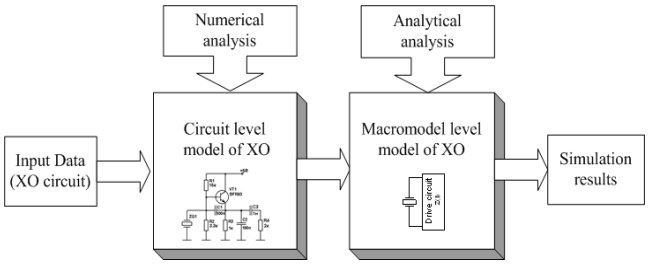


Fig.1 Combined, numeric-analytical method of crystal oscillator analysis

2. MACROMODEL LEVEL MODEL OF DUAL-MODE XO

As a result of any research of self-oscillatory circuit, answers on two interrelated, but methodologically different questions, can be received. of questions deals with amplitude relations in the oscillator: amplitude transients, steady-states and their stability. of questions deals with phase equations, and concerns to stability of oscillations frequency. In this case it is necessary to consider not only the processes in the oscillator circuit, but also the nonlinear acoustic effects in a quartz crystal, including interaction of oscillations modes.

We will be limited only by the first group of questions which supposes to represent the resonator as linear system of independent resonance circuits.
There we use a set of assumptions:

1. The resonator has only two active modes which asynchronous frequencies and amplitude vs. frequency characteristics are not crossed in a pass-band.
2. Q-factors of both resonator modes are high and resonator currents (of each mode) are close to harmonic.
3. One-port drive circuit is broadband (in relation to the resonator pass-band), and its negative resistance in a frequency band of both oscillations depends on frequency poorly.

In this case dual-mode crystal oscillator can be represented as equivalent circuit of Fig.2.

On this circuit the two-frequency oscillator consists of system of two coupling oscillation loops. Each self-oscillating loop characterizes the oscillation excitation process for a corresponding mode.

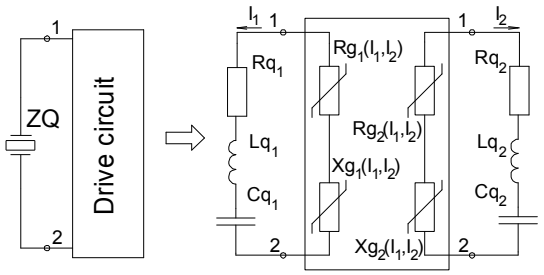


Fig.2 Equivalent circuit of dual-mode crystal oscillator

Both loops consist of equivalent RLC-circuit describing resonant properties of a resonator mode, and the nonlinear two-pole describing the drive circuit. Behavior of the drive circuit for each oscillation are completely defined by functional dependence of its harmonically linearized complex resistance (at the oscillation frequency) vs. amplitudes of mode currents: $Rg_k(I_1,I_2)$ $Xg_k(I_1,I_2)$, k=1..2. Using of two-dimensional dependences allow to take into account the fact of oscillations interaction in the drive circuit.

3. ANALYSIS OF STEADY-STATE LOCAL STABILITY FOR DUAL-MODE XO

Amplitude transients in the above circuit are described by system of two second order differential equations for each self-oscillating loop. For this system the shortened differential equations for amplitude of oscillations has been derived by slowly varying functions method [3]:

$$\frac{dI_n(t)}{dt} = -\frac{I_n(t)\cdot w_{q_n}\cdot(R_{q_n}+R_{g_n}(I_1,I_2))}{2Q_nR_{q_n}}$$

$$w_{q_n} = \frac{1}{\sqrt{L_{q_n}\cdot C_{q_n}}}, \text{ n=1... 2.} \qquad (1)$$

The corresponding to (1) phase space is a plane of I_1, I_2. Steady-state points of this system are defined by phase plane points where phase speed is

$$\frac{dI_1(t)}{dt} = \frac{dI_2(t)}{dt} = 0$$

equal to zero, i.e. .

For two-frequency mode steady-state point finds as the solution of the equations:

$$R_{g_1}(I_1,I_2)+R_{q_1}=0 \quad R_{g_2}(I_1,I_2)+R_{q_2}=0 \qquad (2)$$

Let's define conditions in which dual-mode steady-state will be stable. For this purpose we will analyze system (1) on stability by Lyapunov's method.

For small deviations from the point of equilibrium ΔI_1 and ΔI_2 the nonlinear system (1) may be linearized to the first approximation equations:

$$\frac{d\Delta I_n}{dt} = -\frac{I_n w_{q_n}\left(\frac{\partial R_{g_n}}{\partial I_1}\Delta I_1 + \frac{\partial R_{g_n}}{\partial I_2}\Delta I_2\right)}{2Q_nR_{q_n}}, \text{ n=1... 2.} \qquad (3)$$

There and hereinafter Rg_1 and Rg_2 are means as two-dimensional functions from amplitudes of resonator mode currents $R_{g_1}(I_1,I_2)$, $R_{g_2}(I_1,I_2)$.

Any deviation from the point of equilibrium of linearized systems (3) will fade, if roots of the characteristic equation for this system have a negative real part. For convenience of transformations make following replacements of variables:

$$p_k = \frac{I_k w_{q_k}}{2Q_kR_{q_k}} \quad a_k = \frac{\partial R_{g_k}}{\partial I_1} \quad b_k = \frac{\partial R_{g_k}}{\partial I_2}, \quad k=1,2.$$

Then the characteristic equation of system (3) will become:

$$\begin{vmatrix} -p_1a_1 - z & -p_1b_1 \\ -p_2a_2 & -p_2b_2 - z \end{vmatrix} = 0$$

.

This can be written as:

$$z^2 + (p_1a_1 + p_2b_2)z + p_1p_2(a_1b_2 - b_1a_2) = 0$$

.

Roots of this equation are defined by expression:

$$z_{1,2} = -\frac{1}{2}((p_1a_1 + p_2b_2) \mp \sqrt{(p_1a_1 + p_2b_2)^2 - 4p_1p_2(a_1b_2 - b_1a_2)})$$

The point of equilibrium of oscillatory system is stable for next conditions:

$$p_1a_1 + p_2b_2 > 0$$

$$a_1b_2 - b_1a_2 > 0$$

Expanding variables, receive following conditions for local stability of two-mode steady-state:

$$\frac{I_1w_{q1}}{Q_1R_{q1}}\cdot\frac{\partial R_{g1}}{\partial I_1} + \frac{I_2w_{q2}}{Q_2R_{q2}}\cdot\frac{\partial R_{g2}}{\partial I_2} > 0$$

$$\frac{\partial R_{g1}}{\partial I_1}\cdot\frac{\partial R_{g2}}{\partial I_2} - \frac{\partial R_{g1}}{\partial I_2}\cdot\frac{\partial R_{g2}}{\partial I_1} > 0 \qquad (4)$$

4. SIMULATION PROCEDURE FOR OSCILLATION STABILITY AREA CALCULATION

Apparently from (4), for estimation of crystal oscillator circuit ability to support a two-frequency mode it is enough to obtain and analyze of two-dimensional dependences of negative resistance vs. resonator currents for drive circuit $R_{g_1}(I_1,I_2)$ and $R_{g_2}(I_1,I_2)$. To calculate these dependences it is possible to use the general-purpose EDA (SPICE, MicroCAP). For this purpose in oscillator circuit the resonator is replaced by two equivalent sources of a harmonic current Fig.3.

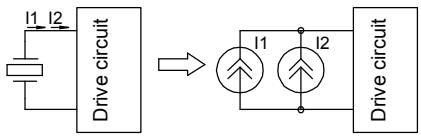


Fig.3

By means of varying of current sources amplitude it is possible to place the drive circuit in the conditions similar to the steady-state mode with this currents.

The real component of complex resistance of the drive circuit can be calculated under the formula:

$$R_{g_n}(I) = \frac{2}{T \cdot I_n} \cdot \int_0^T u_g(t) \cdot \sin(\ _{q_n} \cdot t) dt \text{, n=1.. 2,} \quad (5)$$

where $T = \frac{2}{w_{q_2} - w_{q_1}}$ is the period of integration, $U_g(t)$ is instant value of a voltage on the resonator. The current of signals sources in (5) changes under the sine wave law. The period of integration corresponds to the period of difference frequency, since it is necessary that the integer number of periods of all presented oscillations kept within the period of integration.

Applying the conditions of local stability of two-mode steady-state to the obtained dependences of $R_{g_1}(I_1, I_2)$ and $R_{g_2}(I_1, I_2)$ it is possible to allocate the area of oscillations amplitudes in which the two-frequency regime is possible. The obtained area of oscillations amplitudes, allowed in a two-frequency mode, can be transformed to area of resonator mode C and mode B dynamic resistances using the equations (2).

5. ANALYSIS OF DUAL-MODE OSCILLATIONS STABILITY AREA FOR COLPITTS CIRCUIT

In present work dual-mode oscillation stability areas for some Colpitts oscillators, earlier experimentally investigated in [5] has been calculated. At Fig.4 and Fig.5 the oscillator circuit and the calculated areas of resonator resistance on B-mode and C-mode, for which the circuit is stably excited at two frequencies are resulted. Areas of stability are calculated at various values of a bias voltage Vbias for the resonator with parameters: F=10MHz, $Q_{modeC}=10^6$, $Q_{modeB}=2 \cdot 10^6$

Apparently from plots, the size of stability area of a two-frequency mode is very critical to value of a bias voltage. The best operating point of the circuit lies in a range of bias voltages of 0.8-0.9 V. Overrun of limits of this interval leads to abrupt reduction of area of resonator resistance in which the circuit is capable to be excited at two-frequency. In other words, the oscillator circuit is capable to support two-frequency excitation only at operating point, were volt-ampere characteristic has nonlinearity.

At Fig.6 the modified Collpitts circuit with special RC-circuit for suppression of difference frequency oscillation (Fb-Fc) is presented. This decision allows increase the area of stable oscillation (Fig. 7).

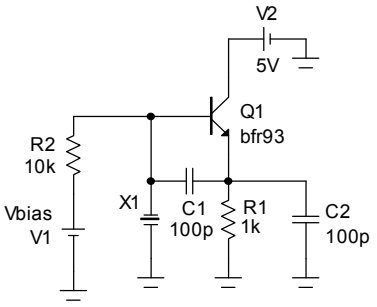


Fig. 4. Colpitts oscillator circuit

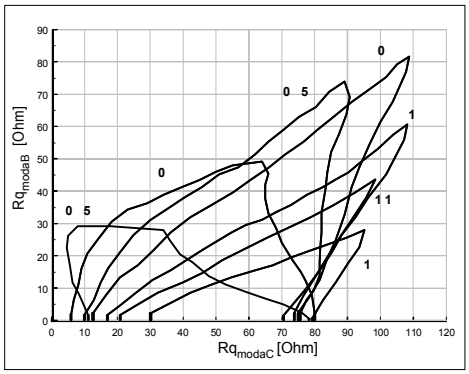


Fig. 5. Stability areas for Colpitts circuit

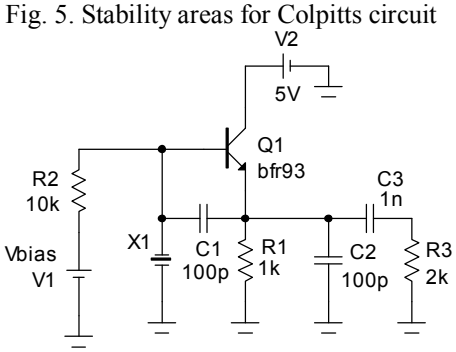


Fig. 6. Modified Colpitts circuit

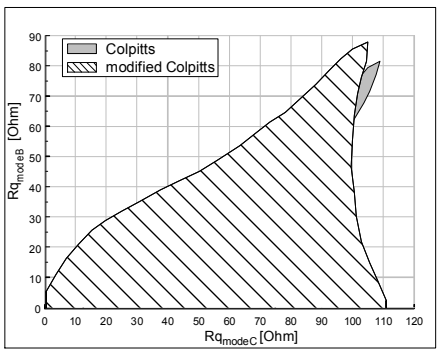


Fig. 7. Stability areas for modified Colpitts circuit

6. CONCLUSIONS

For obtaining information about ability of drive circuit to support a two-mode excitation of the resonator it is enough the next:

1. By means of widespread SPICE simulation program calculate dependences of negative resistance on amplitudes of modes currents $R_{g_1}(I_1, I_2)$ and $R_{g_2}(I_1, I_2)$ for oscillator drive circuit.
2. Analyze the received dependences using criterion of stability (3) and build plot of oscillations stability area in coordinates of resistances of the resonator.

By investigating plots «Oscillation stability area» it is possible obtaining the following information:

- The area of motional resistance of the resonator for each oscillations mode for which it is possible a two-frequency excitation by the given oscillator circuit.
- The area of deviations of resonator motional resistance (for specific resonator) on each oscillation mode for which one, the two-frequency excitation regime is remains.
- Stimulating and worsening conditions factors of joint oscillations excitation.

REFERENCES

1. D.P. Tsarapkin, F.L. Walls. Noise in oscillators with two asynchronous oscillations – Proc. 1999 Joint Meeting EFTF – IEEE IFCS, pp. 1158-1162.
2. A.V. Kosykh, A.N. Lepetaev and S.A. Zavjalov. Investigation of dual-mode excitation of crystal oscillator. – Proc. of 1999 Joint Meeting EFTF – IEEE IFCS, pp. 1154-1157.
3. A..A. Gubarev, A.V. Kosykh, S.A. Zavjalov, A.N. Lepetaev. SPICE simulation of high-Q crystal oscillators: Single and dual-mode oscillator analysis. – Proc. of 2003 joint meeting EFTF – IEEE IFCS, pp. 606-614.
4. Addouche M., Ratier N., Gillet D., Brendel R., Lardet-Vieudrin F., Delporte j. ADOQ: a quartz crystal oscillator simulation software. - Proc. of 2001 IEEE IFCS and PDA Exhibition, pp. 753-757.
5. A.V. Kosykh., I.V. Abramson., V.P. Bagaev. Dual-mode crystal oscillators with resonators, excited on B and C modes // Proc. of 48th IEEE Annual Frequency Control Symposium. – 1994. - P. 578-586.

DUAL-MODE CRYSTAL OSCILLATORS FOR DTCXO AND OCXO

Anatoly V. Kosykh.
Omsk State Technical University. Omsk, Russia

ABSTRACT

Dual-mode crystal oscillators are very attractive and perspective components for precision temperature compensated (TCXO) [1-3] or ovenized (OCXO) [4] reference frequency sources. This is the consequence of the fact that dual-mode oscillator may be used as self-temperature sensing device with excellent static and dynamic measuring resolution. In such devices the best volume coincidence of reference resonator and thermosensor is achieved. However dual-mode excitation of single crystal resonator is vexatious because of complex oscillations interaction in resonator and excitation circuit. At dual-mode oscillation regime the series of technical problems (circuit complication, spurious oscillation, oscillation instability and other) may appear and the solution of this problems is very difficult and non-obvious.

In this paper the complex of problems related to dual-mode oscillation excitation is investigated.

1. OSCILLATION TYPE CHOOSING

The main criterion for oscillation type choosing is good quality of reference mode (high Q-factor, smooth and flat F-T curve) in combination with high F-T curve slope of temperature-sensing mode. The secondary criterion is the absence of oscillation interaction and (as a consequence) stability of motional resistance of both oscillations vs. temperature. In practice the secondary criterion may be a dominant. Some variants of dual-mode resonators may be used. In all cases the reference oscillation (F_R) is C-mode in SC or AT- cut resonator, but as a sensor mode (F_T) may be used one of the next:

- B-mode in SC-cut (or similar orientated) resonator [3, 4, 5, 6]
- Face chear low frequency oscillation
- Anharmonic of C-mode [7]
- Frequency difference between 3-rd and tripled 1-st harmonics of C-mode [1, 2, 8, 9]

In Table 1 main parameters of temperature sensitive oscillation for these cases are shown. The frequency of C-mode is 10 MHz.

Table 1

Oscillation modes	Temperature mode frequency, kHz	F-T curve absolute slop Hz/^{0}C	F-T curve relative slop ppm/^{0}C
B-mode – C-mode	900	-300	-330
3-rd - 3·1-st	0.15	-14	-95

In principal the relative steepness of F-T curve can be made as much as need by subtraction of constant frequency "support" (this may be named as change of nominal value). However, it usually does not change a "signal/noise" ratio of thermosensor and it is not so essential, as a high absolute steepness. Because each oscillation mode has some frequency instability (vs. load, supply voltage, vibration) and this instability leads to temperature measuring error, the better will be the resonation with highest slope of temperature mode F-T curve. In this case frequency drift caused by non-temperature reasons will produce lowest measuring error. Besides when a measuring instrument operates in a mode of a frequency meter (i.e. when the number of the arrived impulses is counted up for fixed time), nominal value of measured frequency does not influence at all on accuracy of measurement.

Being based on this, it is possible to approve, than high will be an absolute steepness of the sensor curve, than more exact will be measurement result and external destabilizing factors will influence less. In this plan the using of B and C modes of SC-cut looks obviously advantageous. Such decision is offered for a long time [3-6], but has encountered a problem connected with instability of motional resistance of a B- mode in an interval of temperatures, leading to catastrophic failure of oscillations of this mode at some temperatures. In spite of the fact that the amazed zones are very narrow, uses of resonators with "dips" becomes impossible. Presence of failures and complexities of struggle against them have caused serious interest to use of frequencies difference of SC-cut harmonics as a thermosensor [1,2,8,9]. This decision has appeared technologically stable, that has overbalanced in developers' opinion a potential advantage of a high steepness of a mode "B" characteristic.

During some time we used resonators with excitation of B and C modes and have come to conclusion, that there are the decisions, allowing to get rid of failures of activity in the characteristic of a B-mode. The reason of failures (or major of the reasons) that at certain temperatures a frequency (or their overtones) of unwanted mode can coincide to frequency of a B-mode and due to arisen oscillation coupling the energy of the basic oscillation will slip over to unwanted oscillation. The decision of a problem consists in regulation of frequencies of unwanted modes so that in a working interval of temperatures their frequencies were not crossed with frequency of the basic fluctuation. In [10] the results of optimization of the sizes of piezoplate of yxbl/23o52′/34o cut by criterion of the maximal percent of resonators without essential failures are given. «Plano-plano» piezoplate had the sizes (10 -

13)x(6.5 - 7.0) mm. The varied parameter was the piezoplate width. On Fig. 1 shown a relative part (examine 100 resonators) of suitable resonators depending on piezoplate width.

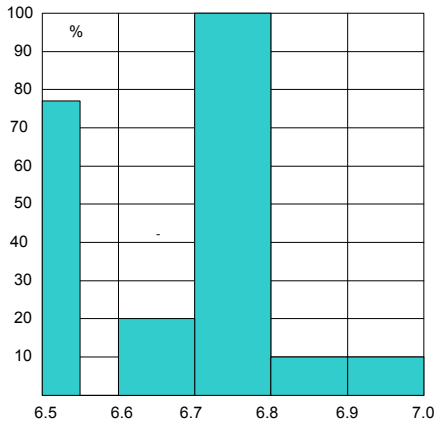


Fig. 1. Statistic data of valid resonators vs. plate width

The other way is changing the dioptries of Plano-convex plate. On Fig 2 are shown motional resistance of mode B vs. temperature for 5 resonators for radius of plate curvature of 100, 150, and 200 mm.

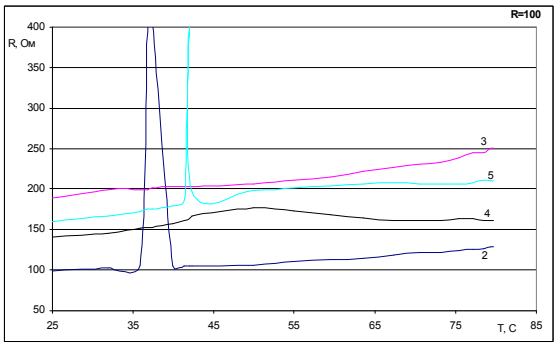


Fig. 2 Motional resistance of B-mode for plano-convex resonator with R=100 mm.

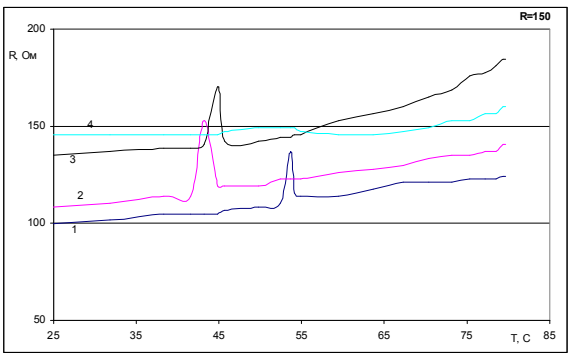


Fig. 3 Motional resistance of B-mode for plano-convex resonator with R=150 mm

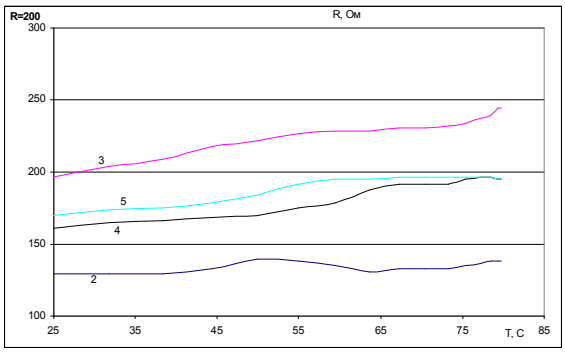


Fig. 4 Motional resistance of B-mode for plano-convex resonator with R=200 mm

These data are published for the first time. It is well noticeable that there is strong dependence of activity dips (at diagram it looks like a peaks of resistance) on plate curvature.

As it may be concludes from this there are some technological methods for improving of monotonous of motional resistance vs. temperature characteristic of B-mode. This means that resonators with simultaneous excitation of B and C modes may be used as a base for high stable OSXO and DTSCO.

OSCILLATION CIRCUIT CHOOSING.

It is not enough to have the good resonator, it is necessary to excite it correctly. Excitation of two frequencies simultaneously is very difficult technical problem. For its decision a plenty of various ways is offered. The most well knowing ideas are further described.

2 . 1 . Dual-frequency oscillators with separated nonlinear elements

A dual-frequency oscillator of this type is a device of two independent oscillators coupled to each other via dual-mode crystal resonator. Each oscillator comprises a selective filter providing the oscillations to arise only at certain frequency. A block diagram of this oscillator is shown in Fig. 5.

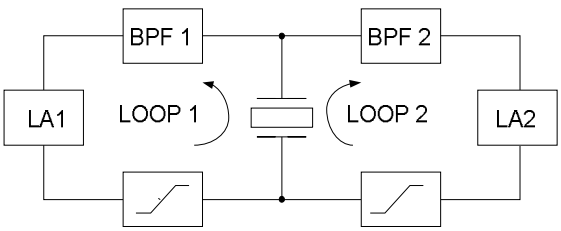


Fig.5. Dual-frequency oscillators with separated nonlinear element block-diagram

Each loop consist of band-pass filter (BPF), linear amplifier (LA), and amplitude limiter

Good mutual filtration of excited modes is an obligatory condition of oscillations stability. With equal oscillations activity the selectivity of separation filters should not be worse than 20 - 30 dB. If one of the modes is weaker than the other one, the stronger mode level at the input of the oscillator exiting the weaker mode should be suppressed still more. In practice, LC-circuits are used as selective elements, their quantity being up to 3 or 4 [2]. In modern construction high technological piezoresonance filters are used [11].

2 .2 Dual-frequency oscillators based on mutually synchronized oscillators

The seeking to get rid of tuned circuits and to obtain a diagram suitable for microelectronic implementation led to an idea to use mutually synchronized oscillators for the sake of filtration. It is known that in two mutually coupled oscillators synchronous oscillations are possible to arise under certain conditions. Synchronism region is determined by the oscillator phase characteristic and by the coupling level. This coupled system exhibits striking properties and can be used for separation of dual-frequency oscillations. I. V. Boudiakov (Russia) suggested a diagram providing the excitation of the SC-cut resonator in B and C-modes. [12].

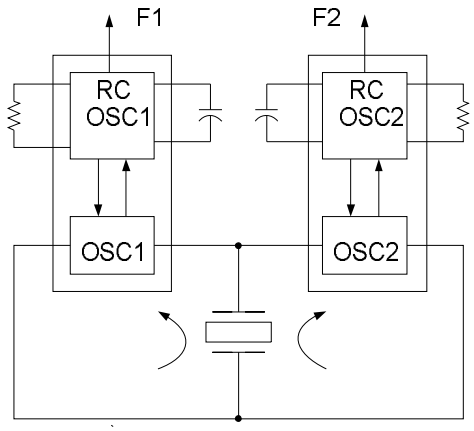


Fig.6. Dual-frequency oscillators based on mutually synchronized oscillators.

The system of mutually coupled oscillators OSC1 - RC-OSC1, OSC2 delivers the frequency of B-mode at the output and the system of oscillators OSC2 - RC-OSC2- delivers the frequency of C-mode at the output. The RC-OSC1 and RC-OSC2oscillators are of RC-type. Their natural oscillation frequencies are the same as the B-mode and C-mode frequencies respectively.
If the modes being excited are of near frequencies, a situation is possible when all four oscillators are acting at the same frequency. In this case synchronization with LC -oscillator may be necessary instead of RC-oscillator synchronization

2 .3 . Dual-mode modulation type oscillators

When using B-mode not for frequency driving but for filtering, the change of its motional resistance shows itself in lesser degree. Pure passive filtering method does not give good thermosensor signal-to-noise ratio and, therefore, is not being used. Combined modulation-adaptation method developed by Y. S. Shmaliy [13, 14] and Y. S. Ivantchenko [15] seems to be more interesting. Here C-mode is excited in simple self-oscillation regime. To obtain B-mode frequency, oscillation Fc is modulated (by amplitude or by frequency) by modulator with oscillation of (Fb-Fc) frequency coming from a certain VCO. In result, Fb frequency component appears in the output spectrum of the modulator. Resonator's B-mode is used as a frequency discriminator controlling VCO frequency. The feedback circuit being closed, VCO frequency always follows B-mode natural frequency, changing with temperature (See Fig. 6)

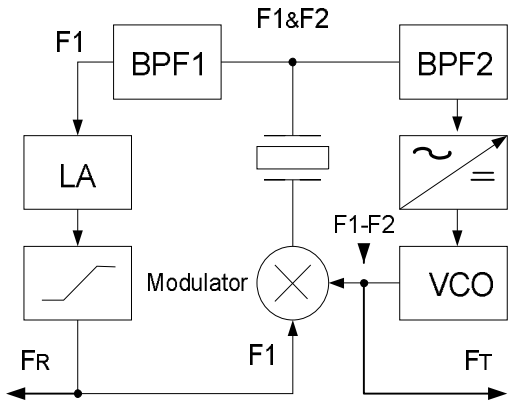


Fig.7. Dual-frequency modulation type oscillator block-diagram.

Adjusting modulation index we can set minimum possible Fb -signal level in the output spectrum, where still surely temperature signals Fb and Fc are formed. This method essentially diminishes requirements to the constancy of B-mode motional resistance, but does not simplify the problem of mutual filtration of oscillations (in particular, of strong Fc -oscillation in the loop for modulating signal formation) It needs to complicate the circuit for their separation

2.4. Keying dual-mode oscillator

On Fig. 8 the seldom-used in frequency stabilization systems circuit of dual-mode oscillator is presented.

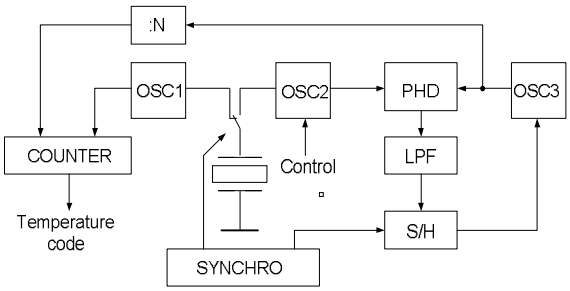


Fig. 8. Dual-mode oscillator with resonator keying

In this circuit the resonator is periodically excited in two independent oscillator circuits in a singe-mode regime: once on a basic mode F_R, and once on a thermosensitive mode F_T. (Oscillators are switched by the key). Obviously, oscillator circuits OSC 1 and OSC 2 should contain any frequency selectivity, allowing exciting the necessary mode. The output oscillator OSC 3 is controlled under frequency of oscillator OSC 2 by circuit of pulse PLL. As the temperature of crystals changes slowly, it is possible to disconnect OSC 2 for short time from the PLL circuit (at this time other mode of crystal is excited and there is a measurement of temperature of a crystal) freezing the operating influence by the sample and hold (S/H) circuit. Advantage of the circuit consists that oscillators are excited in single-mode regime that reduces a level of collateral oscillation in output signal. Such circuit is especially interesting to construction of high-frequency oscillators on the basis of frequency synthesizers, about such as [3].
Historically this circuit was offered one of the first [16] and in the original was used with resonator of the AT-cut with excitation as a temperature mode various low-frequency oscillations (flexure, extension, and shear). Certainly, on its basis can be excited resonators of other cuts.

2.4. Automatically balanced modes, asynchronously excited dual-frequency oscillators

The dual-frequency oscillation methods considered above, regime rather complicated excitation and discrimination circuitry. At the same time there is a possibility to excite and maintain two asynchronous oscillations in single-transistor oscillator. V. Anisimov in his theoretical paper [17] proved the existence of this possibility provided that an amplifier element non-linearity is expressed by 5-th degree polynomial:

$$i = \alpha \cdot u + \beta \cdot u^3 + \gamma \cdot u^5$$

and there is a non-zero solution to the system of equations describing oscillation process in the oscillator. If there is a self-bias circuit and a cut-off angle is less than 90^0, such type of system can produce mild asynchronous self-oscillation.

It is very important that the area of stable dual-mode oscillation is very narrow and depend strongly on circuit regime and resonator parameters. During some years we are worked on improvement both circuit and resonators. For automatically balanced near the optimal point of dual-mode excitation it is proposed to use a special feedback circuit [18,19]. This circuit extracts the low-level component of (F_b – F_c) frequency from output signal, make a necessary amplification of this signal, and adjust the DC regime of oscillator so to maintain a amplitude of this low-frequency signal at low, but stable level.

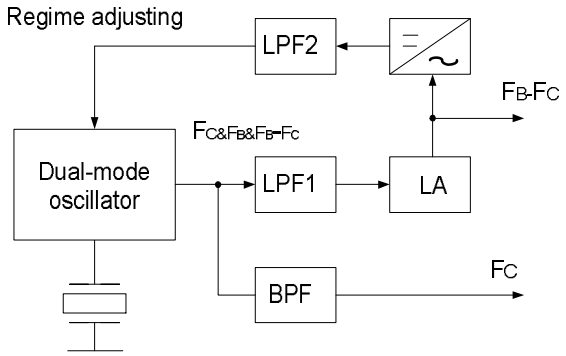


Fig. 10. Oscillator with automatically balanced mode level.

The method of simultaneous excitation of two asynchronous oscillations in one self-oscillator is, from our point of view, the most economic and attractive and, therefore, all our efforts were aimed at its development.

CONCLUSIONS

On the basis of above-stated it is possible to draw the following conclusions:

1. From the point of view of accuracy characteristics in dual-mode thermocompensated and ovenized oscillators it is desirable to use a resonators of SC-cut with excitation of B and C modes.
2. The resonator design should be exposed to special optimization for the purpose of suppression of splashes in motional resistance of a B-mode (activity failures).
3. The great attention should be given to the oscillator circuit, allowing carrying out the stable oscillation of two frequencies in width temperature interval.

The decision of complex of these questions must increase temperature stability of oscillator frequency essentially.

REFERENCES

1. M.Bloch. M.Meirs, J.Ho. The Microcomputer Compensated Crystal Oscillator (MCXO). Proc. 43-rd A.F.C.S.,1989, p. 16-19
2. A. Benjaminson and S. C Stallings, A microcomputer-compensated crystal oscillator using a dual-mode resonator, Proc. 43rd Ann. Frequency Control Symp,, 1989, pp. 20-26
3. A. Kosykh, V. Bagaev, B. Ionov, A. Lepetaev, S. Zavialov, A. Vasiliev. The New Method of Statistic Piecewise-Linear Interpolation and its Application to DTCXO Creation. Proc. 47-th A.F.C.S., 1993
4. Y. Watanabe, K. Ozaki, S. Goka, T. Sato, H. Sekimoto. Frequency Stability and Phase Noise Characteristics of Digital OCXO Using Dual-Mode Excitation. Proc 2001 IEEE FCS and PDA Exhibition, pp. 790 – 79
5. Kusters J.A., Leach J.G., Firher M.S. Quartz Resonator Cut to Compensate for Static and Dynamic Thermal Transient. US Patent N 40792801978
6. A. Kosykh, I. Abramson, V. Bagaev. Dual-mode Crystal Oscillators with Resonators Excited on Band C Modes. Proc. 1994 AFCS, pp. 578 – 586.
7. Y. Shmaly, S. Nedorezov, R. Besson. Statistical Model of an Anharmonic Sensor of Precision Crystal Oscillator. 2001 IEEE IFCS & PDA Exhibition. Pp758 – 766.
8. S. Shodovski. Resonator Self-Temperature-Sensing Using a Dual-Harmonic-Mode Crystal Oscillator. Proc. 1989 AFCS, pp. 2 – 7
9. E. Jackson, H. Phillips, B. E. Rose, The microcomputer compensated crystal oscillator - a progress report, Proc. 1996 IEEE Intl. Freq. Contr. Symp., pp. 687-692, 1996
10. I. Abramson. Two-Mode Quartz resonator For Digital Temperature Compensated Oscillator. Proc. 1992 AFCS, pp. 442 – 447.
11. Y. Watanabe, H. Sekimoto, S. Goka, I. Nimi, A Dual-Mode Oscillator Based on Narrow-Band Crystal Oscillators with Resonator Filters, Proc. 1997 IEEE Int'l Frequency Control Symp,. 1997, pp. 932-937
12. Будяков И.В., Липко С.И.. Прокопенко В.Г. Основные структуры двухмодовых кварцевых генераторов. (The Base Structures of dual- mode oscillators) /Д.С.Л. ВИНИТИ N 5040- В86 от 14.07.86
13. Ю.С.Шмалий, Ю.И.Евдокименко. Модуляционный метод адаптации опорных кварцевых генераторов (Modulation Method of Adaptation of Reference Crystal Oscillators). В кн. Стабилизация частоты. Часть 1, ВИМИ, стр. 75-78. (USSR)
14. Y. Shmaly. The Modulation Method of Quartz Crystal Oscillator Frequency Stabilization. IEEE Trans. On Ultrason., Ferroel. And Frequency Control. Vol. 45, pp 1476 – 1484. Nov. 1998.
15. USSR Invention Certificate N 1709492. Устройство термокомпенсации кварцевого генератора (Crystal oscillator thermocompensation device). Ю.С.Иванченко, В.Ф.Самойленко, И.И.Буймала //Открытия, изобретения, 1992, N 4.
16. R. Rubach, Dual-mode digitally temperature compensated crystal oscillator, Proc. 36th Ann. Frequency Control Symp., 1982, pp. 571-575
17. V. V. Anisimov, "On biharmonic oscillations excitation in an oscillator with two degrees of freedom," Vestnik MGU, ser. physics, astronomy, 1956, n. 1, pp. 137-146 (in Russian).
18. Patent N 1145450 Thermocompensated Crystal Oscillator. /V. Samoilenko Бюллетень №10 1985 (USSR)
19. Patent N 1573524 (USSR) Двухмодовый кварцевый генератор (Dual - mode Crystal oscillator /В.П. Багаев, С.А. Завьялов., Бюллетень изобретений № 23, 1990 г

SUB NANOSECOND CALIBRATION OF REGIONAL TIME LINKS USING CONTINUOUS GNSS CARRIER PHASE OBSERVATIONS

Carsten Rieck, Per Jarlemark , Kenneth Jaldehag, Jan Johansson
SP Swedish National Testing and Research Institute, Box 857, S-501 15, BORÅS, Sweden
Phone: +46 33 165440; Fax: +46 33 125038; e-mail: carsten.rieck@sp.se

Abstract - **The calibration of GPS or TWSTFT links is an important requirement for accurate time transfers. A simple method commonly used is to transport a clock between the two ends of the link and perform time interval measurements between the carried clock and the respective laboratory clock. As the main disadvantage one can see the difficulty to model the behavior of the carried clock during the transport. An accurate prediction usually requires highly stable clocks to be transported with all the risks and costs involved. One solution to this is to estimate the carried clock's drift during transportation relative to a fixed clock using precise GNSS observations.**
This paper describes the method of a GPS aided clock transport experiment with the aim of calibrating the GPS-P3 time link between SP Swedish National Testing and Research Institute and Onsala Space Observatory, which are separated by a distance of about 70 km. In a first attempt data from a local round trip transport from and to SP were analyzed in order to test and verify the method. A moderate rubidium oscillator was connected to and transported together with a geodetic GNSS receiver system. The collected data were post-processed with a common view GNSS code and carrier phase filter yielding precise differential time series between the rubidium clock and the reference clock at SP. These results were used to link the two time interval measurements taken between the clocks at the beginning and the end of each transport together. An achievable sub nanosecond precision of the link calibration is of value for applications requiring highly accurate time synchronization such as distributed time scales.

Keywords - **GPS, carrier phase, time, frequency, temperature, delay, accuracy**

I. INTRODUCTION

Clock transportation is a historically proven time transfer method. However, with the dawn of satellite based transfer methods it has lost its practical use a long time ago. In order to match for instance the (few) nanosecond capability of a calibrated GPS Common View (CV) time transfer it is necessary to transport an expensive frequency standard in environmentally stable condition. A long transportation is risky for the clock and the longer the duration of the transport the more difficult it is to predict the clock behavior during that time. It is obvious that the accuracy of the prediction of the clock behavior during transport sets the ultimate limit of the accuracy of the time transfer. Being able to measure the transported clock instead of predicting it would ease some of the difficult aspects of the clock transportation method:

- the use of inexpensive and small clocks for transportation,
- no need for environmental conditioning of the used clock,
- the duration of the transport is not critical as long as the transported clock can continuously be measured.

One solution to the clock measurement problem during transport is the use of precise GNSS carrier phase observations (CP).

II. METHODOLOGY

Figure 1 depicts the principle of the GNSS aided clock transport method. All involved GNSS receivers track all satellites in view during the whole experiment. Three phases can be distinguished: pre-, transportation- and post-phase. During pre- and post-phases the transport clock and its receiver system are stationary. In the pre-phase the reference clock is compared with the transport clock by using time interval measurements. In the post-phase this comparison is done with the calibration clock at the other site. During the transport-phase GNSS CP observations are used to track the drift of the transport clock versus the reference clock. This makes it possible to link the time interval estimations from the pre-phase to those of the post-phase.

Our current CP processing can be summarized with the following points:

- use of GPS observation and navigation data from at least two Javad GGD geodetic receiver systems (REF, (CAL) and ROVER), use of IGS final orbits [1],
- fixing of integer cycle differences at L1 and L2 for all satellite observations for the pre-phase and for all cycle slip occasions (Figure 4) for the duration of the experiment, ambiguity elimination is aided by code data and sequential clock drift prediction,
- estimation of antenna orientation and correction for the apparent clock drift caused by antenna rotation and circular polarization of the carrier signals [2],
- Kalman filtering of the ionospheric "free" linear combination L3 with state variables for differential clock drift, East, North and Vertical coordinate components and neutral atmospheric delay for the rover,

no states for phase ambiguities needed, use of IGS orbits and a simple earth-tide model [3], [4], [5],

- "Rauch-Tung-Striebel" smoothing of forward Kalman series results in a time series describing the differential clock drift between reference oscillator and transport oscillator [3].

The GPS measured clock drift has an arbitrary offset to the "absolute" TIC measurements at the respective sites, i.e. a reference-offset x and a calibration-offset y. The GPS based drift estimates between the reference and the transportation clocks are "connected" to the TIC values by using a fit in the overlapping interval providing x. If the calibration site was also equipped with a GPS receiver then the same procedure can be done for the post-phase TIC measurements. Here the GPS based drift between the transported clock and the calibration clock is fit to the respective TIC measurements of the post-phase resulting into the calibration-offset y. In this way a time transfer between the reference clock and the calibration clock, $REF(t_0)$-$CAL(t_0)$, can be easily achieved:

$$REF(t_0) - TRANS(t_0) = GPS_{REF}(t_0) + x$$
$$CAL(t_0) - TRANS(t_0) = GPS_{CAL}(t_0) + y$$
$$REF(t_0) - CAL(t_0) = GPS_{REF}(t_0) - GPS_{CAL}(t_0) + x - y$$

GPS_{REF} and GPS_{CAL} denote the respective differential GPS time series between reference/calibration clocks and the transportation clock. The reference time t_0 can with advantage be set at any time during the experiment.

In case that only the reference clock is used to produce a GPS based drift time series, then estimates of $REF(t_0)$-$CAL(t_0)$ are only possible for the post-phase:

$$REF(t_0) - TRANS(t_0) = GPS_{REF}(t_0) + x$$
$$CAL(t_0) - TRANS(t_0) = TIC_{CAL-TRANS}(t_0)$$
$$REF(t_0) - CAL(t_0) = GPS_{REF}(t_0) - TIC_{CAL-TRANS}(t_0) + x$$

$TIC_{CAL\text{-}TRANS}$ denotes the TIC measurements during the post-phase of the experiment.

III. Results

In a first approach we have run an experiment comparing a local reference clock, UTC(SP), with a Rubidium clock, SP_TRANS, which was transported around urban and highway environments for 55 min and a length of about 75 km, see Figure 2-5. The outline of the track gave valuable information about the behavior of the used GPS equipment in a kinematic application. Figure 6 shows part of the experimental setup used. The TIC, which normally is part of the rover kit, was left in the lab environment since in this case study the reference clock and the calibration clock were the same. It was therefore possible to evaluate the performance of the method. Figure 7-9 present results from this first experiment. In Figure 9 we estimated an error of about 0.6 ns for the round trip time transfer.

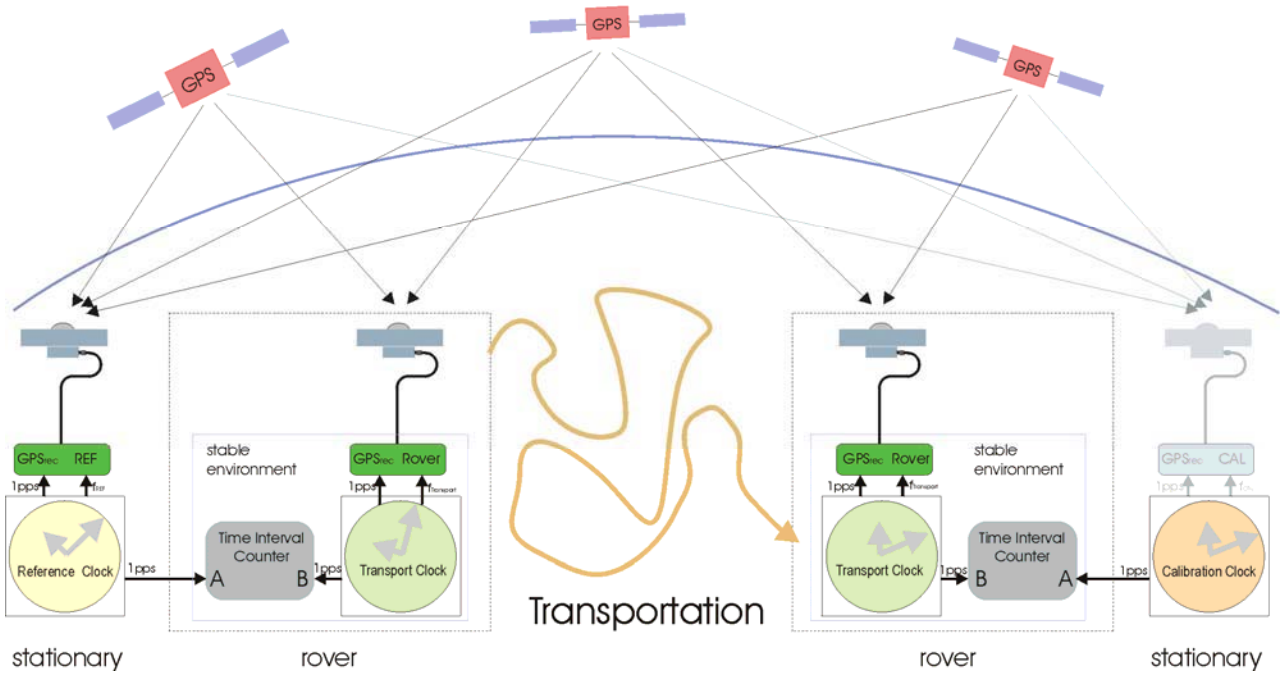


Fig. 1. A principle view of the GNSS aided clock transport method. In order to evaluate the performance of the approach in our experiment the reference clock and the calibration clock are the same clock (UTC(SP)). The Time Interval Counter (TIC) is part of the rover equipment, which together with the GPS receiver should be held in a constant environment.

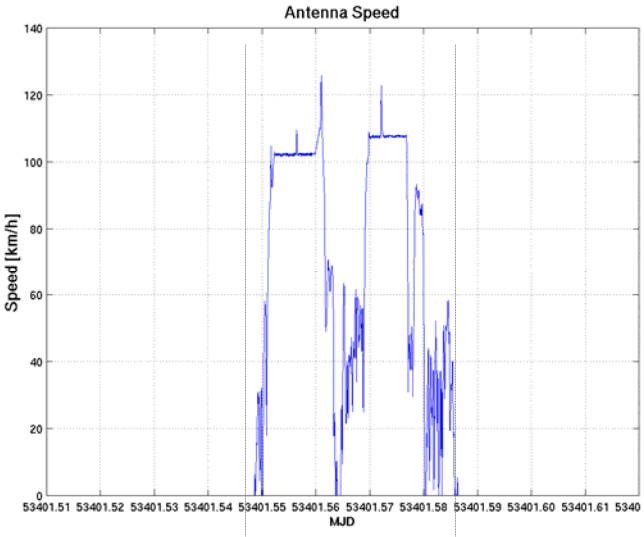


Fig. 2: Speed of the vehicle (antenna) during the experiment derived from GPS code measurements. In order to test a representative transport, both urban and highway environments were included into the track.

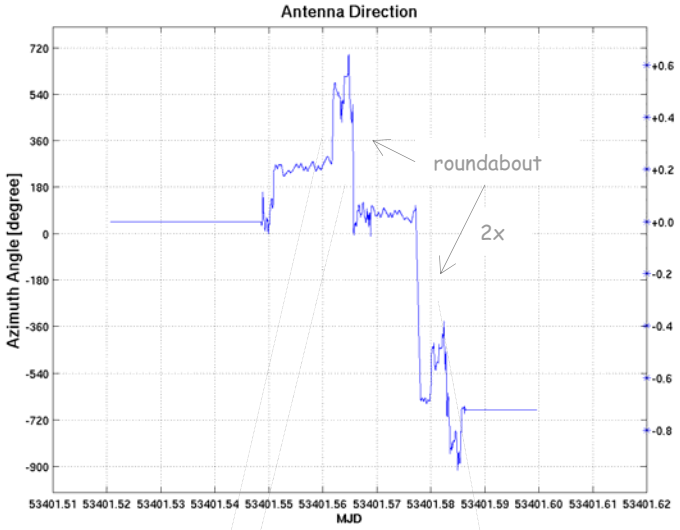


Fig. 3: Knowledge about the relative antenna orientation at start and stop including the full number of antenna rotations during the transport are important for a correct time transfer with carrier phase GPS. The right hand axis tics show the feed-rotation compensation in ns.

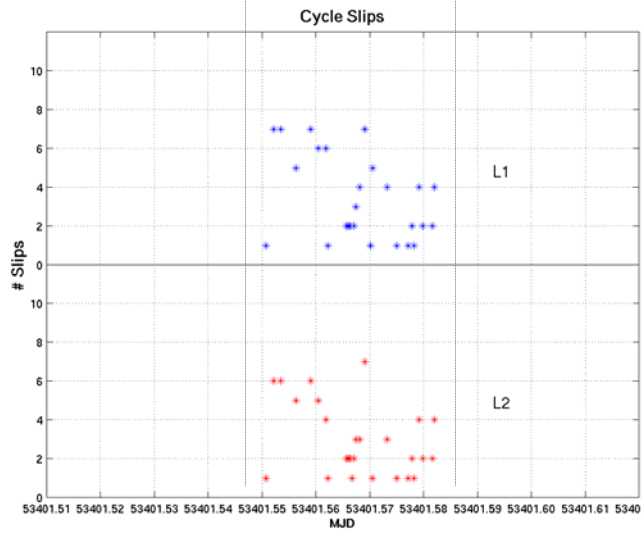


Fig. 4: Number of cycle slips during the experiment. Cycle slips only occurred during the transport phase and are due to obstructions, vehicle acceleration in combination with incorrect receiver settings and variations in the immediate surroundings of the vehicle.

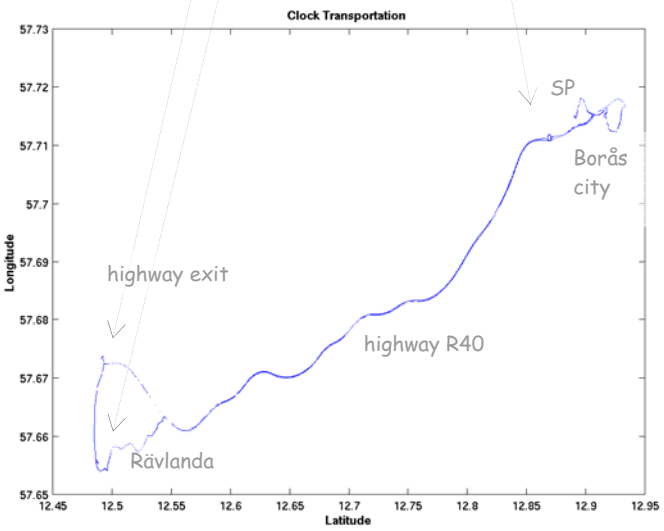


Fig. 5: Latitude-Longitude plot of the track chosen. The track had a length of ca 75 km and the transport lasted about 55 min. Clearly visible are the highway exit and the roundabout traffic near SP. The graph is derived from GPS code measurements.

IV. Conclusion

We have developed a time transfer method with potential sub nanosecond accuracy, which is based on a GPS carrier phase aided clock transport. This first local experiment shows promising results. But there are some issues to be considered in order to utilize the full potential of the method described.

Even though a *stable environment* is not really needed for the transported clock, the receiver chain and the TIC need to be held in an as constant environment as possible. For instance all parts of the receiver chain exhibit a thermal signal delay dependence that limits the achievable accuracy of the measurement. Geodetic GPS receivers can have thermal coefficients of several hundreds of picoseconds per Kelvin change of ambient temperature. Other, partly correlated, environmental parameters are for example humidity, pressure and the quality of the circuit Voltage for receiver and antenna. [6] investigates the thermal influences on the GPS receiver chain. For example the used Javad GGD receiver has a measured thermal coefficient of –170 ps/K at the L1 carrier frequency, which is a relevant error source for this kind of application, see Figures 10 and 11.



Fig 6: Experimental setup: A Dorne Margolin choke-ring antenna was mounted on top of a automotive vehicle in order to offer the best possible view of the sky. The antenna was directly connected to a geodetic two frequency GPS/Glonass receiver (Javad GGD) that uses the external frequency (10MHz) of a Rubidium oscillator (Symmetricon X72, SP_TRANS). All parts were DC power supplied by a sealed lead-acid battery. The reference- and the calibration clock are the same physical clock (UTC(SP)), which is situated in SP's time laboratory. A portable computer was used to store the data from the receiver. For a real GPS aided clock transport with different clocks for REF and CAL a TIC is needed to be transported as well. During this experiment a Pendulum PM6681 within the same lab environment as the REF/CAL clock was used. In the future an air-conditioned container will house all the equipment providing stable thermal conditions. The stationary GPS receiver at REF, SP01, was also a Javad GGD using the IGS/SWEPOS antenna point SPT0 [1].

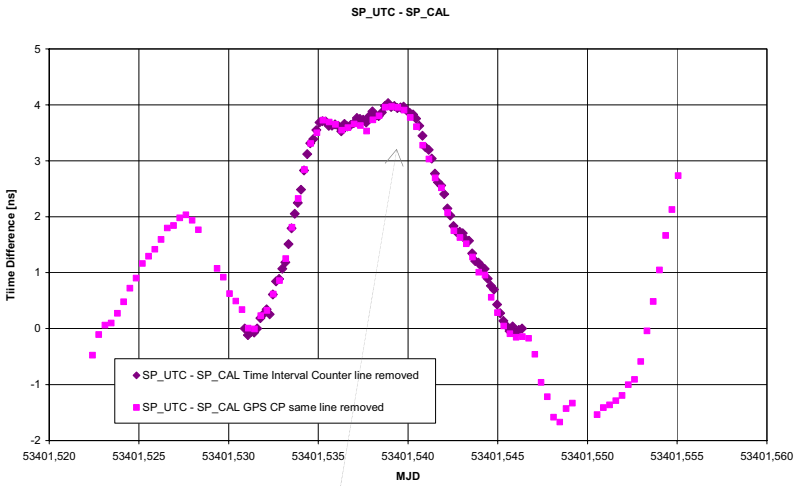


Fig 7: The rover is in a fixed position before the start of the transportation phase. In the graph a linear drift calculated from the TIC measurements is removed from both the GPS CP and the TIC values. GPS CP is aligned to the TIC by least squares yielding a fix REF-offset *x*.

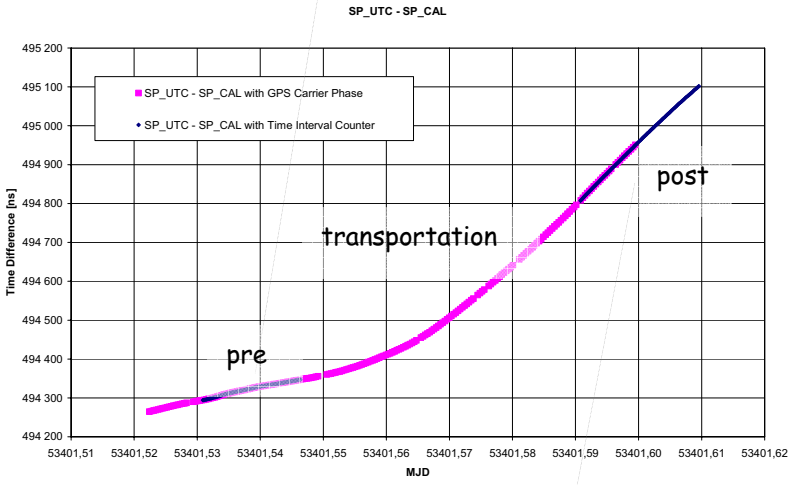


Fig 8: The result of the experiment: In blue, the TIC values at the beginning and the end define the "absolute" offset of the Rubidium oscillator SP_TRANS to UTC(SP). During those times the rover was in a fixed position (Figure 7 and 9). The pink graph shows the GPS CP measured drift during transport aligned to the TIC by means of a least squares fit.

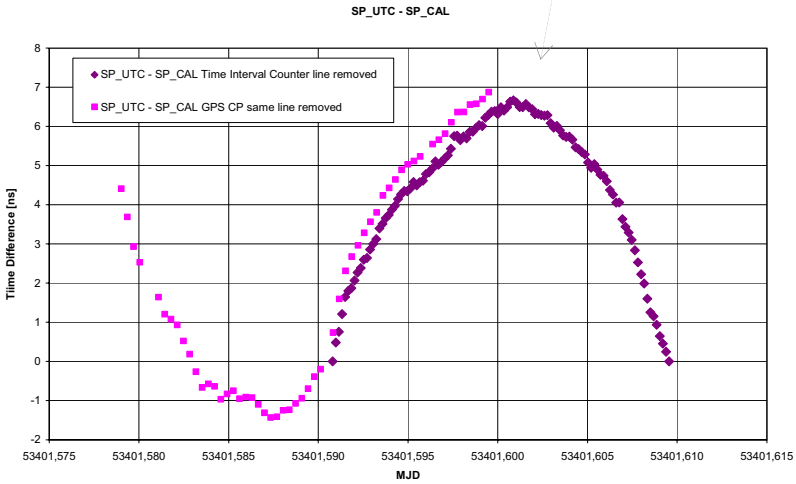


Fig 9: The rover is in a fixed position after the stop of the transportation phase. Again a linear drift calculated from the TIC series is removed from both the GPS CP and TIC values. GPS CP is offset by the REF-offset *x*. The mean difference between TIC and GPS CP is about 0.6 ns.

Most of the error of 0.6 ns in this experiment was likely due to environmental changes around the receiver.

GNSS observations are dependent on the *free line of sight* between antenna and satellite. Any interruptions require a numerically intensive recovery of all integer ambiguities. Very long interruptions, e.g. tunnels, can under certain circumstances make it impossible to track the transported clock. Unfortunately it is impossible to avoid interruptions in most of the applicable environments. For example crossing underneath highway bridges causes special attention. Due to the relatively high antenna velocity and that re-locking of satellites is a time consuming process, most of the satellite observations are lost at the same time. Thus the rover receiver settings should therefore be optimal for kinematic applications.

A *fast sampling*, in the order of seconds, of the GPS observables is advantageous for cycle slip recovery and antenna orientation calculations. The same sampling rate should be used at the stationary REF receiver and for the TIC. TIC measurements should overlap the GPS estimates in order to provide a good fit.

Multipath is a considerable error source for both code and carrier phase. The antenna mounting on top of a metal car roof is especially critical for carrier phase measurements. Reduction of multipath can also help the process of ambiguity resolution.

It would be possible to reduce the instrumental setup by the transportation oscillator. Short experiments can be done by using the internal receiver oscillator instead. This requires, among other things, to keep the receiver from resetting its internal time-base on a regular bases.

Generally, one can claim that GPS carrier phase time processing has a potential geodetic accuracy of a couple of tens of picoseconds, i.e. the influence of differential channel noise. Nevertheless, unpredictable and individual receiver hardware delay variations reduce the achievable accuracy dramatically and it is difficult to predict its lower limit. In this clock transport application a careful chosen environment and good planning certainly allows accuracies in the order of a couple of hundred picoseconds.

V. FUTURE WORK and THANKS

In a next step we first of all intend to improve the experimental setup by stabilizing the receiver environment and the TIC. Together with the transportation clock all equipment will be mounted into an air-conditioned housing. The use of phase-stabilized cables (>20m) for the TIC measurements to the REF/CAL clocks will improve the uncertainty for the indoor/outdoor interface. Improvements to the immediate multipath environment around the antenna and optimization of the receiver settings for kinematic use will help better tracking and thus ease the post processing. The software also needs certain improvements for automation and data handling, the use of GLONASS data will be included in order to support a more robust solutions.

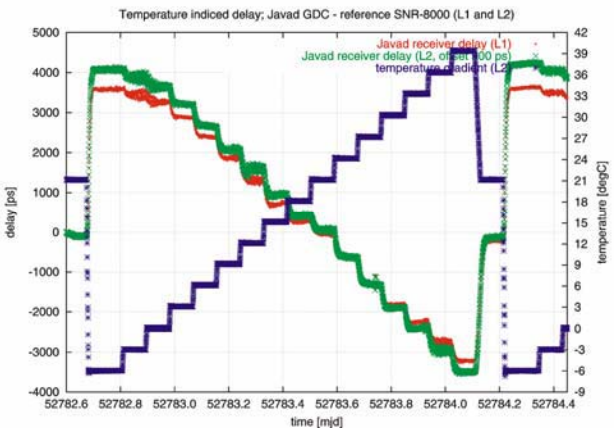


Fig. 10. Javad Lexon – GGD receiver, temperature induced delay show differences between the carrier frequencies -170 ps/K at L1, -150 ps/K at L2, measurements in zero baseline to a SNR-8000.

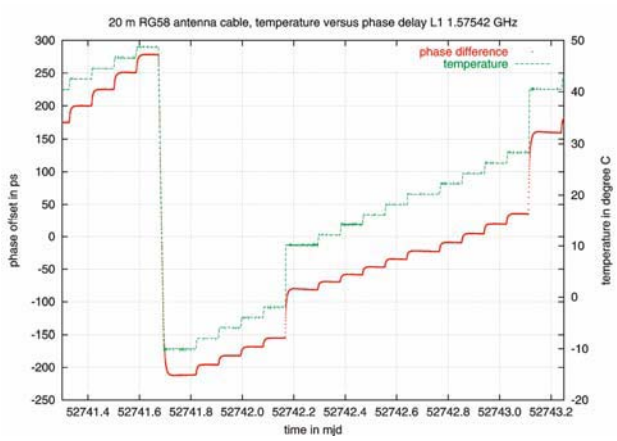


Fig. 11. Temperature induced delay on an Rg58 antenna cable at the L1 frequency, nearly linear behavior over the temperature range: 0.42ps/Km; comp. phase stable Andrews LDF2-50: 0.01ps/Km.

We would like to thank the Swedish Defense Material Administration (FMV) and Onsala Space Observatory for loan of equipment, and Peter Löthberg, STUPI AB, for fruitful discussions and providing data.

REFERENCES

[1] http://igscb.jpl.nasa.gov.org/

[2] Wu JT, SC Wu, GA Haij, WI Bertinger, SW Lichten, 1993 "Effects of Antenna Orientation on GPS Carrier Phase" Manuscripta Geodaetica 18, 91-98

[3] RG Brown, PYC Hwang, 1997 "Introduction to Random Signals and Applied Kalman Filtering" Wiley & Sons, New York.

[4] A Niell, 1996, "Global Mapping Functions for the Atmospheric Delay at Radio Wavelengths" J. Geophys. Res. 101, 3227-3246

[5] B Hofmann-Wellenhof, H Lichtenegger, J Collins, 2001 "Global Positioning System, Theory and Practice" 5th rev. edition, Springer-WienNewYork

[6] C Rieck, POJ Jarlemark, K Jaldehag, J Johansson, 2003 "Thermal Influence on the Receiver Chain of GPS Carrier Phase Equipment for Time and Frequency Transfer" Proceedings IEEE FCS/PDA/17th EFTF

TIME METROLOGY IN ROMANIA

Anca Niculescu
National Institute of Metrology, Sos. Vitan Barzesti no. 11, 042122 Bucharest, Romania
niculescu@inm.ro, fax: (+4021) 334 53 45

ABSTRACT

Since 1966 the National Institute of Metrology Bucharest (NIMB) has been responsible for the Romanian Time and Frequency Standard. In the beginning the national standard relied on two quartz oscillators Rhode & Schwarz. Between, 1980-1984, in order to satisfy the increasing requirements of the economy and technology, an atomic commercial frequency standard (HP 5061A) was used as national standard for time and frequency. This latter was replace, further on, by two hydrogen masers build up by the specialists of Atomic Horology Laboratory of the National Institute of Laser, Plasma and Radiation - INFLPR (dr. Octav Gheoghiu, dr. Cipriana Mandache, dr. Liviu Giurgiu and others). For the moment, the local realization of UTC (the national time scale) relays on one cesium clock HP 5071A.
This paper is a report on the present metrological parameters of the national standard and on the methods which we intend to use to disseminate the data/time information at the national scale.

1. SHORT HYSTORY OF TIME MEASUREMENTS IN ROMANIA

In early years the time measurement was based in astronomical measurements. In our country the fast time measurement system was discover at Gradistea Muncelului. It is known as the great Dacian sanctuary and it is similar with other sanctuary found in England, China and India. The exactitude of time measurement it was one of the better for this moment.
Later on the time measurements in our country followed the European trend. In 1891 by royal decree was adopted the official time as time of the second meantime zone and later on, in 1957, was set up the first Time Service at the Astronomic Observatory of the Academy and the first data was include in BIH bulletins in 1958. At this moment the national time standard was a Belin quartz horology. Since 1961 the Astronomic Observatory was included in the international system. In 1951 was set up the Institute of Metrology which became, in 1974, the National Institute of Metrology as it is know today. Concomitantly it was set up the time and frequency laboratory. At the beginning, the only time and frequency standards had been two marine chronometers and two Rhode & Schwarz quartz horology. Hardly, in 1975, was buy the first atomic clock which represent an important jump in the evolution of the metrological. In
80 years the time and frequency laboratory had using three hydrogen masers, manufactured by the National Institute for Laser, Plasma and Radiation (INFLPR). Because of their poor reliability they were never used as national time and frequency standards.

2. THE PRESENT SITUATION OF TIME AND FREQUENCY IN ROMANIAN METROLOGY

The National Institute of Metrology Bucharest (NIMB), has been working in this metrological field for more than 50 years and it is responsible for Romania of the realization and dissemination of SI units for nearly all-physical quantities. The Time and Frequency Laboratory of NIMB establishes, since 1966, the reference frame for a national time scale trough out the whole territory of Romania. The national time and frequency standard independently reproduce the national unit of time, generates and keeps the national time scale *UTC* (NIMB) so it is ensuring the unification of time and frequency measurements.
In present, the national time scale is basing on a commercial cesium frequency standard HP 5071A (put in work in 1995), since October 2002 we began to contribute to the international time scales TAI and UTC established by BIPM.

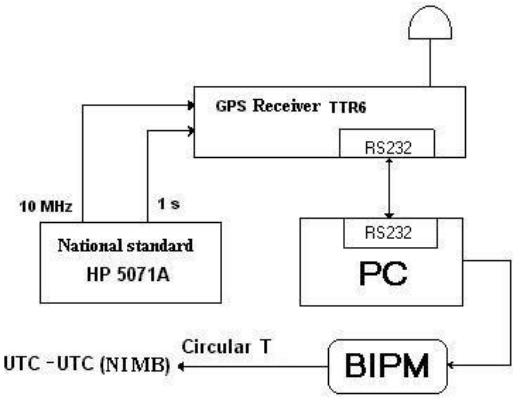


Fig. 1 Block scheme of comparison system

Initially the comparisons were set up using a GPS receiver manufactured by Odetics, but this receiver had not been fully compatible with the method used by BIPM. The comparisons were performing uninterrupted by self-activating, every hour, via an HPIB interface

using a PC and dedicated software. The uncertainty evaluation of the difference between *UTC*(NIMB) and *UTC* was very difficult. It must take account, as much as possible, of all error sources. Therefore the national time scale *UTC*(NIMB) was maintained with a uncertainty grater than 100 ns and this value is entirely due to the method used. Later on, from May 2003, the comparisons had been done using a GPS receiver, type TTR-6, which permit us to improve the uncertainty of the comparison results (Fig.1). The aim of maintaining the synchronization of *UTC*(NIMB), within 100 ns, to *UTC*, as recommended by the CCDS, had been achieved at 1σ level. Fig.2 shows the results of the comparison between the national time scale *UTC*(NIMB) and the international time scale *UTC* as they arise from data published in Circular T of BIPM [1].

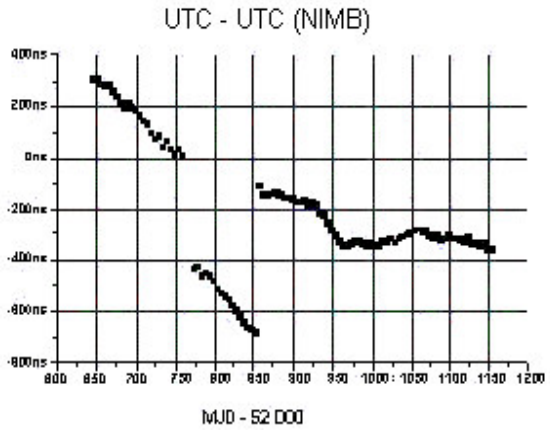


Fig. 2 Comparison of *UTC* (NIMB) vs. *UTC* via BIPM

The two jumps that appear on the graphic were induce the first by putting in work of the GPS receiver TTS 6 and the second one by the major failure of the power supply.

The future developments foreseen in the activities concerning the national time scale realization will follow two directions:

- the improving the reliability in real time generation of and the stability of realization of *UTC*(NIMB) and the stability of *TA*(NIMB) by including a new commercial cesium frequency standard, manufactured by Agilent, in the national standard. Since March 2005 the NIMB time scale are routinely compared to UTC by means of the GPS common-view comparisons with one C/A code receiver TTS 2 which was recently put in operation.
- the development of time dissemination services transmissible to all users.

In the future, the new primary frequency standard (cesium atomic fountain) that will be built up within the frame of cooperation between INFLPR and NIMB will improve the metrological performance of the national time and frequency standard and will increase the accuracy of national time unit [2, 3, 4.].

3. TRANSMITION OF TIME INFORMATION

The transmission of time information at national scale is also an important part of the NIMB activity. The development of electronic commerce and of the secure electronic document exchange also requires that a trusted data and time information is included as part of digital signature of the documents exchanged [5]. To meet the requirements of the Time Stamping Authorities (TSA) that are going to be established in Romania, than NIMB must implement a trusted time distribution service for the remote certification of these authorities that must have their time sources traced to *UTC*(NIMB) within 1 s. The approach proposed by NIMB assumes that a TSA will have some sources of data and time synchronized e.g. by GPS, DCF77, to have the necessary redundancy, and will call some times a day the NIMB trusted time service by a secure dial-up connection or INTERNET, to have a certification of its time information traceable to UTC. With the solution proposed, NIMB will be able to measure the rate of TSA clocks and to provide a certification of their time error. As follow, recently, we realized and put into provisional operation a system that allows the transmission of time information via telephone lines. The system consists of one time distribution system (TDS), manufactured by the Technical University of Graz (Austria), a telephonic modem and a PC together with special software. This dissemination service, designed to synchronize computer clocks or digital terminal, is capable to reach accuracy ranging from low to medium (100 ms to 1 ms), as reported in the papers dealing with synchronization systems [6]. The experimental results were quite satisfactory the timing accuracy obtain ranging from 400 ms (for CMOS clocks) to 60 ms (for DOS clocks).

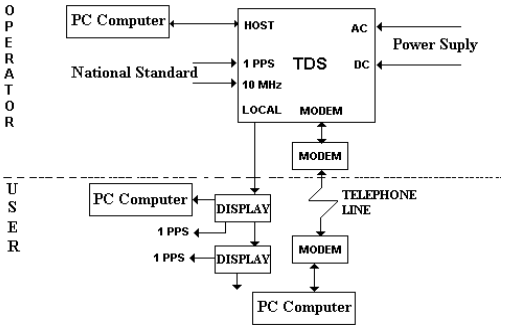


Fig. 3 Time distribution system concept using TDS

4. CONCLUSIONS

This paper want to demonstrate the "state-of-the-art" of Romanian Time and Frequency National Standard regarding its metrological parameters and both external and internal traceability. The time and frequency activities performed at NIMB cover most of the

demands coming from the research and user communities. Some other activities will be more significant in the future for Time and Frequency Laboratory of NIMB:

- to improve the accuracy and stability of the national time and frequency standard;
- to develop our frequency stability measurement capability, with particular attention to short and medium averaging time (1…1000 s);
- to realize a dissemination service able to provide a trusted time reference for the future activities that will use electronic exchange of information.

REFERENCES

[1] BIPM Circular T (monthly), available on-line from www.bipm.org
[2] C.Mandache, T Acsente, M. Bengulescu, Research on cesium cold atoms in Bucharest, Romania, Roum. Journal of Physics, vol. 42, nr.7-8, pp. 501-507, 1997
[3] C. Mandache, T. Acsente, M. Bengulescu, Cesium magneto-optic trap in Romania, Proc. Of the 11th EFTF, Neuchatel, Suisse, pp. 605-609, 1997
[4] T.Acsente, A. Niculescu, C. Mandache, P.Thomann, Laser cooling of alcaly atoms for time metrology application, Proc. of the Fifth General Conference of the Balkan Physical Union, pp981-984, 2003
[5] F. Codara, M. Mascarello, V. Pettiti, G. Vizio, R.Costa, Trusted time service for Italian time stamping authorities, Proc. Of the 14th EFTF, Torino, Italy, pp. 403-407, 2000
[6] D.Kirchner, Genaue Zeit für rechner über telephonmodems, Proc. Mikroelectronik 89, Springer-Verlag, 1989, p.103-108

IONOSPHERIC EFFECTS AND MODELING ON GPS TIME TRANSFER

Hsin-Min Peng, Fang-Dar Chuu and Chia-Shu Liao
National Standard Time and Frequency Lab, Telecommunication Laboratories,
12, Lane 551, Min-Tsu Road Sec.5, Yang-Mei, Taoyuan, Taiwan, 326, R.O.C.
E-mail: hmpeng@cht.com.tw, Tel: 886 3 424 5174, Fax: 886 3 424 5474

ABSTRACT

In this paper, the ionospheric effects on GPS time transfer and an improved model for this correction are studied. It is known that the stability of time transfer using GPS common-view technique in low latitude region, such as in Taiwan, is worse than that in other region. This phenomenon is mainly caused by the great ionosphere fluctuation within ± 20º of the equator. For demonstrating this effect, the ionospheric corrections obtained from broadcast ephemerides, global ionosphere map (GIM) and from actual GPS measurements are computed and the GPS common-view data from two timing laboratories are analyzed. Based on our tests, the time comparison result is degraded by using the GIM correction in low latitude region. On the other hand, using the measure ionospheric delay, the time transfer result in low latitude region is as good as that in high latitude. To compensate the differences between GIM and the measure ionospheric delay, the higher-order ionospheric effects are tested and the satellite-dependent model is proposed. Using this new model, the ionospheric effects can be reduced; furthermore, this model can be adopted in local region to improve the frequency stability when only single frequency GPS receiver is used in frequency comparison.

1. INTRODUCTION

Using GPS code measurements for time comparison is studied for more than twenty years. In the conventional common view structure, only one channel C/A code observables are adopted and the most significant error is the ionospheric delay. Since there are no measurements for computing the effect, we only have to estimate the delay from the ionospheric model of broadcast ephemerides and the error only can be compensated by about 50% [1].

To overcome this drawback, two typical methods are adopted. This first one is using the dual frequency GPS receiver to obtain the P1 and P2 code measurements and the ionospheric delay can be determined by the linear combination of these measurements [2]; however, this kind of receiver is usually more expensive than a single frequency receiver. The other method is using the IGS global ionosphere map products. This product supports the global total electron content (TEC) values and then the ionospheric delay can be computed with an adequate model. Although the delay computed by the IGS GIM product is more accurate than that by the broadcast model, some obvious errors still exist. For reducing the errors, a modified model based on nonlinear model decomposition is proposed and some tests are performed to evaluate the performance of our model. Based on the test data, we can conclude that our model is sound and effective, and can be easily applied to the surrounding stations of the reference station.

2. REVIEW OF IONOSPHERIC MODELS

Two typical ionospheric models are described as follows and are implemented in our test program. The first one is the GPS broadcast model and the other is the procedure to use the GIM TEC values for computing the ionospheric delay.

2.1. Broadcast Model

The broadcast model, also called as the Klobuchar model, represents the zenith delay as a summation of a half-cosine function in the daytime and a constant value at nighttime. For reducing the computation load, the first three terms of Taylor series expansion for cosine function is used in the following equation.

$$T_{iono} = \begin{cases} F \cdot \left[5.0 \cdot 10^{-9} + (AMP)\left(1 - \frac{x^2}{2} + \frac{x^4}{24}\right)\right], & |x| < \pi/2 \\ F \cdot (5.0 \cdot 10^{-9}), & |x| \geq \pi/2 \end{cases}, \quad (1)$$

$$F = 1.0 + 16.0 \cdot (0.53 - E)^3, \quad (2)$$

$$x = \frac{2\pi(t - 50400)}{PER}, \quad (3)$$

where F is the obliquity factor which is expressed in terms of the satellite elevation angle E, while E is in units of semi-circles; t is the local time; AMP and PER are the amplitude and period of the cosine function and can be computed by the broadcast parameters. The detail formulae can be found in [1], and the relationships between zenith delay, AMP and PER are shown in Fig 1.

2.2 Single-Layer Model and Global Ionosphere Map

For obtaining a simple model, the ionosphere can be regarded as a thin shell and the total electron content along a signal path is compacted in terms of the vertical TEC and a multiplier, called obliquity factor, to account for the longer path length. Fig.2 shows the concepts of single layer model and the corresponding TEC values can be modeled as

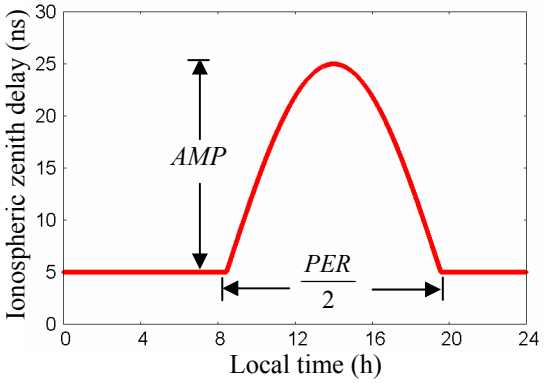


Fig. 1: The ionospheric zenith delay model

$$TEC(\delta)=\frac{1}{\cos\delta'}\cdot TECV\ , \tag{4}$$

where δ' can be determined by

$$\frac{\sin\delta}{(R_E+H)}=\frac{\sin\delta'}{R_E}\ , \tag{5}$$

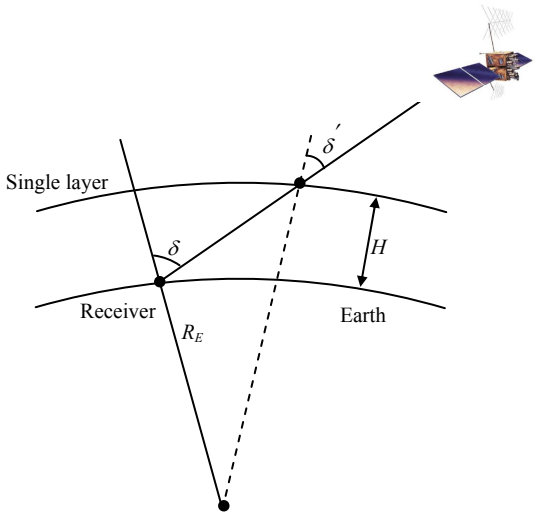


Fig. 2: Single layer model for ionosphere

The international GPS service (IGS) provides the TEC of the ionosphere on a global scale with the differential code biases of satellites and stations. Using this product, the ionospheric zenith delay can be computed by

$$I_Z=\frac{40.3\cdot TECV}{f^2}\ , \tag{6}$$

where f is the carrier frequency and $TECV$ is the vertical TEC which is given by IGS [3]. The vertical TEC at specified coordinates and universal time can be computed by

$$TECV(\beta,\lambda,t)=$$
$$\frac{T_{i+1}-t}{T_{i+1}-T_i}TECV_i(\beta,\lambda_i')+\frac{t-T_i}{T_{i+1}-T_i}TECV_{i+1}(\beta,\lambda_{i+1}')\ , \tag{7}$$

where β,λ,t is the geocentric latitude, longitude and universal time. $\lambda_i'=\lambda+(t-T_i)$. The $TECV_i$ can be computed by the simple 4-point interpolation formula,

$$TECV(\lambda_0+p\Delta\lambda,\beta_0+q\Delta\beta)=(1-q)(1-p)TECV_{0,0}+$$
$$p(1-q)TECV_{1,0}+q(1-p)TECV_{0,1}+pqTECV_{1,1}\ , \tag{8}$$

3. CORRECTIONS OF IONOSPHERIC MODEL

Based on the above equations, two different models can be used to compute the ionospheric effects. The GPS observables from two timing laboratories are chosen for comparison. One is National Institute of Information and Communications Technology (NICT) in Japan and the other one is Telecommunication Laboratories (TL) in Taiwan, and their coordinates are N 35.7068°, E 139.4885° and N 24.9536°, E 121.1645°. Fig. 3 and 4 show the differences between measured ionospheric delay and two different models of two laboratories. Based on these results, we can observe that the IGS model is better than broadcast model in mid-latitude region; however, in low-latitude region, the differences between measured ionospheric delay and IGS model result are still large.

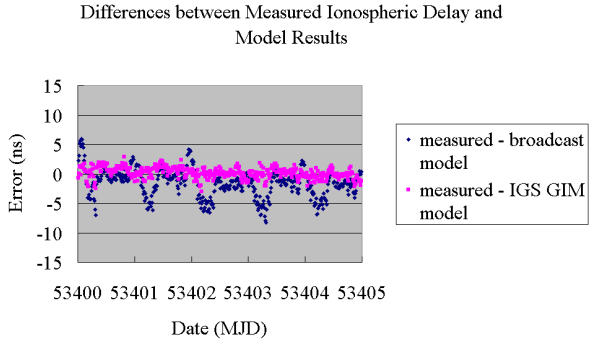


Fig. 3: Ionospheric delay test results of NICT

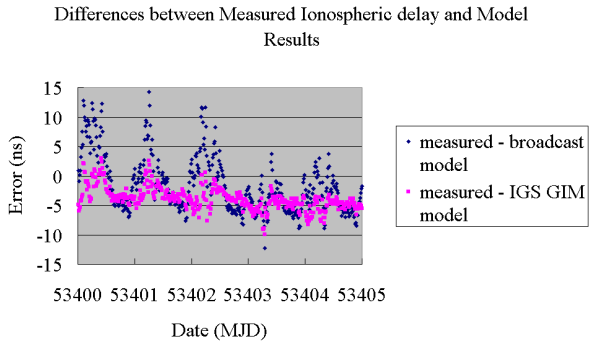


Fig. 4: Ionospheric delay test results of TL

For compensating the differences between the measured ionospheric delay and IGS GIM model with an adequate model, we use the empirical mode decomposition method [4] to extract the significant features of the differences. Fig. 5 shows the differences between measured ionospheric delay and IGS model results and

their corresponding decomposition components. After performing the Hilbert transform [5] on each component, the data can be expressed in the following form:

$$a_j(t)\exp(i\int\omega_j(t)dt, \quad (9)$$

where the sub-index j means the component number, $a_j(t)$ is the time-varying amplitude of the component j and $\omega_j(t)$ represents the instantaneous frequency of the jth component. Fig. 6 and 7 are the amplitude and instantaneous frequency of the components 1-3 in Fig. 5. After the different mode functions are obtained and the corresponding amplitude and instantaneous frequency functions are computed, our algorithm will check the criteria to select suitable components for further use. For the time comparison applications, two criteria are set for selection. The first one is the amplitude range, and the second one the range of instantaneous frequency. Since the performance of short-term time comparison is limited by the accuracy of the code measurements, we only concern about the mid-term and long-term performance. Thus, we can give an upper bound of the instantaneous frequency. The lower bound of amplitude also can be chosen for accuracy consideration. After selecting the components functions, we can give a table with the modified values at specified time, and the interior point in succession epochs can be determined by the interpolation method.

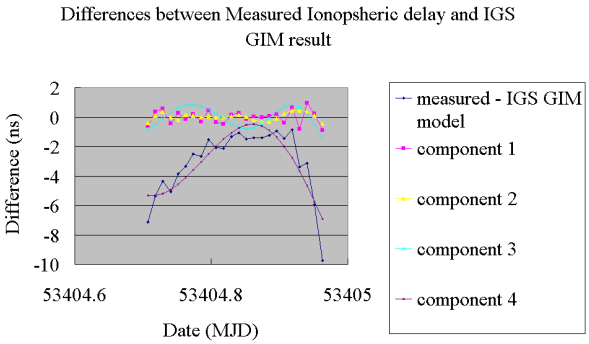


Fig. 5: Measured ionospheric delay minus the IGS model and their decomposition components of PRN1 data

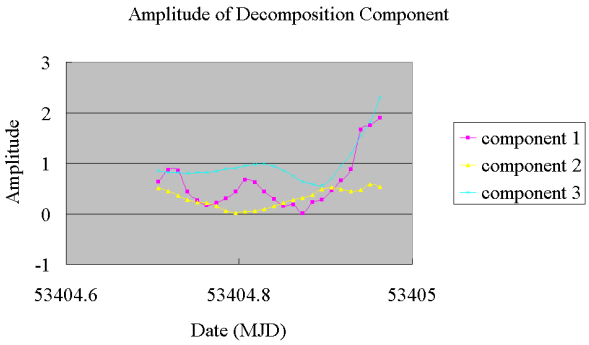


Fig. 6: Amplitude of the components of PRN1 data

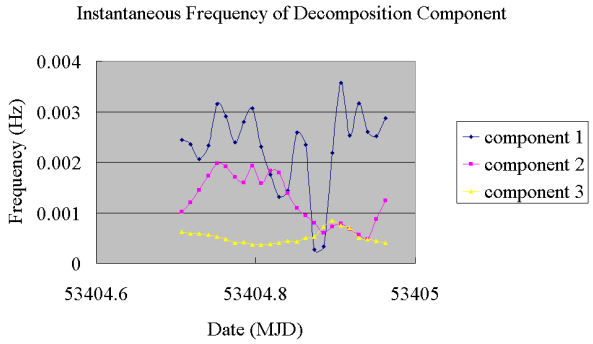


Fig. 7: Instantaneous frequency of the components of PRN1 data

4. TEST RESULTS

For verifying the performance of time comparison with different ionosphere models, the data from two timing stations are analyzed. Fig. 8 and 9 show the results from NICT station. The standard deviation of time comparison results by using measured ionospheric delay, broadcast model and IGS GIM product are 1.8 ns, 2.8 ns and 1.7 ns, respectively. Fig. 9 shows the differences between the time transfer results using measured values and IGS model values, and their differences is less than ± 2 ns in most epochs.

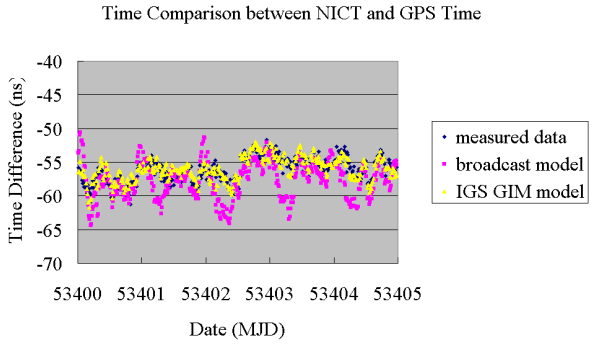


Fig. 8: Time comparison between NICT and GPS time

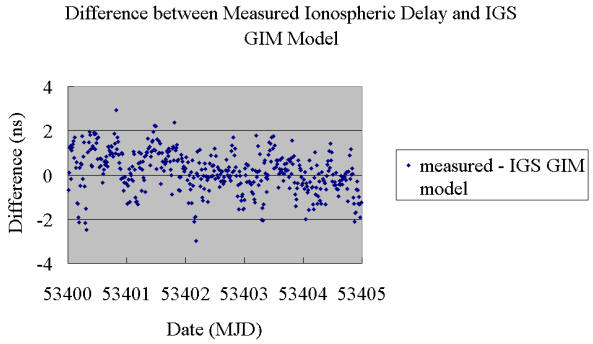


Fig. 9: Differences between time comparison results by using measured ionospheric delay and IGS model

On the other hand, the ionospheric effect is more serious

in low-latitude region. Fig. 10 shows the time comparison results between TL and GPS time from MJD 53400 to 53405. The standard deviation of time comparison by using measured ionospheric delay is 1.7 ns; however, they are 4.9 ns and 2.2 ns by using broadcast model and IGS model, respectively. Based on our algorithm, we can obtain the modified values for IGS model and re-compute time comparison results. Fig. 11 shows the final results based on our algorithm and the differences between measured data and modified model data are shown in Fig, 12. It is noted that the differences show in Fig. 12 is the average of unselected components due to our algorithm criteria.

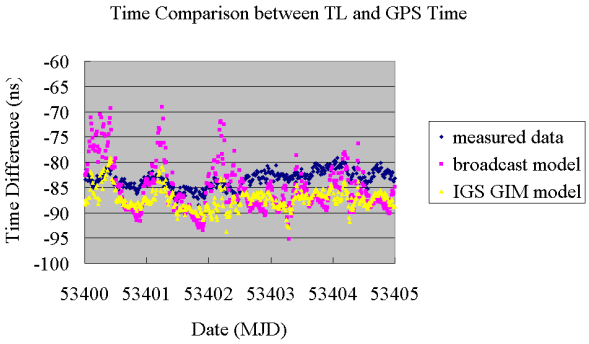


Fig. 10: Time comparison between TL and GPS time

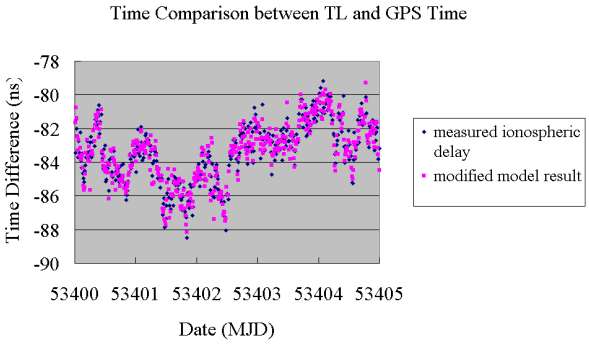


Fig. 11: Time comparison between TL and GPS time by using measured ionospheric delay and modified model

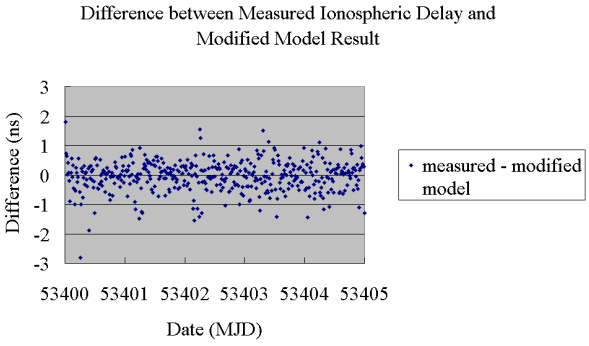


Fig. 12: Differences between time comparison results by using measured ionospheric delay and modified model

5. CONCLUSION

In this paper, the ionospheric effect on time comparison is studied. The broadcast model and IGS GIM products are used to assess the ionospheric delay and the linear combination of dual frequency data are used for comparison. Based on the test data, it is obvious that the performance of time comparison in low-latitude region is degraded by using the ionospheric model to estimate the delay. For reducing this error, a dual frequency reference station is necessary. After decomposing the difference between measured ionospheric delay and IGS model results, we can select suitable mode for model correction. The final results show that the algorithm can be applied to a local region for time comparison applications.

REFERENCES

[1] SPS, Global Positioning System, Standard Positioning Service Signal Specification, 2nd Edition, U.S. Department of Defense, 1995.

[2] P. Misra and P. Enge, Global Positioning System: Signals Measurements, and Performance, Ganga-Jamuna Press, 2001.

[3] S. Schaer, W. Gurtner and J. Feltens, IONEX: The IONosphere Map Exchange Format Version 1, Astronomical Institute, University of Berne, Switzerland, 1998.

[4] N. E. Huang, et al., The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis, @ Proceedings: Mathematical, Physical and Engineering Sciences, vol. 454, no. 1971, pp. 903-995, March 1998.

[5] J. S. Bendat and A. G. Piersol, Random Data: Analysis and Measurement Procedures, 2nd Edition, New York: Wiley, 1986.

TIMING AND OPTICAL FIDUCIAL SYSTEMS FOR THE *LASER MEGAJOULE*

Pierre Raybaut (Tel. : +33 1 69 26 54 08, Fax. : +33 1 69 26 70 43, e-mail : pierre.raybaut@cea.fr),
Michel Prat (Tel. : +33 1 69 26 61 15, Fax. : +33 1 69 26 70 43, e-mail : michel.prat@cea.fr),
Virginie Allouche, Pascal Leclerc, Jean-François Pastor, Jean-Yves Salmon
CEA/DIF, BP 12, 91680 Bruyères-le-Châtel, FRANCE

1. ABSTRACT

Design of timing and optical fiducial systems for the forthcoming *Laser MégaJoule* (LMJ) facility is outlined. Recent results obtained with the LMJ prototype (LIL–*Ligne d'Intégration Laser*) timing system are reviewed together with future improvements planned for the LMJ facility. The future LMJ optical fiducial system is then presented.

2. INTRODUCTION

The LMJ facility, a 240-beam laser, is currently under construction in France. It will be mainly used for Inertial Confinement Fusion (ICF) experiments, i.e. to study high energy and density plasma produced by imploding deuterium-tritium targets with high energy laser beams. In this paper, we focused on two distinct LMJ subsystems: the timing system, and the optical fiducial system. The first is intended to synchronize laser beams, diagnostics, and other devices with trigger signals, while the second provides an optical time reference on imaging diagnostics (like streak cameras). The common goal of these two subsystems is to guarantee a precise timing of ICF experiments.

3. LMJ TIMING SYSTEM

In addition to pulse energy or spatial focusing resolution requirements, a very precise beam-to-beam power balance and timing are required to guarantee a highly homogeneous target compression. Furthermore, being built on a large ground area of 30,000 m², LMJ requires a specific timing system to generate many trigger pulses (around 5000 signals) to control the laser and diagnostics with a high time precision (from some microseconds to some picoseconds, for a dynamic range of up to 2 s) throughout the facility.

3.1 Principle of timing system

The timing system we have studied is based on a master clock from which the time is transferred to distant areas in the facility by sending an optical bit sequence through single-mode fibers. In each of these areas, clock is regenerated from this sequence and coupled with highly accurate slave delay generators. Since mid-2003, this concept has been carried out for the LIL facility (which is presently a 4-beam LMJ prototype), with 800 trigger signals [3-1]. However we would like to improve some parts of this timing system for the LMJ facility. Three important improvements are planned. Firstly a new optical receiver is tested to reduce the influence of optical power on generated delays and jitter. Then, a new architecture is designed to increase the number of local areas where the trigger signals are distributed: 16 local areas for the LIL facility and 256 for the LMJ facility. Finally we are planning to develop a specific timing system with a significantly low jitter for pulse shaping system (generation of the 240 synchronized Laser pulses).

3.2 New optical receiver

A specific optical receiver is included in slave delay generators. This device converts the optical bit sequence to electrical local clock. However, the delay generated by a slave unit changes with the level of received optical power. Furthermore, the jitter of generated signals increases markedly, for a power lower than -30 dBm. Consequently we designed a new optical receiver to reduce this effect. We are testing the new device and the first results are promissing. Indeed, the delay and the jitter are stable if the power is lower than -20 dBm : the peak to peak variation for the delay is less than 200 ps and the jitter is constant (see figure 3-1). We will continue to improve the optical receiver to reduce the effect of the power for a level higher than -20 dBm. Moreover, we will correct the remaining delay variation with a numerical model.

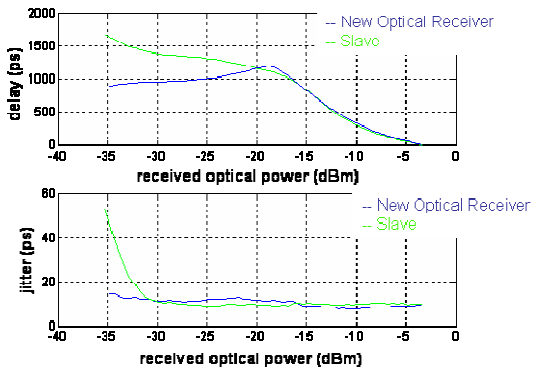


Fig. 3-1. Delay and jitter vs optical power for a slave unit and the new receiver.

3.3 New architecture and local area design

We are currently designing future architecture improvements for the LMJ timing system. We will use 3 levels of optical splitters to distribute the optical message from the master clock to the 256 different areas of LMJ building (see figure 3-2).

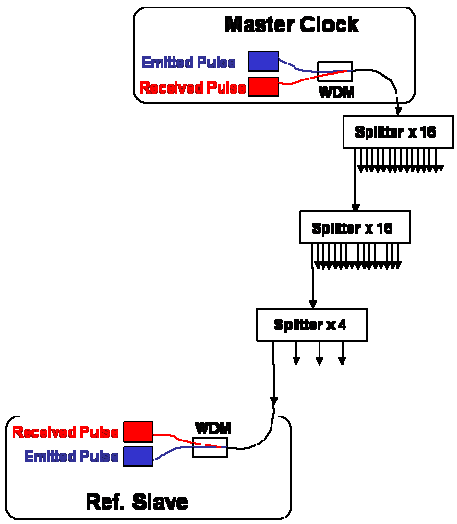


Fig. 3-2. Architecture of LMJ Timing System

There will be four Slave generators, in each local area : three standard slave units and one reference slave unit. A reference unit will be able to measure time differences between the master clock and one LMJ area, exactly as in the LIL timing system. Moreover, a chronometer will be included in each local area to measure time difference between the different slave generators of the area. Furthermore, we are planning to integrate a specific chronometer inside each slave unit to measure temperature and long-term drifts (see figure 3-3). This device will be able to send alarm to the timing system monitoring if the drifts increase too high. In this case, the slave generator will have to be calibrated again.

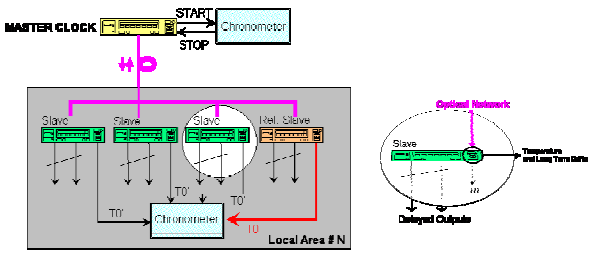


Fig. 3-3. Local Area Design

3.4 Specific timing system for pulse shaping

It is necessary to use a specific timing system for pulse shaping function. This function allows to generate the 240 laser pulses with a specific shape. The triggers of pulse shaping devices will have to be very precise (jitter lower than 10 ps rms) to synchronize the 240 pulses. Consequently, we will design a high accuracy timing system with low jitter. This timing system will be based on passive electro-mechanical devices without jitter : a high voltage pulse generator will distribute 60 trigger signals for 60 pulse shapers (one shaper for 8 laser beam). The signals will be transmitted with coaxial cables. The delay of each signal will be adjusted with a specific delay line. And a specific circuit will be designed to reduce signal risetime, after cable transmission (see figure 3-4). Some parts of this specific system will be tested this year (cable transmission and delay line) to estimate the precision and the drifts.

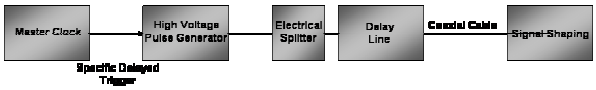


Fig. 3-4. Principle of Specific Timing System

The LMJ timing system will be more advanced than the LIL system, allowing us to improve the way of using it and to reduce the cost per channel, while keeping the same accuracy or improving it.

4. LMJ OPTICAL FIDUCIAL SYSTEM

4.1 Introduction

Highly stable and extremely low-jitter laser-based system is required for designing the LMJ optical fiducial system. The main issue in conceiving such system is to guarantee peerless timing specifications in harsh environmental conditions [4-1].

As mentioned above, a very precise timing is required for LMJ. Thus, the LMJ 240 laser pulses have to be synchronized with a precision of 50 ps (single-shot RMS jitter), and the timing system has to trigger the laser driver and diagnostics with a precision of less than 15 ps. The so-called diagnostics use detectors such as streak cameras to study the laser-produced plasma. For simplicity, we will only consider the case of streak cameras fiducial in the remaining (we will not consider the case of framing cameras which also require fiducial).

Triggering diagnostics is generally not sufficient in laser-plasma interaction experiments [4-2]. Interpretation of experimental results actually requires that the temporal evolution of the observed x-ray emission be correlated exactly with respect to the incident laser pulse. One can add that timing system is neither precise enough nor well adapted to this purpose. On LMJ, this correlation will be achieved in streak cameras by using a laser pulse fiducial monitor system triggered by the timing system. This laser system produces a signal (namely the optical fiducial signal) which is sent through multimode optical fibers to streak camera photocathodes. This optical fiducial signal is a set of eight laser pulses with known relative time delays, which are used to calibrate the sweep speed of streak cameras (Fig. 4-1).

4.2 Specifications, architecture and technical issues

Specifications of LMJ optical fiducial signal are given in Table 4-I. This optical signal consists of eight 50-ps to 100-ps pulses, with relative time delays of 500 ps or 2 ns (depending on streak camera time range), and with a shot-to-shot peak power stability less than 20%. The first pulse of this sequence is the fiducial time reference, which can be related with a good precision to the LMJ laser pulse reference time (thanks to a specific low power experiment, or using a specific diagnostic at each gain shot). The remained pulses are used for the sweep speed calibration (Fig. 4-1 and Fig. 4-2).

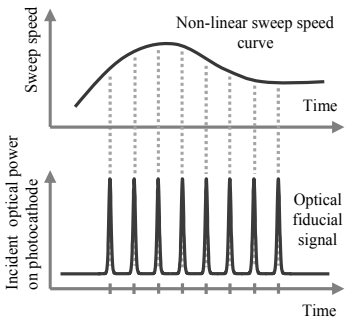


Fig. 4-1. On a streak camera, optical fiducial signal not only provides a reference time (given by the first pulse position), but also allows calibrating the streak camera sweep speed.

Detector	Wavelength	Pulse energy	Channel number
Visible streak camera	~526nm	10 pJ	24
X-ray streak camera	~351nm	200 nJ	10-12

Table 4-I. LMJ optical fiducial specifications for streak cameras.

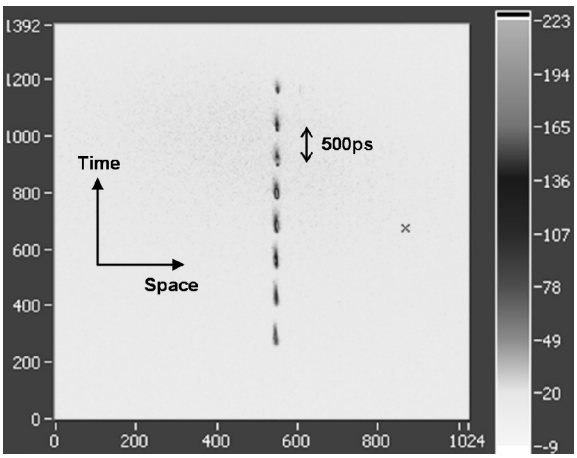


Fig. 4-2. 2-GHz optical fiducial signal at 532 nm. This picture was acquired on a visible streak camera (S20 photocathode).

An overview of what could be the future LMJ optical fiducial system is shown on Fig. 4-3. Picosecond optical pulse generation can be achieved with various laser systems, but, to the best of our knowledge, the most stable ones are based on Semi-Conductor Saturable Absorber Mirror (SESAM) technology [4-3]. Thus, the "amplified laser source" on Fig. 4-3 is a Nd:YLF laser passively mode-locked by SESAM and frequency-synchronized to the LMJ timing system at around 77 MHz (Fig. 4-4), followed by two amplification stages (single-shot Nd:YLF regenerative and multipass amplifiers). An ultra-fast Michelson-based pulse multiplexer [4-4] generates 2 GHz and 0.5 GHz fiducial signals, and then a third harmonic generation stage allows us to produce both 526 nm and 351 nm fiducial signals, which are divided in N channels (where $N = 12$ for 526 nm, and $N = 24$ for 351 nm), and sent through 60-m optical fibers to diagnostics.

In LMJ optical fiducial system, there are three main subsystems that could introduce significant single-shot jitter (i.e. jitter *and* long-term drift) :

- LMJ timing system : 15-ps RMS jitter ;
- 60-m optical fibers : ~2-ps RMS long-term drift (in ±0.5°C environment) ;
- and laser system, which is the subject of this paper.

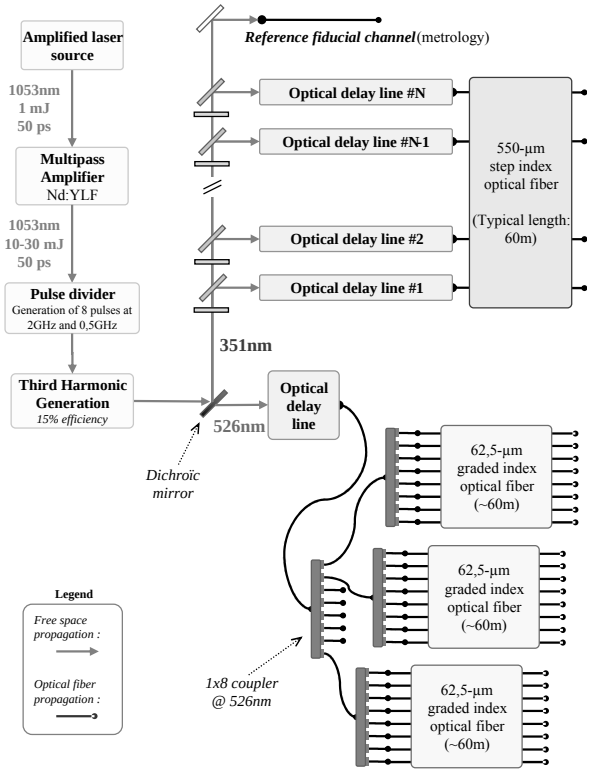


Fig. 4-3. Simplified architecture of the future LMJ optical fiducial system (only the streak camera fiducial channels are shown).

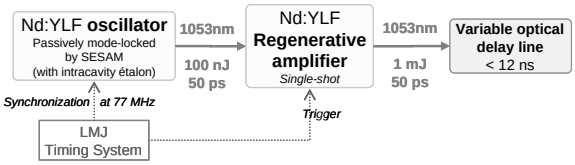


Fig. 4-4. LMJ optical fiducial system laser source ("Amplified laser source" in Fig. 4-2).

Our laboratory has been addressing the issues related to LMJ optical fiducial system design for three years. This work includes for instance timing jitter and long-term stability issues in laser system, or pulse broadening and damage (or non-linear) threshold issues in optical fibers. All these issues have to be solved in LMJ harsh environmental conditions [4-1], and fiducial specifications (jitter and pulse energy) have to be guaranteed in single-shot operation.
In this paper, we shall limit ourselves to experimental results relative to laser system stability, which is one of the main technical issues for designing the LMJ optical fiducial system.

We report an original design, that we used as a prototype of LMJ optical fiducial system, based on a Nd:YVO_4 laser passively mode-locked by SESAM and synchronized to a master-slave timing system (similar to LMJ timing system). This laser oscillator produced 7-ps pulses at three wavelengths (1064 nm, 532 nm, and 355 nm), at a repetition rate around 77 MHz.

4.3 Optical fiducial prototype long term drift measurements

As shown in Fig. 4-5, using a 12-ps rise time photodiode ("Allegro-20-FS", Advanced Laser Diode Systems) and the 7-ps rise time sampling module of a signal analyser ("CSA-8000", Tektronix, with the following settings : 200-fs resolution, 200-ps time range), we measured the long term drift in a temperature-stabilized environment (better than ±0,5°C).

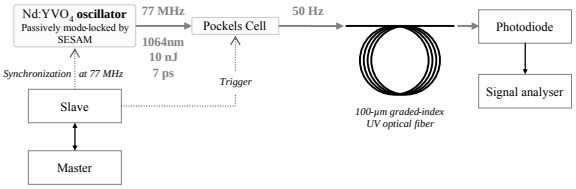


Fig. 4-5. Experimental set-up for laser long-term drift measurement

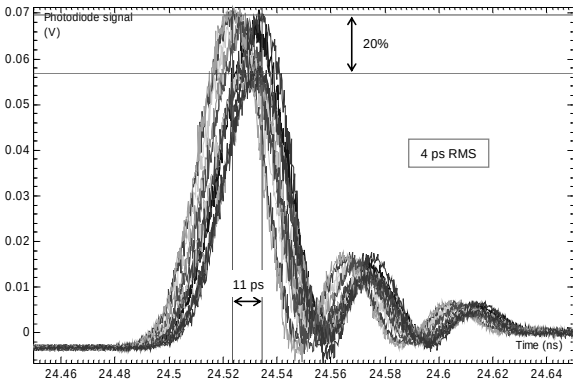


Fig. 4-6. Long-term stability of our SESAM-based laser system over 3 days is around 4 ps RMS (see text).

Results are shown in Fig. 4-6. In the graph, each curve (raw data) represents a 45-minutes acquisition of the output pulse (averaging allowed us to separate long-term drift from short-term jitter), which is broadened up to 20 ps FWHM because of the limited detection bandwidth. These 14 curves were acquired over 3 days, at different moments of the day – given that all instruments were turned off at the end of the day, and then turned on the next day morning. The maximum pulse position deviation was 11 ps, and the standard deviation was 4 ps, which is quite low. In fact, this establishes that jitter due to the laser was negligible, because these values corresponds to the best performances of our master-slave timing system (in

temperature-controlled environment). With regard to signal amplitude, peak-to-peak variation was about 20 %, which is also a good result for a single-shot operation laser system.

4.4 Optical fiducial prototype jitter measurements

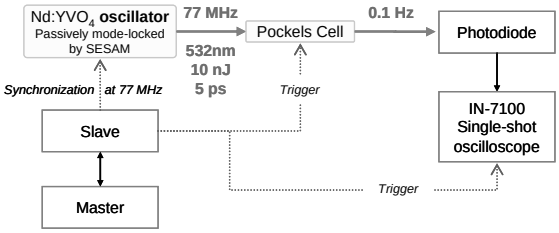


Fig. 4-7. Experimental set-up for laser jitter measurement.

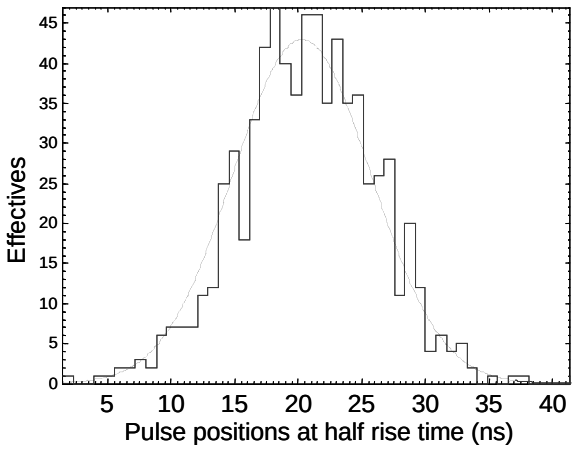


Fig. 4-8. Histogram of pulse positions at half rise time. According to the Gaussian fit, pulse position jitter is around 5 ps RMS.

We measured timing jitter of our system using a 7-GHz single-shot oscilloscope (1-ns time range, triggered at 0.1 Hz), and the same photodiode as in previous section (Fig. 4-7). Thus, 720 experimental curves were acquired over a 120-min period. Pulse positions histogram is shown in Fig. 4-8 : jitter was 5 ps RMS, which corresponds again to the best performances of our timing system.

4.5 Conclusion

With regard to jitter or long-term drift of the optical fiducial laser system prototype, we were not able to measure any contribution from the laser (laser timing jitter is therefore less than 1 ps RMS).
The whole system – including optical fiber for long-term drift measurements (10-m UV graded-index optical fiber) – was stable for hours and could allow us to build an optical fiducial system for the LMJ, with a single-shot RMS jitter less than 2 ps (excluding timing system).

REFERENCES

[3-1] P. Leclerc *et al.*, Performance of the picosecond timing system for the LIL Laser Facility, 16th EFTF, March 2002.

[4-1] J.L. Bourgade *et al.*, New constraints for plasma diagnostics development due to the harsh environment of MJ class lasers (invited), Rev. Sci. Instrum., vol. 75 (10), pp. 4204-4212, October 2004.

[4-2] S.A. Letzring *et al.*, An absolutely timed x-ray streak camera for laser fusion experiments, J. Appl. Phys., vol. 54 (11), 1983.

[4-3] U. Keller *et al.*, Semiconductor Saturable Absorber Mirrors (SESAM's) for femtosecond to nanosecond pulse generation in solid-state lasers, IEEE J. Sel. Top. Quantum Electron., vol. 2 (3), pp. 435-453, September 1996.

[4-4] C.W. Siders *et al.*, Efficient high-energy pulse-train generation using a 2n-pulse Michelson interferometer, App. Opt., vol. 37 (22), pp. 5302-5305, August 1998.

DYNAMIC PARAMETERS OF SYNCHRONIZATION SIGNALS FOR NON-OVERLAPPING AND OVERLAPPING DATA SEGMENTS

Andrzej Dobrogowski, Michal Kasznia

Institute of Electronics and Telecommunications
Poznan University of Technology
ul. Piotrowo 3A, 60-965 Poznan, Poland
phone: +48 61 6652 293, fax: +48 61 6652 572
dobrog@et.put.poznan.pl, mkasznia@et.put.poznan.pl

ABSTRACT

This paper deals with the methods of characterization of timing signals using dynamic parameters. First the idea of dynamic Allan deviation and dynamic time deviation is described. Then different ways of the parameters' calculation, taking into account non-overlapping and overlapping segments of time error data are presented. The results of computation experiment performed for different data series are presented and discussed. The proper solutions of computing procedures are suggested.

1. INTRODUCTION

Allan deviation ADEV and time deviation TDEV are the parameters, which describe the quality of synchronization signal in the telecommunication network [1, 2, 3]. The parameters allow to assess the variations of time interval provided by the synchronization signal and recognize the type of phase noise affecting the signal. The estimates of the parameters are calculated for a series of observation intervals using the sequence of time error samples measured between two timing signals at the network interface.

The characterization of the timing signal using the dynamic parameter was proposed by L. Galleani and P. Tavella at the Joint Meeting of IEEE FCS and EFTF in 2003 [4]. This characterization with dynamic Allan deviation (DADEV) allows to recognize the variations of the phase noise affecting the analyzed signal. This approach was extended by the authors of this paper at the 18th EFTF in 2004 [5]. The dynamic time deviation (DTDEV) was introduced and the time effective methods of calculation of the dynamic parameters were proposed.

The evaluation of timing signal using dynamic parameters consists in the calculation of the Allan deviation or time deviation values for the segments (slices) of time error sequence, with a change of the initial point of calculation and presentation of the results of calculation in the form of three-dimensional graph as a function of observation interval τ and time t. We can distinguish between calculation for overlapping and non-overlapping segments of data. The arrangement of the segments depends on the number of segments, the time shift between the segments, their length and the length of the whole data sequence. This arrangement (resulting in specific calculation procedure) influences the form of presented calculation result (i.e. we can obtain different forms of the 3D graphs for the same data sequence).

In the paper the influence of the computation procedure (determined by the data segments arrangement) on the form of results obtained is analyzed. The results of DADEV and DTDEV calculation performed for several different time error sequences using different calculation procedures (for overlapping and non-overlapping data segments), are presented and discussed. The proper computation procedures are suggested.

2. IDEA OF DYNAMIC PARAMETERS

The idea of dynamic parameters is quite simple. Instead of one curve representing the values of ADEV or TDEV as a function of observation interval τ, the set of the curves plotted in the form of three-dimensional graph as a function of observation interval τ and time t is considered. As a result we can recognize the changes of the type of phase noise affecting analyzed timing signal. The changes of the slope of the parameter's curve indicate the changes of the noise type [4].

In order to obtain this form of graph, the following computation procedure must be performed. First, the data sequence (time series of time error measured at some network interface) is divided into equal segments (slices) with the length of T_s. We can distinguish between overlapping data segments (Fig. 1) and non-overlapping data segments (Fig. 2). Then the calculation of the parameter's value for a required range of observation intervals for each data segment is done. The results of calculation are plotted in a form of three dimensional graph.

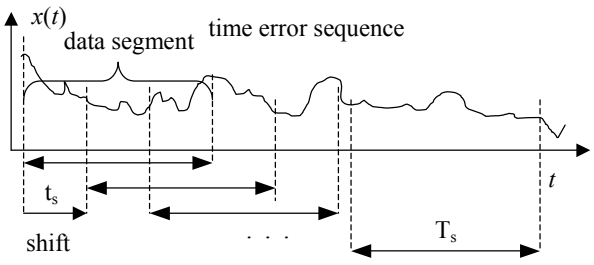


Fig. 1. Overlapping segments of time error sequence

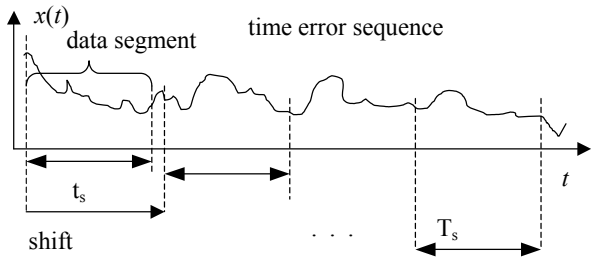


Fig. 2. Non-overlapping segments of time error sequence

3. COMPUTATION PROCEDURES

The formulae for the estimators of the Allan deviation ADEV and the time deviation TDEV take the form:

$$\hat{A}DEV(\tau)=\sqrt{\frac{1}{2n^2\tau_0^2(N-2n)}\sum_{i=1}^{N-2n}(x_{i+2n}-2x_{i+n}+x_i)^2} \quad (1)$$

$$\hat{T}DEV(\tau)=\sqrt{\frac{1}{6n^2(N-3n+1)}\sum_{j=1}^{N-3n+1}\left[\sum_{i=j}^{j+n-1}(x_{i+2n}-2x_{i+n}+x_i)\right]^2} \quad (2)$$

where $\{x_i\}$ is a sequence of N samples of time error function $x(t)$ taken with interval τ_0; $\tau=n\tau_0$ is an observation interval.

The procedure of dynamic ADEV and dynamic TDEV calculation depends on the set of quantities. These are: length of the whole data series T, length of the data segments created within the whole data series T_s, time shift between the initial points of the data segments t_s, and range of observation intervals (τ_{min} - τ_{max}).

The relation between the whole data series length, segments' lengths and the time shift t_s determines the number of segments and the arrangement of the segments (overlapping or non-overlapping). We must also take into consideration the relation between T_s and the maximum observation interval τ_{max}. In practice, the value of the parameter's estimate can be calculated, when the length of data series (in this situation: segment's length) is two times longer for ADEV and three times longer for TDEV. According to the telecommunication standards and recommendations, the length of data sequence used for ADEV or TDEV calculation must be 12 to 15 times longer than the maximum observation interval τ_{max} [1, 2, 3]. The minimum observation interval τ_{min} is determined by the sampling interval τ_0: τ_{min} must be three times longer than τ_0.

In order to analyze the influence of the procedure on the form of calculation results, four different computation procedures were established. The quantities determining these procedures are presented in the Table 1. Two procedures (A and B) are with non-overlapping data segments and two next (C and D) are with overlapping data segments. The data sequences used for computation contain the time error samples taken with sampling interval τ_0=1/30 s during the time T of 20000 s.

The data sequence was divided into the five non-overlapping segments with the length of 4000 s for the procedure A. The range of observation intervals is from 0.1 s till 1000 s, so the relation between T_s and τ_{max} fulfils only the practical requirement. For the procedure B we have more non-overlapping segments (20) than for A, with shorter length (1000 s). Therefore the range of observation intervals is smaller than in the procedure A, but the relation between T_s and τ_{max} is better. As result we can obtain more curves on the 3D-graph for the shorter range of τ but with better quality of the estimate.

The number of the overlapping data segments for the procedure C and D is the same (11), but they differ in the segment's length and shift. The segment's length for the procedure C is 4000 s (as for the procedure A) with the segment's shift of 1500 s. For the procedure D the length of the data segment is 10000 s (half of the length of the whole data series) with the shift of 1000 s. Therefore the first segment for this procedure covers the first half of the whole data series, the last one covers the second half.

Table 1. Details of computation procedures

procedure	over-lapping segments	observation intervals τ_{min}- τ_{max}	segment's length T_s [s]	segment's shift t_s [s]	number of segments
A	no	0.1-1000	4000	4000	5
B	no	0.1-100	1000	1000	20
C	yes	0.1-1000	4000	1500	11
D	yes	0.1-1000	10000	1000	11

4. EXPERIMENT

In the computing experiment the procedures described in the previous section were used to calculate both parameters, dynamic ADEV and dynamic TDEV. The ordinary Allan deviation and time deviation were also computed for better comparison of the obtained results. The calculations were performed for several different data sequences. First sequence, denoted as WPM represents the samples of white phase noise, which changes its standard deviation for the period of 2000 s after the half of the sequence. The Allan deviation for this sequence is presented in Fig. 3. The plots of dynamic ADEV computed using procedures A, B, C, and D are presented in Fig. 4–7. The time deviation for the WPM sequence is presented in Fig. 8. The plots of dynamic TDEV for this sequence are presented in Fig. 9–12. The second sequence (WPFM1) represents white phase noise which changes its type to the white frequency noise for the period of 2000 s after the half of the sequence. The results of ADEV and DADEV calculation are presented in Fig. 13–17. The results of TDEV and DTDEV calculation are presented in Fig. 18–22. The third sequence (WPFM2) represents the change of dominating noise. In the middle of the sequence the domination of WPM (WFM has lower level) changes into the domination of WFM (WPM takes lower level). The results of ADEV and DADEV calculation are presented in Fig. 23–27. The results of TDEV and DTDEV calculation are presented in Fig. 28–32.

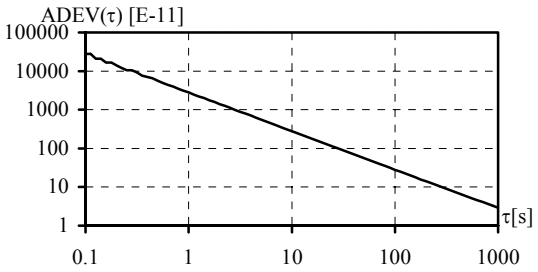


Fig. 3. Allan deviation for WPM sequence

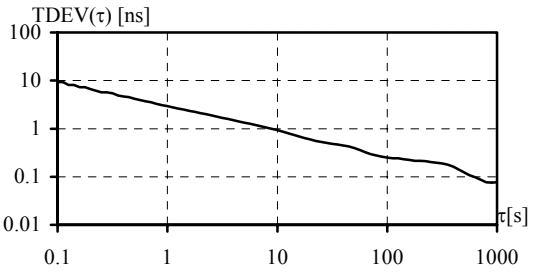


Fig. 8. Time deviation for WPM sequence

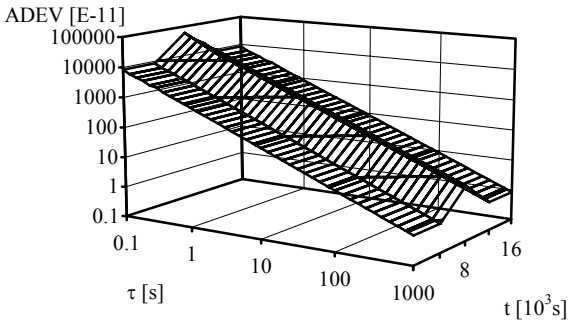


Fig. 4. DADEV for WPM sequence – procedure A

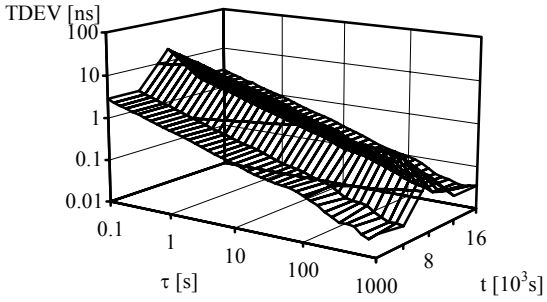


Fig. 9. DTDEV for WPM sequence – procedure A

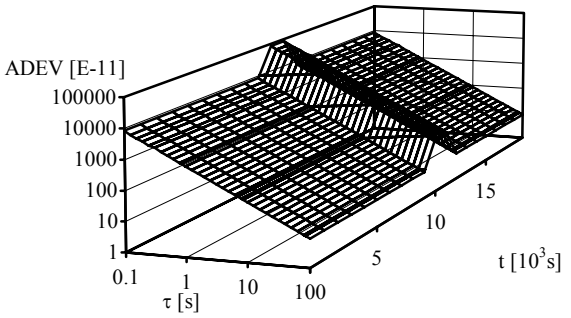


Fig. 5. DADEV for WPM sequence – procedure B

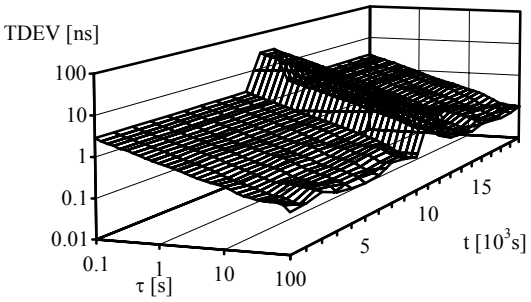


Fig. 10. DTDEV for WPM sequence – procedure B

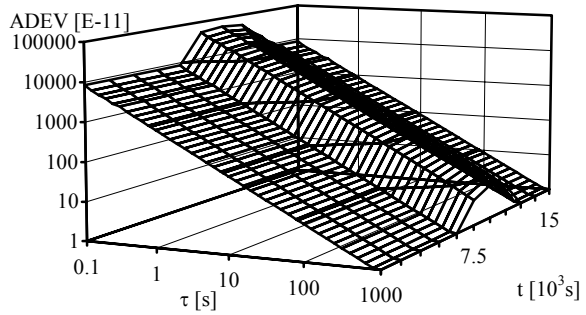


Fig. 6. DADEV for WPM sequence – procedure C

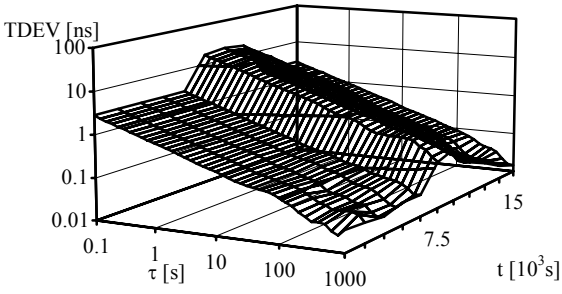


Fig. 11. DTDEV for WPM sequence – procedure C

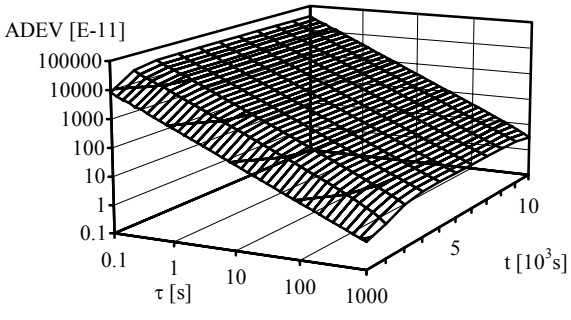


Fig. 7. DADEV for WPM sequence – procedure D

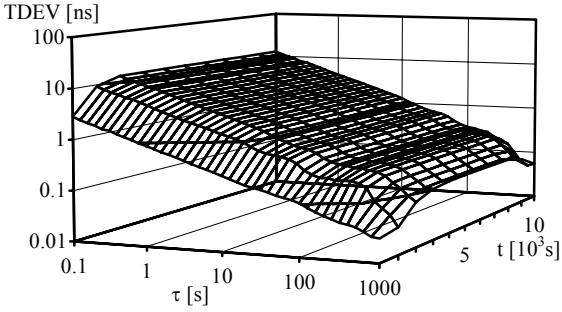


Fig. 12. DTDEV for WPM sequence – procedure D

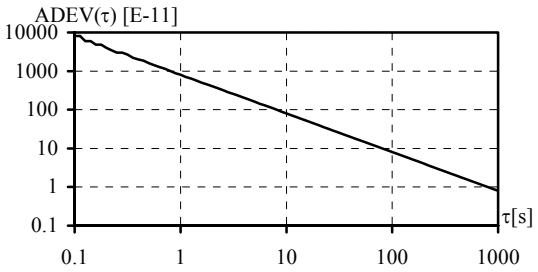


Fig. 13. Allan deviation for WPFM1 sequence

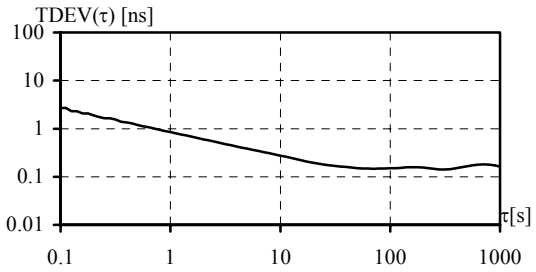


Fig. 18. Time deviation for WPFM1 sequence

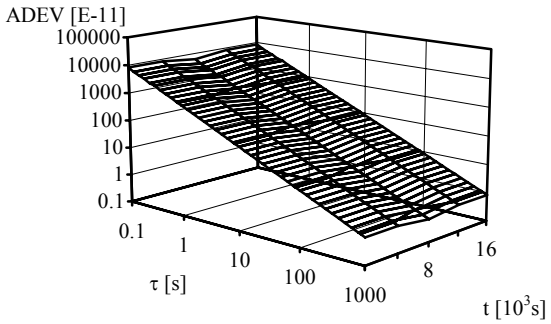


Fig. 14. DADEV for WPFM1 sequence – procedure A

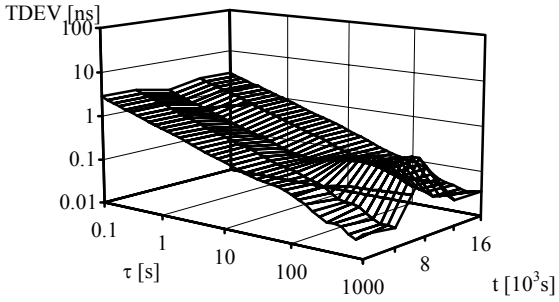


Fig. 19. DTDEV for WPFM1 sequence – procedure A

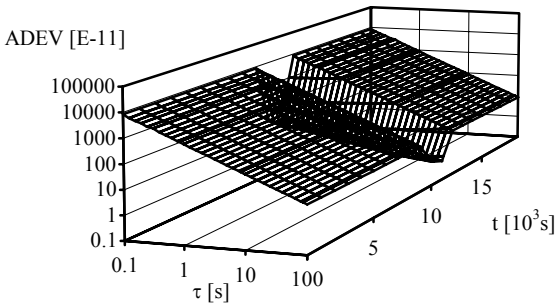


Fig. 15. DADEV for WPFM1 sequence – procedure B

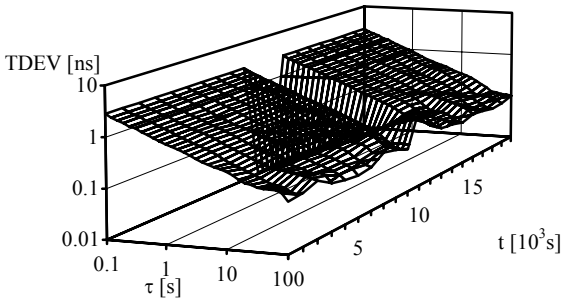


Fig. 20. DTDEV for WPFM1 sequence – procedure B

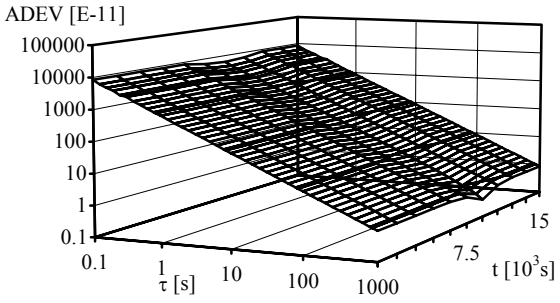


Fig. 16. DADEV for WPFM1 sequence – procedure C

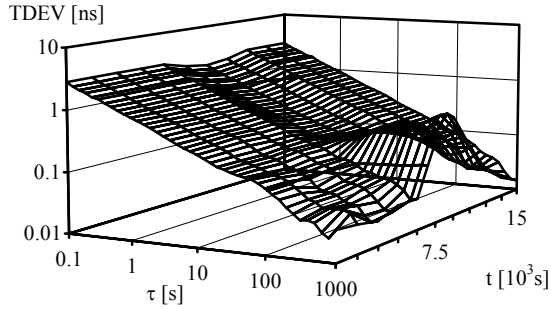


Fig. 21. DTDEV for WPFM1 sequence – procedure C

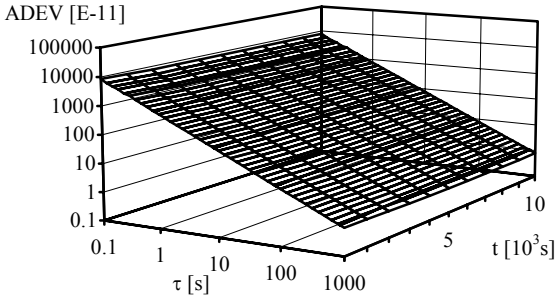


Fig. 17. DADEV for WPFM1 sequence – procedure D

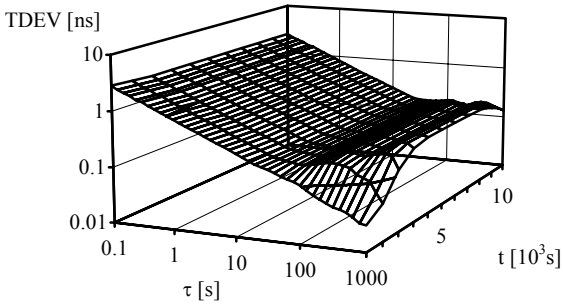


Fig. 22. DTDEV for WPFM1 sequence – procedure D

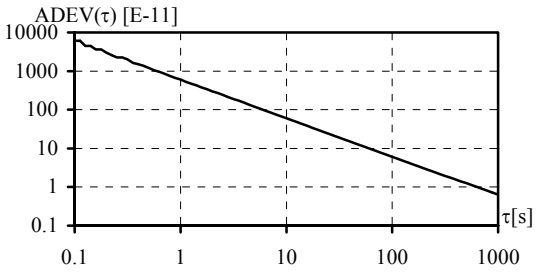


Fig. 23. Allan deviation for WPFM2 sequence

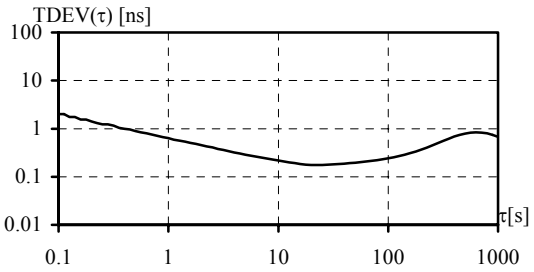


Fig. 28. Time deviation for WPFM2 sequence

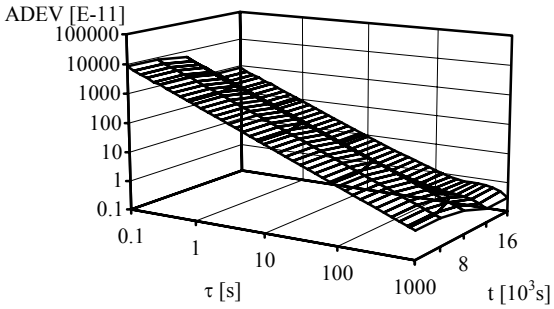


Fig. 24. DADEV for WPFM2 sequence – procedure A

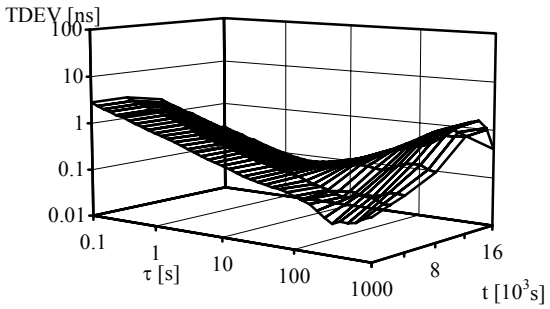


Fig. 29. DTDEV for WPFM2 sequence – procedure A

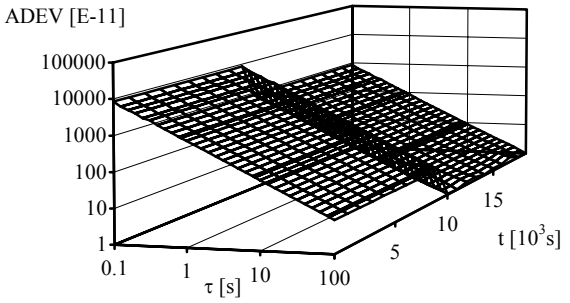


Fig. 25. DADEV for WPFM2 sequence – procedure B

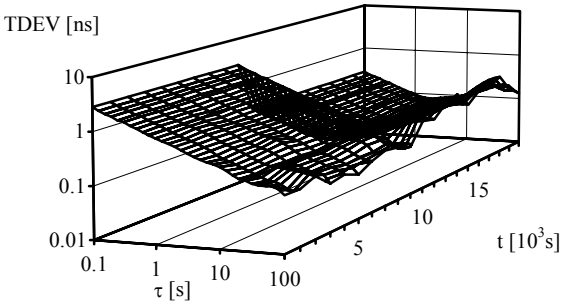


Fig. 30. DTDEV for WPFM2 sequence – procedure B

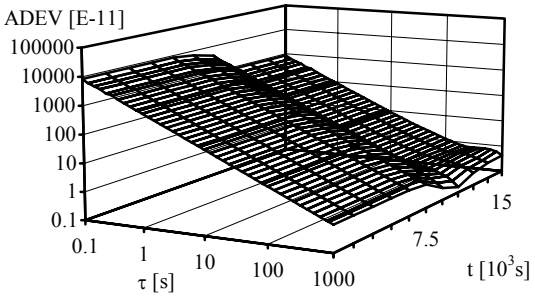


Fig. 26. DADEV for WPFM2 sequence – procedure C

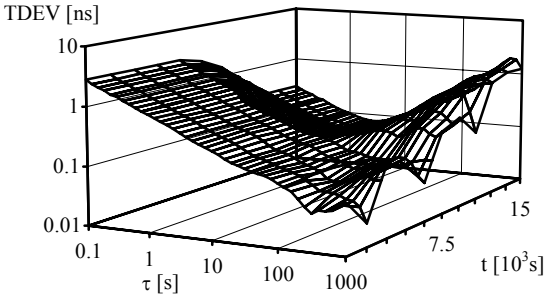


Fig. 31. DTDEV for WPFM2 sequence – procedure C

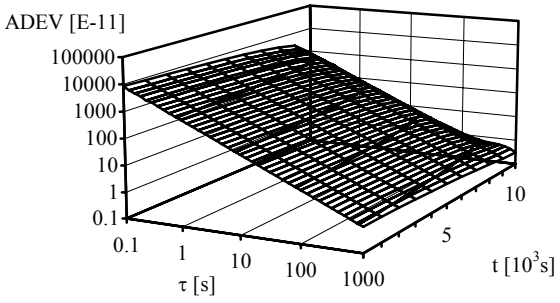


Fig. 27. DADEV for WPFM2 sequence – procedure D

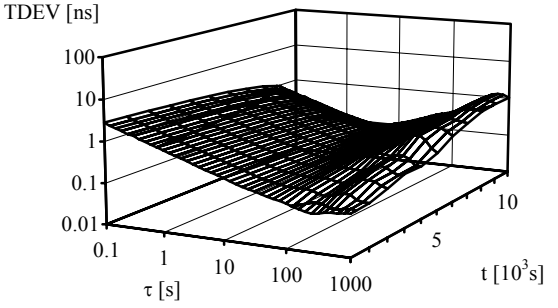


Fig. 32. DTDEV for WPFM2 sequence – procedure D

5. DISCUSSION OF THE EXPERIMENT RESULTS

The results obtained by means of the proposed procedures will be described as follows.
The slopes of common Allan deviation (Fig. 4) and time deviation (Fig. 9) computed for the WPM sequence are typical for the white phase noise and do not indicate the change of the noise level in the middle of the sequence. This change was recognized using DADEV and DTDEV, computed for the sequence. The best illustration of the changes was obtained using the procedure B (non-overlapping short segments). The number of curves obtained using procedure A (Fig. 5 and 10) is too small to indicate the instant of the change. The length of the segments in the procedure D is too long. The shapes in Fig. 8 and 13 illustrate some change, but it is not possible to fix its initial and final instant. All segments, excluding the first one, cover the period of magnified noise, which affects the computation results. The variation of the slope of DTDEV for observation interval τ=1000 s in Fig. 12 (procedure C) is observed because the requirement for the relation between τ_{max} and T_s is not fulfilled.
The change of the common TDEV curve for WPFM1 sequence (Fig. 20) indicates the presence of white frequency noise. The form of this presence can be concluded from the plots of dynamic parameters. However, the forms of dynamic TDEV (especially for procedure B, Fig. 22) recognize the presence of WFM noise better. In the case of DADEV only the result obtained by means of the procedure B (Fig. 17) recognizes the presence of the period of changed noise.
The change of the dominating noise representing by the WPFM2 sequence is well recognized by the results of DTDEV calculation (Fig. 32 – 35). The curve of common TDEV (Fig. 31) indicates the presence of both types of noise only. The slope of the DADEV and DTDEV obtained using procedure D (Fig. 30 and 35) suggests rather soft change of the dominating noise. The rapid change of the dominating noise type is better illustrated by the shapes obtained using the procedures B and C (Fig. 28, 29, 33, and 34).
The results of the experiment allow us to characterize the proposed procedures of calculation of dynamic ADEV and dynamic TDEV. The number of curves provided by the procedure A seems to be too small to show its dynamic feature. The number of curves in the case of procedure B expresses the dynamic character of the parameters better. Short length of segments allows to identify the changes in the analyzed signal, but limits the range of observation intervals τ – the maximum observation interval τ_{max} must be shorter.
In the case of procedures, which use the overlapping data segments for computation, better results were obtained by means of procedure C. The number of curves is enough to express the dynamic character of the parameters. The segment's length (T_s=4000 s) enables the calculation for rather wide range of observation intervals (form 0.1 s till 1000 s), although the requirement for the relation between τ_{max} and T_s was not fulfilled. The segment's length (T_s=10000 s) in the case of procedure D was too long with the relation to the data sequence length (T=20000 s). The changes affect the results of calculation for almost all segments. It was hard to identify the form and instant of the change.
The results of calculations using procedures B and C enable better recognizing the variations in the analyzed data sequences. In general, it is very hard to find some compromise between the length of data sequence T, the length of data segment T_s and the maximum observation interval τ_{max}. If we want to obtain good dynamic character of the results using non-overlapping segments (as in the case of procedure B), the length of segments must be rather short (T_s<<T) and therefore τ_{max} is rather short. In order to perform the analysis for longer observation intervals, we must use the overlapping data segments with extended length.

6. CONCLUSIONS

The application of the dynamic ADEV and dynamic TDEV to the evaluation of timing signals enables to recognize some features of the signals, which cannot be found using the common Allan deviation or time deviation. Particular lengths and quantities should be chosen carefully before the calculation process. In order to obtain "dynamic" character of the expected result, a number of segments greater than 10 is suggested. The segment's length should be very short in comparison with the length of the whole data sequence in the case of non-overlapping segments (e.g. T_s<0.1 T). The segment's length in the case of overlapping segments cannot exceed the quarter of the data sequence's length.

REFERENCES

[1] ANSI T1.101-1999, Synchronization interface standard
[2] ETSI EN 300 462, Generic requirements for synchronization networks (1998).
[3] ITU-T Rec. G.810, Considerations on timing and synchronization issues (08/96).
[4] L. Galleani, P. Tavella, The characterization of clock behavior with the dynamic Allan variance, in Proceedings of the 2003 IEEE International Frequency Control Symposium and 17th European Frequency and Time Forum, Tampa, USA, 5-8 May 2003, pp. 239-244.
[5] A. Dobrogowski, M. Kasznia, Time effective methods of calculation of dynamic parameters of synchronization signal, in Proceedings of the 18th European Frequency and Time Forum, Guildford, UK, 5-7 April 2004.

CALIBRATION SYSTEM OF INTERNAL DELAY DIFFERENCE FOR TWSTFT STATION

Miho Fujieda, Masanori Aida, Hideo Maeno and Lam Quoc Tung
National Institute of Information and Communications Technology
4-2-1, Nukui-Kitamachi, Koganei, Tokyo, 184-8795, Japan
tel : +81-42-327-6316, fax : +81-42-327-6664
e-mail : miho@nict.go.jp

Tomonari Suzuyama and Michito Imae
National Institute of Advanced Industrial Science and Technology
1-1-1, Umezono, Tsukuba, Ibaraki, 305-8563, Japan

ABSTRACT

Two Way Satellite Time and Frequency Transfer(TWSTFT) enables us to perform time transfer precisely, for atmospheric and geometric delays between the earth station and the satellite are almost canceled because of the path reciprocity. However, the delays in the transmission and the reception paths of the earth station cannot be cancelled due to the different signal paths. The internal delay difference makes the absolute value of the time difference unknown, while the variation in the delay difference degrades the stability of the time transfer.

We have developed an internal delay-difference calibration system using a multi-channel TWSTFT modem. Because the signals for the calibration are processed with the received time-transfer signals from other stations, not only a short-term variation but also an annual change in the delay difference can be measured. It is found that the diurnal variation of the delay difference becomes a few hundred ps even if the low noise amplifier works in the temperature-controlled box. As for the time transfer result between two H-masers, its stability was improved by the correction of the delay difference. We confirmed the calibration system is important to improve the precision of TWSTFT.

1. INTRODUCTION

It is considered that Two-way Satellite Time and Frequency Transfer(TWSTFT) is most precise method at this moment because of its path reciprocity. In the recent TWSTFT system, measurement of code phase is widely used and its precision is almost fixed from its code rate and the line quality of the link via satellite. If more precise time transfer is necessary, we can make the code rate and the line quality higher. The internal-delay variation in the earth station for TWSTFT limits the precision and degrades the stability, however. Time transfer result between the earth stations A and B is presented as follows;

$$\begin{aligned} \Delta t^A - \Delta t^B &\cong \frac{1}{2}[(t^A - t^B) + t_{Sagnac} \\ &+ (d^A_{Tx} - d^A_{Rx}) - (d^B_{Tx} - d^B_{Rx})]. \quad (1) \end{aligned}$$

Here, $\Delta t^A - \Delta t^B$ means the time difference and t^A, t^B do their reception time. t_{Sagnac} means the delay difference due to the Sagnac effect. d^A_{Tx}, d^A_{Rx} are the internal delays of the transmission and reception paths in the earth station A, respectively. And d^B_{Tx}, d^B_{Rx} are similar. Because the internal delay-difference, $(d_{Tx} - d_{Rx})$, is included in the equation (1), it is necessary for its estimation in order to obtain an accurate time difference. Besides, its variation need to decrease as low as possible to improve the stability of time transfer. It is reported the time transfer results showed diurnal variations arriving at a few ns[1]. Not only to stabilize the environment for equipments but also to measure the variation and compensate it is effective for stability of TWSTFT.

We have developed the calibration system of the internal delay-difference using a multi channel TWSTFT modem[2, 3]. The modem can generate second transmission signal to measure the reception path delay and another transmission signal is used for usual time transfer and measurement of the transmission path. Those delay measurements are performed simultaneously with time transfer. This paper reports about our system and the measurement results.

2. CALIBRATION SYSTEM

2.1 Multi-channel TWSTFT modem

We have developed a multi-channel modem called as NICT modem and started to report its time-transfer results to BIPM on trial[4]. If the trial run is succeeded, TWSTFT in the Pacific-Rim region will be performed by using the NICT modem instead of the ATLANTIS modem.

The NICT modem has two transmission channels and eighth reception channels. One transmission channel is used for usual time transfer and the measurement of the delay in the transmission path. The other is used for measurement of the delay in the reception path. The eighth reception channel is used for the reception of the signal to measure the delay in the transmission path. The eighth reception channel has an independent input connector from other reception channels, for the structure makes us to differentiate the round-trip-delay from the calibration signals with same PRN(Pseudo Random Noise) number.

2.2 Internal delay measurements part

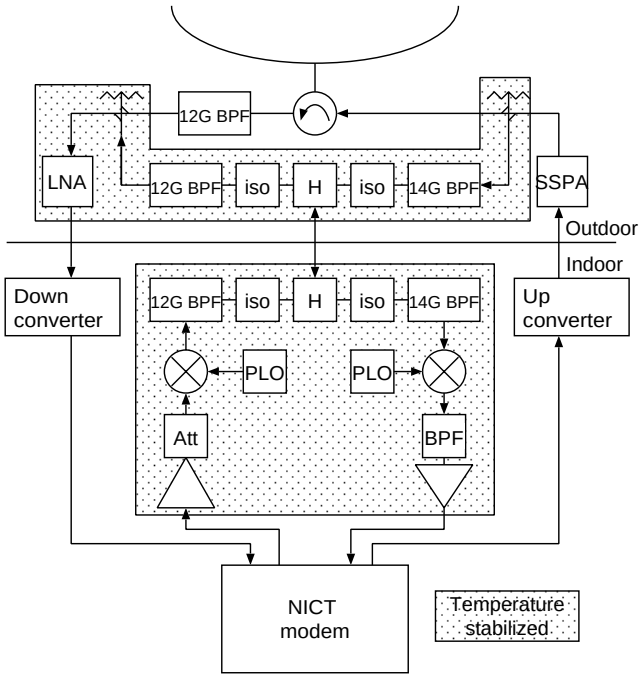


Figure 1: Block diagram of the earth station. Components in the dotted area are mounted in the temperature-controlled chamber.

Figure 1 shows the block diagram of the earth station in NICT including the calibration system, where "iso" and "PLO" mean isolator and phase locked oscillator, respectively. The signals to measure the delays in the transmission and reception paths are named as Txcal and Rxcal here, respectively. Their paths of Txcal and Rxcal are outlined below;

• Txcal
The Txcal is common with usual transmission signal for time transfer. Its portion is clipped by the coupler in front of the antenna and inserted into the path for Txcal. It is transmitted through the cable for calibration to indoor and converted to IF frequency. Then it is inputted to the eighth reception channel specialized for Txcal and processed there.
• Rxcal
The Rxcal is generated by the second transmission channel of the modem. It is converted to the RF frequency of reception and transmitted to outdoor, where the same RF cable as Txcal is used. Then the Rxcal is inserted into the usual reception path by the coupler before the LNA. It is overlapped with the reception signals from other stations and processed simultaneously at the modem.

Main constitutions of the calibration system are third RF cable between indoor and outdoor, indoor equipments including the frequency converters and PLO and outdoor equipments including couplers in front of the antenna. In the system, the modem, up and down converters are placed indoor room. The room temperature variation is controlled within ±1°C. It is varied if someone gets into the room only. The SSPA, LNA and others are placed outdoor and the LNA is mounted in the temperature-controlled chamber where the specification of temperature variation is within ±1°C and the actual variation is much lower. The RF cables with low temperature coefficient, which are the Andrew FSJ1-50A, connected with the indoor and outdoor equipments. The RF components for the calibration are installed the temperature-controlled chamber which is placed indoor room. The temperature of the chamber changes only ±0.1°C. It is said that the delays in the RF cable and SSPA vary mainly about the Txcal and that in the RF cable causes variation mainly in the Rxcal accordingly.

In order to compensate the time-transfer result, necessary delay-difference value, ΔT, is presented using the description seen in Figure 2 as follows;

$$\begin{aligned}\Delta T &= \frac{1}{2}(d_{Tx} - d_{Rx}) \\ &= \frac{1}{2}[(tx_modem + IF_up + up_conv + hpa \\ &+ up_feed) - (down_feed + lna + down_conv \\ &+ IF_down + rx_modem)]\end{aligned}$$

On the other hand, the measured delays of Txcal and

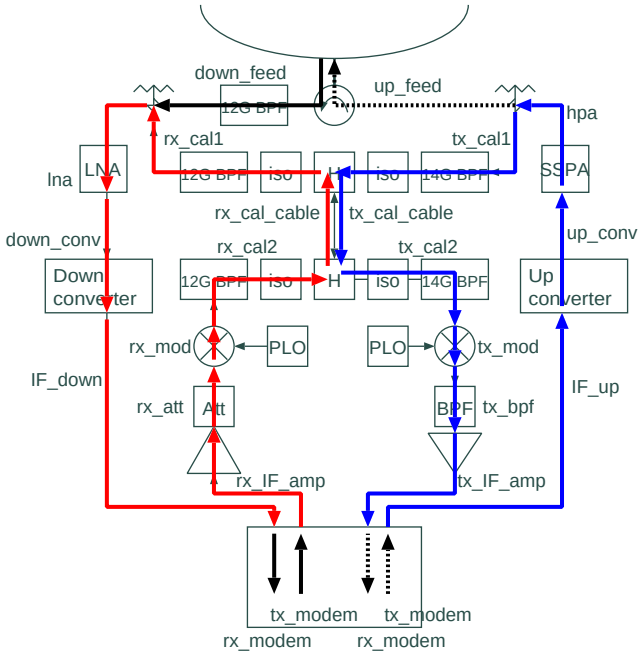


Figure 2: Block diagram of internal delay-difference calibration system.

Rxcal, ΔT_{Tx} and ΔR_{Rx} were written as follows;

$$\begin{aligned}\Delta T_{Tx} &= (tx_modem + IF_up + up_conv + hpa \\ &+ tx_cal1 + tx_cal_cable + tx_cal2 + tx_mod \\ &+ tx_bpf + tx_IF_amp + rx_modem) \\ \Delta T_{Rx} &= (tx_modem + rx_IF_amp + rx_att + rx_mod \\ &+ rx_cal2 + rx_cal_cable + rx_cal1 + lna \\ &+ down_conv + IF_down + rx_modem)\end{aligned}$$

Using the measured value of ΔT_{Tx} and ΔT_{Rx}, ΔT is written as follows;

$$\begin{aligned}\Delta T &= \frac{1}{2}[(\Delta T_{Tx} - \Delta T_{Rx}) + (up_feed - down_feed) \\ &+ (tx_modem - rx_modem) + t_{bias}] \qquad (2) \\ t_{bias} &= (rx_att + rx_cal1 + rx_cal_cable + rx_cal2) \\ &- (tx_cal1 + tx_cal_cable + tx_cal2 + tx_bpf) \qquad (3)\end{aligned}$$

The term, $(tx_modem - rx_modem)$, remains in the equation (2). If the other station uses the same modem and their individual difference is negligible, the term is cancelled when the time difference is calculated. And if we can measure the delays of up_feed and $down_feed$ before construction of the earth station, the absolute value of delay in the earth station is possible to be fixed. We have not measured them, however. The delays in the IF amplifier is cancelled because they are inserted into both paths of Txcal and Rxcal. Those in the modulator, tx_mod and rx_mod, are done similarly. Those in the residual components, for examples, BPF, cables, are not cancelled because of the frequency difference in the transmission and reception. As for delays in the not-common path with Txcal and Rxcal, t_{bias} the group delay measurements are necessary in advance by using a network analyzer consequently. Because the equipments in their not-common paths are mounted in the temperature-controlled bath, their delays are able to be assumed as constant values. The value of t_{bias} is 48.563 ns in this system.

3. RESULTS

3.1 Internal delay variation

Figure 3 shows the time series plots of delays of Txcal and Rxcal from 11/20/2004 to 11/24/2004. The delays were measured every second and the outdoor temperature was done every one minute. Top means half of internal delay-difference, $\frac{1}{2}(\Delta T_{Tx} - \Delta T_{Rx})$, second is delay of Rxcal and third is delay of Txcal. Black dashed line shows the outdoor temperature the antenna placed. In order to get a good visibility, biased values were plotted. Because we stopped the temperature control for chamber the LNA mounted and adjusted the signal level around day 2 in the horizontal axis, their delays changed and there were jumps in data. Because the Txcal was processed in the independent channel and not overlapped with other signals, it was measured with higher line quality of about 90dB·Hz. On the other hand, the measurement of Rxcal was done with about 60dB·Hz. We can see the difference in their data spreads in figure 3.

The diurnal variation in the outdoor temperature reached to about 12 °C in the period, while the indoor temperature varied about ±0.2°C. We can see that the delay variations synchronized with the outdoor temperature change in figure 3. There was no correlation with the indoor temperature. When the temperature-controlled chamber the LNA mounted worked, the delay variations of Txcal and Rxcal were about 200 ps and 150 ps, respectively. And when the temperature control for the chamber turned off, those of Txcal and Rxcal were about 300 ps and 400 ps, respectively. As for ΔT, its variations were about 180 ps and 360 ps with and without the temperature control, respectively. Using those results, we roughly estimated the temperature coefficient assuming that they were linear for the outdoor temperature. The temperature coefficient of SSPA and LNA were -30 ps/°C and 30 ps/°C, respectively. It is found that the delay-difference variation

becomes a few hundred ps when the LNA is installed into the temperature-controlled chamber and its compensation is required to improve the stability of TW-STFT.

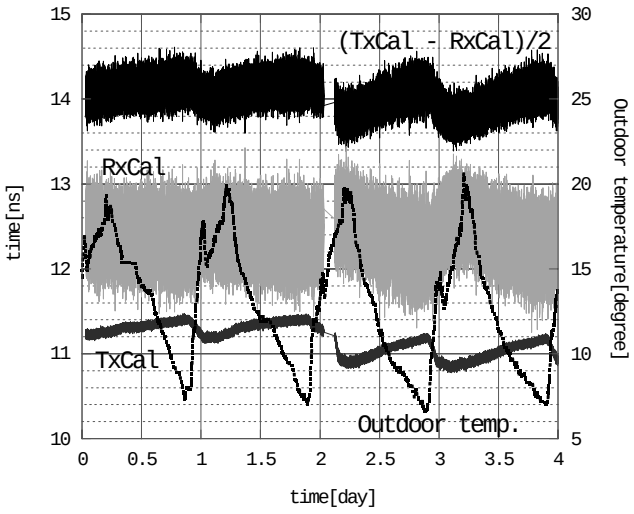


Figure 3: Time-series plot of internal delays. Horizontal axis is time, left vertical one means delay in ns and right vertical one is temperature. Top means half of internal delay-difference, second is delay in reception calibration path and third is delay in transmission calibration path. Black dashed line means outdoor temperature.

3.2 Application for time transfer

We implemented the time transfer experiment between NMIJ and NICT as well as the internal delay-difference measurement in the foregoing section. The calibration system was installed only in the earth station of NICT. The H-maser signals were used as the references for the modems in both sites. When the reference was given from Cs clock, the variation referred from the clock was much larger than that caused by the internal delay difference and it was difficult to see whether it made the time-transfer result degraded or not. Figure 4 shows the time-transfer results using H-masers. We removed the linear gradient due to the frequency difference from data. Top and bottom show results without and with correction of the internal delay-difference variation, respectively. Black and gray lines mean data taken in every seconds and white dotted lines mean 1 hour averaged data. At that time, the temperature control for the LNA chamber turned off. In case without the correction, we can see the diurnal variation in the time-transfer results synchronous with the variation of outdoor temperature. On the other hand, the diurnal variation decreases in case with the correction.

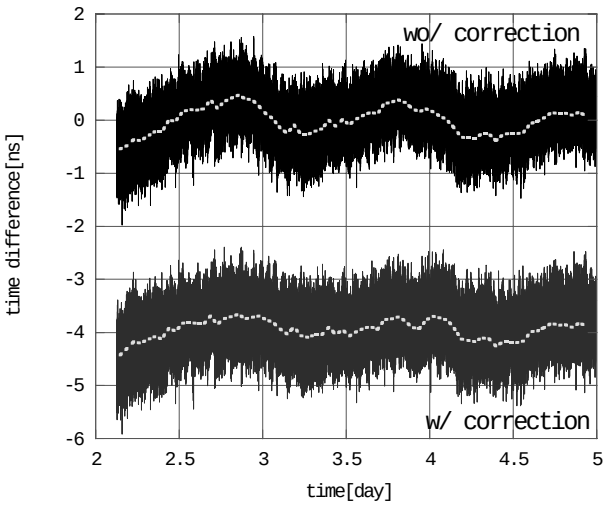


Figure 4: Time-transfer results between NICT and NMIJ using H-masers. Horizontal axis means time and vertical one is time difference in ns. Top and bottom show results without and with correction of internal delay-difference, respectively. White dotted lines mean 1 hour averaged values.

The results agree that the correction works well.

Figure 5 shows the stabilities of time transfer, which was calculated from the values of 1 hour averaged of the figure 4. Black circles and white squares mean without and with the correction, respectively. The correction of time-transfer result by using the measured delay-difference values makes the data spread larger as seen in the equation (1). The short term stabilities degrade with the correction because there is no correlation between the signal to measure the internal delay and its round-trip-delay via satellite. The stabilities of the averaging time of 1 hour show the similar values with and without the correction. It is found that to use the averaged data over 1 hour is effective not to degrade the stability of time transfer.

If there is a diurnal variation like a sine wave, the stability at the averaging time of 43200 s worsens most. The stabilities with the correction around there improve in the figure 5.

The stabilities at the averaging time of 43200 s without and with the correction are presented as σ_o and σ, respectively. The stability of the internal delay-difference is written as σ_{cal}. We estimated them from the measurements as follows;

$$\sigma_o = 1.38 \times 10^{-14}$$
$$\sigma = 7.00 \times 10^{-15}$$

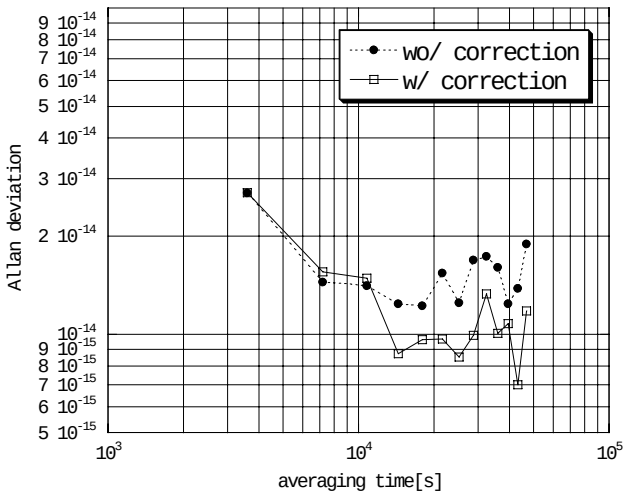


Figure 5: Stability of time transfer between NICT and NMIJ using hydrogen masers. Horizontal axis means averaging time and vertical one is Allan deviation.

$$\sigma_{cal} = 7.54 \times 10^{-15}$$

If there is no correlation between the time-transfer result and the internal delay-difference, the stability is presented as the root sum square, as below;

$$\begin{aligned} \sigma_o &= (\sigma^2 + \sigma_{cal}^2)^{1/2} \\ &= 1.03 \times 10^{-14} \end{aligned}$$

The number above noted is almost consistent with the estimated value from the measurement. Then, we made the 1 hour averaged data FFT and checked the amplitudes of diurnal variation. Their amplitudes are as follows;

- wo correction : 102 ps
- w correction : 50 ps
- internal delay difference : 59 ps

From this result by FFT, the sum of amplitude with correction and that of the internal delay difference becomes 109 ps, which is roughly equal to the value of 102 ps. These examination results prove that the correction is effective and decreases the amplitude of diurnal variation in the time transfer by half. The stabilities at the averaging times over 10^4 s became flat, however, which did not reach at the stability of $\sim 3 \times 10^{-15}$ of H-maser. It is clear that any variation remain in the time transfer result. The internal delay-difference was compensated for only one station at this moment. It seems that the delay-difference variation in the other station degrades the stability of time transfer.

4. CONCLUSION

We have developed the calibration system of internal delay-difference in the earth station for TWSTFT. The NICT modem employed in the system has dual transmission channels, seventh reception channels and one reception channel specialized for the measurement of delay in transmission path. The system has been composed of common RF components consequently. The delays in the transmission and reception paths were measured simultaneously with time transfer between NICT and NMIJ using H-maser signals as references. The LNA was mounted in the temperature-controlled chamber and other instable equipments for temperature were placed in the indoor room. The variation of internal delay-difference became 15 ps/°C, however. The result shows that the delay-difference may change a few hundred ps not only during a day but also between summer and winter. The stabilities of time transfer around 43200 s were improved by the calibration using 1 hour averaged data. It encourages us to compensate the delay-difference variation in order to upgrade the stability of TWSTFT. The delay-difference measurements have been tried in the odd site at this moment. If the calibration systems work in both sites, the effect for time transfer might be clearer.

Acknowledgments

We would like to thank TWSTFT staffs in NICT, NMIJ, KRISS and TL.

References

[1] D. Matsakis, "Time Transfer Research and Development at USNO", 44th CGSIC meeting, Timing Subcommittee, Long Beach, CA, September, 2004.

[2] M. Imae et al., "DELAY CALIBRATION METHOD FOR PRECISE AND ACCURATE TWO WAY SATELLITE TIME AND FREQUENCY TRANSFER", in the conference digest on CPEM, 2002, pp. 450 - 451.

[3] H. Maeno et al., "Development of Multi-channel Time Transfer Modem for TWSTFT", in the proceedings on ATF 2004, pp. 267-272.

[4] H. Maeno et al., "TWSTFT plans in the Pacific-Rim region and Europe link", in this proceedings.

DETERMINATION OF GPS/GALILEO TIME OFFSET: INSIGHTS FROM THE GSTB-V1 RESULTS

I. Sesia[*°], D. Orgiazzi[*], P. Tavella[*]

[*] Istituto Elettrotecnico Nazionale "Galileo Ferraris", Strada delle Cacce 91, Torino, Italy, sesia@ien.it

[°] Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy

ABSTRACT

Some preliminary evaluations on the feasibility of the GPS/Galileo Time Offset (GGTO) determination are taken into account. Using the data coming from the Galileo System Test Bed (GSTB) V1 experiment as well as additional measures and IEN experience on timing aspects, possible scenarios to estimate GGTO were experienced also estimating the uncertainty, along with possible advantages of the different estimation procedures.

The first experimental results are reported and discussed pointing out the criticalities encountered and the solutions suggested for reaching the demanding goals.

1. INTRODUCTION

The GPS Joint Program Office and the Galileo Project Office have recently signed an agreement [1] on the interoperability of the two navigation satellite systems asking, among other things, for a close coordination of the two system time scales, namely the GPS Time and the Galileo System Time (GST).

In particular, the offset between GPS Time and GST, named as GGTO, has to be independently determined by both systems and then broadcasted in the navigation message of each system. By that, both systems can ensure to the user receiving both GPS and Galileo signals an improved three dimensional position accuracy and service coverage [2, 3].

During the Galileo In-Orbit Validation Phase (IOV) scheduled in 2007, the four Galileo satellites will begin broadcasting the GGTO information in the navigation message. The current GGTO specification asks for both an accuracy less than 5 ns (2σ) and a frequency stability better than 8×10^{-14} over 1 day, even if improved figures are expected by the time Galileo becomes operational [1].

At the Istituto Elettrotecnico Nazionale (IEN) "Galileo Ferraris", Turin (Italy), some preliminary evaluations on the feasibility of GGTO determination at the requested level of accuracy as well as different measurement scenarios have been experienced, relying on first experimental results from the GSTB-V1.

The GSTB-V1 was the first experimental phase of the Galileo project supported by the European Space Agency (ESA) with the aim to mitigate the programme risks [4].

In the GSTB-V1 infrastructure, an Experimental Precise Timing Station (E-PTS) has been designed by IEN in collaboration with Alenia Spazio, and it has been installed and continuously operated since 2003 at IEN premises in Turin, Italy. Scope of E-PTS element was to generate a real-time hardware Experimental Galileo System Time (EGST), which is the reference time for the whole GSTB-V1 allowing the realization of a large set of experiments in the timing domain. In particular, EGST was based on IEN atomic clocks, namely an ensemble of Cesium clocks and Hydrogen maser, applying a proper algorithm also ensuring the steering of EGST to the international reference time TAI and UTC [5, 6].

Taking advantage of the fact that EGST was a real-time physical signal available at IEN during year 2004, possible scenarios to estimate GGTO have been experienced outside the main scope of GSTB-V1 using additional data and evaluation at IEN, also with the contribution of U.S. Naval Observatory (USNO) data. The GGTO was determined by three different methods:

- by means of a multi-channel single-frequency GPS timing receiver, connected to UTC(IEN) time scale, performing "common view" measurements on GPS signal in space (SIS) which are referred to EGST by using local measurements at E-PTS,
- by means of the E-PTS Sensor Station, that is a multi-channel double-frequency geodetic GPS receiver directly connected to the hardware realization of EGST,
- by means of a Two Way Satellite Time and Frequency Transfer (TWSTFT) link between UTC(IEN) and UTC(USNO) time scales, using then the UTC(USNO) to GPS Time offset as provided by USNO to obtain the UTC(IEN) to GPS Time offset, and finally using E-PTS local measurements to refer to EGST.

Each scenario above is detailed in the following sections, where some experimental results on GGTO evaluation are also reported in terms of accuracy estimated considering the BIPM data as a reference.

In fact, using the BIPM data reported in Circular T, namely GPS Time versus UTC and UTC(IEN) versus

UTC offsets, an additional GGTO determination can be achieved:

$$GGTO_{BIPM} = (UTC - UTC(IEN))_{BIPM} + \\ -(UTC - GPS_{Time})_{BIPM} + \\ -(EGST - UTC(IEN))_{Meas.\ System} \quad (1)$$

where the last term results from the local measurement system at E-PTS. This evaluation may be obtained on BIPM Standard Dates only, spaced by 5 days, and the estimates are to be intended as instantaneous values dated at 00h UTC. Furthermore, a frequency stability analysis in term of Allan deviation has been conducted for the medium-long term (in order to check versus the GGTO requirements) as well as for the short term, dealing about the potential of a "near-real time" estimation.

2. FIRST SCENARIO: GSTB-V1 NOMINAL PROCESSING

According to the tasks allocated to GSTB-V1 experimentation, the offset between GPS Time and EGST was evaluated using the local UTC(IEN) realization of UTC as intermediate time scale by means of the following expression:

$$GGTO_{GSTB\ V1} = (GPS_{Time} - UTC(IEN))_{3SN} + \\ -(EGST - UTC(IEN)_{Meas.\ System} \quad (2)$$

where (see Figure 1) :

- $(GPS_{Time} - UTC(IEN))_{3SN}$ was estimated once per day (at 00h UTC) by a specific algorithm developed by IEN [7], processing the raw measurements (over 24-hours interval centered at 00h UTC) provided by a 3S Navigation (3SN) multi-channel single-frequency GPS timing receiver, using UTC(IEN) as external time reference,
- $(EGST-UTC(IEN)_{Meas.\ System}$ was the calibrated measure at 00h UTC of the same day taken by the E-PTS measurement system.

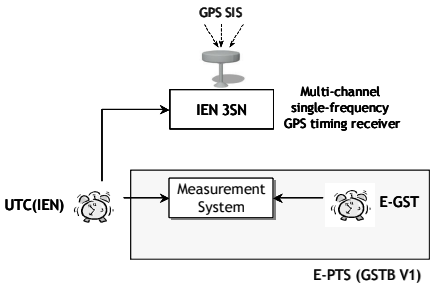


Figure 1: Architecture of GSTB-V1 Nominal Processing scenario.

The resulting $GGTO_{GSTB\ V1}$ for this scenario is plotted in Figure 2, where the estimated $GGTO_{BIPM}$ (1) is also reported as a reference.

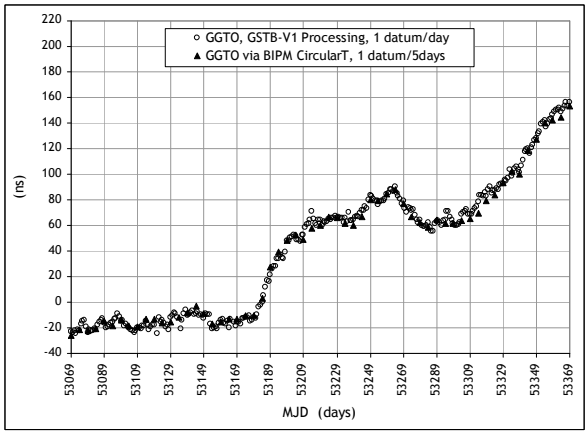


Figure 2: Time series of $GGTO_{GSTB\ V1}$ daily estimates as provided by the GSTB-V1 Nominal Processing scenario (symbol **o**). $GGTO_{BIPM}$ (5 days spaced, symbol ▲) is also reported. The analysis period is from MJD 53069 (05/03/2004) to MJD 53369 (30/12/2004).

It can be seen that the GSTB-V1 Nominal Processing estimates are in good agreement with BIPM Circular T data. In particular, the difference between the two estimates (plotted in Figure 3) falls within a few nanoseconds, around a 0.8 ns mean value and with a 3.6 ns standard deviation.

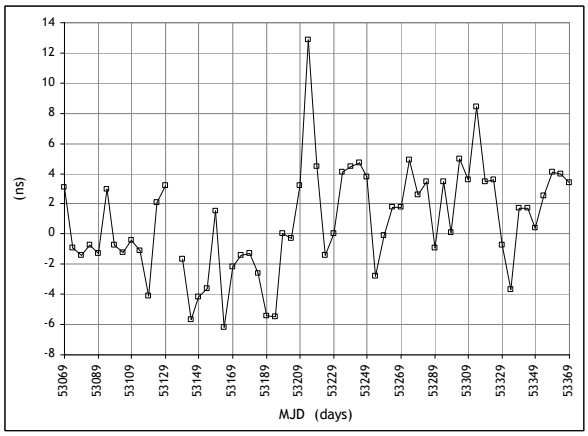


Figure 3: Differences between $GGTO_{GSTB\ V1}$ and $GGTO_{BIPM}$ estimates, 1 datum/5 days, from MJD 53069 (05/03/2004) to MJD 53369 (30/12/2004).

As expected, a very close to zero mean difference is achieved thanks to the calibration of both the 3SN receiver, regularly performed by BIPM [8], and the measurement system at E-PTS.

Besides, it follows that a GGTO estimate at the level of 5 ns (1σ) uncertainty seems then feasible. However, it's worth to notice that such a result is twice the requested uncertainty, asking for 5 ns (2σ).

As a matter of fact, also the BIPM Circular T reports UTC versus GPS Time with a global uncertainty of 10 ns,

therefore it seems really the obtainable limit of this technique.

2.1 GGTO uncertainty

For the GSTB-V1 Nominal Processing scenario, the uncertainty associated to each daily $GGTO_{GSTB\ V1}$ estimate comes out from the two different contributions in equation (2), namely:

- the uncertainty on the $(GPS_{Time} - UTC(IEN))_{3SN}$ estimate,
- the uncertainty on the $(EGST\text{-}UTC(IEN))_{Meas.Sys.}$ measure.

Because of the high-resolution time interval counter (TIC) and the careful calibration of the EGST distribution chain on which the E-PTS measurement system is based on, the latter contribution can be negligible in this context.

On the contrary, the receiver noise and the uncertainty of the calibration of the receiver itself makes significant the former contribution.

As investigated in [7], the IEN 3SN multi-channel GPS timing receiver shows a typical white phase noise with a figure of 1.6 ns (1σ) at one day observation interval.

Besides, according to the BIPM Circular T data, a 5 ns (1 σ) uncertainty amount has to be considered for the calibration of the receiver. Summing up these results, an uncertainty of about 5.2 ns (1 σ) can be obtained in the GGTO determination, mainly relied to the GPS receiver calibration capability.

2.2 GPS smoothing algorithms for GGTO evaluation

When using a GPS timing receiver in the GGTO evaluation, it is known that the added measurement noise may be significant and some estimation and smoothing techniques have to be applied. In this study we considered a one day batch smoothing algorithm (fully described in [7]), whose main processing steps are the following:

1. outliers identification and removal,
2. average of the different observation at the same epoch,
3. estimation of a daily averaged value.

In particular, the daily value is computed using a linear regression of all the available measures within a 24 hours interval centred at 00h UTC of each day, aiming to filter out any short-term instabilities.

The considered method allows to evaluate a mean daily value; nevertheless, the epoch values from step *2* can be used as an iterative "real time" estimation of GGTO. A GGTO dynamic estimation technique based on Kalman filter application is actually under investigation (other studies on different GGTO filtering techniques are addressed in [2]).

3. SECOND SCENARIO: E-PTS SENSOR STATION BASED

On March 23rd 2004 (MJD 53087), a dual-frequency geodetic GPS receiver (Javad Legacy with timing option) operated by IEN has been physically connected to the hardware realization of EGST, as required by the final operation of the E-PTS Sensor Station (Figure 4).

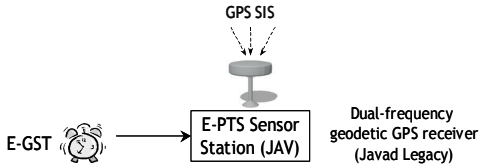


Figure 4: Architecture of E-PTS Sensor Station based scenario.

Since the switch-on, the E-PTS Sensor Station phase/pseudorange measurements were routinely collected with 1 Hz sampling rate and then processed according to the procedure described in [9] and by the IEN algorithm [7], providing then daily estimates of the GPS Time to EGST offset:

$$GGTO_{E\text{-}PTS\ SS} = (GPS_{Time} - E\text{-}GST)_{JAV,P3} \qquad (3)$$

This scenario has the advantage of being based on a single measurement system (the geodetic GPS receiver) directly estimating GGTO. The resulting $GGTO_{E\text{-}PTS\ SS}$ is shown in following Figure 5, together with the $GGTO_{BIPM}$ obtained by means of BIPM Circular T data (1).

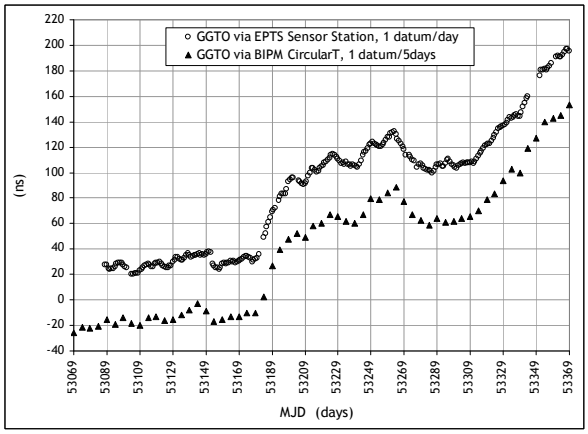


Figure 5: Time series of $GGTO_{E\text{-}PTS\ SS}$ daily estimates, as evaluated in the E-PTS Sensor Station scenario (symbol **o**). $GGTO_{BIPM}$ is also reported (symbol ▲). The analysis period is from MJD 53088 (24/03/2004) to MJD 53369 (30/12/2004).

The difference between the two considered estimates of GGTO is reported in Figure 6.

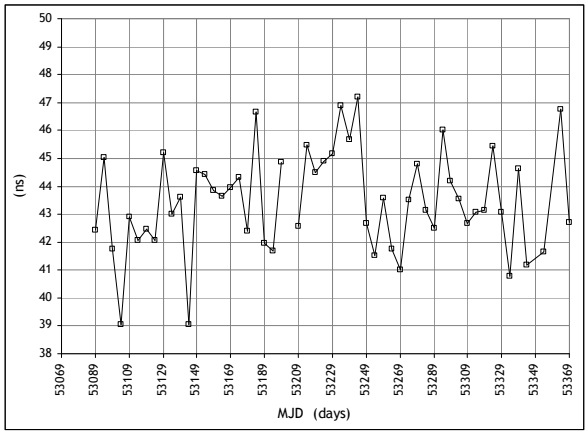


Figure 6: Differences between $GGTO_{E\text{-}PTS\ SS}$ and $GGTO_{BIPM}$ estimates, 1 datum/5 days, from MJD 53089 (25/03/2004) to MJD 53369 (30/12/2004).

As clearly pointed out from the time differences in Figure 6, the behaviour of the two estimates is very similar (only 1.7 ns standard deviation), even if a 43.5 ns mean value is noticed. Such bias with respect to BIPM Circular T data is due to the fact that the E-PTS Sensor Station was not calibrated in the frame of GSTB-V1 experiment, showing then its own electrical delay which is not compensated in this scenario.

3.1 GGTO uncertainty

Since EGST is a physical time scale directly connected to the geodetic GPS receiver, the uncertainty of the $GGTO_{E\text{-}PTS\ SS}$ estimated in this scenario is mainly driven by the capability of the receiver to estimate GPS Time from the GPS SIS.

To this figure, two additional contribution at least are to be added:

- the uncertainty of the receiver calibration,
- the intrinsic uncertainty in the dissemination of GPS Time by GPS SIS.

As stated before, the E-PTS Sensor Station was not fully calibrated in the frame of GSTB-V1. Nevertheless, the first term could be assumed to be 2.5 ns (1σ), which is the typical uncertainty affecting the differential calibration of geodetic GPS receivers [10]. Besides, the stability due to the receiver itself, over observation period of one day, was estimated to be 0.8 ns (1σ) [7].

Moreover, the latter term has to be considered in every scenario involving a GPS receiver (either timing or geodetic) to access to GPS Time. This is due to the fact that GPS Time, as broadcasted through GPS SIS, is actually a prediction of its own offset with respect to each satellite on-board clock.

Such a prediction is daily updated only possibly not in the same instant on all GPS satellites, so the GPS time that can be inferred by SIS is a mix of old and newly predicted time offsets. However, considering an average daily estimate (as in the previous and current scenario), this intrinsic uncertainty may be negligible, but it is worth mentioning that it may result in a very crucial uncertainty source in case of more frequent ("near real time") evaluation of GGTO.

Considering only the effects of the calibration and the receiver instability, the uncertainty of the method would amount to 2.6 ns (1σ), therefore not too much outside the specification of 5 ns (2σ), but this evaluation needs careful study of other possible different components.

4. THIRD SCENARIO: TWSTFT LINK WITH USNO

Following one of the approaches proposed in the agreement between the GPS Joint Program Office and the Galileo Project Office [1], GGTO has also been evaluated by means of direct comparison of appropriate reference clocks using a TWSTFT link between IEN and USNO laboratories (Figure 7).

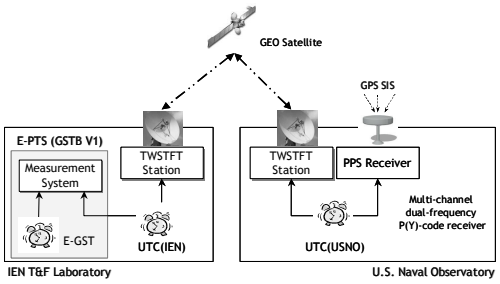


Figure 7: Architecture of the scenario based on the TWSTFT link with USNO.

With reference to the architecture above, GGTO results from the following expression:

$$GGTO_{TW} = -(EGST - UTC(IEN))_{Meas.\ System} + -(UTC(IEN) - (UTC(USNO))_{TWSTFT} + -(UTC(USNO) - GPS_{Time})_{USNO\ data} \quad (3)$$

where:

- $(UTC(USNO) - GPS_{Time})_{USNO\ data}$ are the data from the USNO time monitor station reference receiver, once per day centred at 00h UTC (provided by USNO [11]),
- $(UTC(IEN) - (UTC(USNO))_{TWSTFT}$ is the measure at 00h UTC of the same day, obtained by the IEN TWSTFT link with USNO, which has been calibrated a posteriori with respect to BIPM Circular T data,
- $(EGST - UTC(IEN))_{Meas.\ System}$ is the calibrated measure at 00h UTC of the same day taken by the E-PTS measurement system.

The daily $GGTO_{TW}$ estimates provided by this scenario are reported in Figure 8, along with the $GGTO_{BIPM}$ (1). Resulting differences from these two estimates are given in Figure 9, where significant lack of TWSTFT data can be noticed at the end of the considered period.

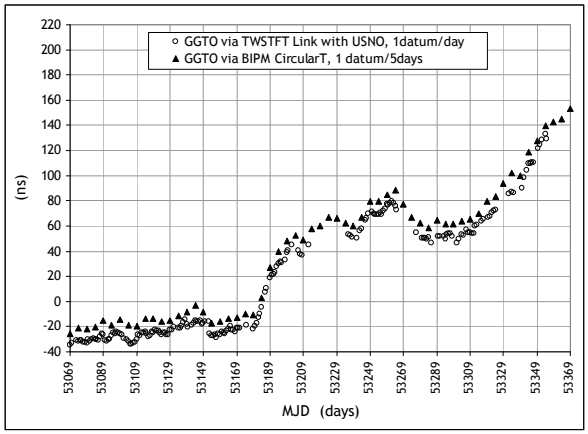


Figure 8: Time series of $GGTO_{TW}$ daily estimates as provided by the TWSTFT link with USNO scenario (symbol **o**). $GGTO_{BIPM}$ is also reported (symbol ▲). The analysis period is from MJD 53069 (05/03/2004) to MJD 53369 (30/12/2004).

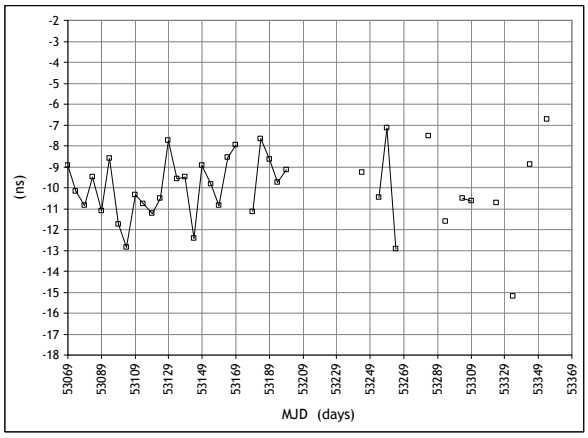


Figure 9: Differences between $GGTO_{TW}$ and $GGTO_{BIPM}$ estimates, 1 datum/5 days, from MJD 53069 (05/03/2004) to MJD 53369 (30/12/2004).

However, focusing on the first period only (up to MJD 53199 (13/07/2004)), it can be seen that the $GGTO_{TW}$ estimates from the TWSTFT link with USNO agree with $GGTO_{BIPM}$ within a 1.8 ns standard deviation. Besides, the 10 ns offset with respect to Circular T data could be probably due to the different data processing at BIPM and at USNO side, contributing a different reconstruction of the GPS Time.

4.1 GGTO uncertainty

GGTO uncertainty for this scenario is obtained as the sum of three contribution:

- uncertainty of $(EGST - UTC(IEN))_{Meas.\ System}$, which is negligible in this context (as seen before),
- uncertainty of $(UTC(IEN)-(UTC(USNO))_{TWSTFT}$, which currently is about 4 ns (1σ, according to Circular T data), but it can be reduced to 1 ns (1σ) in case the TWSTFT link is appropriately and repeatedly calibrated,
- uncertainty of $(UTC(USNO) - GPS_{Time})_{USNO\ data}$, which is about 2 ns [11] considering SIS and receiver errors and excluding the errors in absolute calibration (which would add several more nanoseconds).

Summing up these contributions, an uncertainty of about 5 ns (1σ) seems to be reachable, which is again twice the requested one.

5. COMPARISONS AMONG THREE SCENARIOS

The mean and the standard deviation values of the differences among the three GGTO estimates obtained considering the scenarios described above and $GGTO_{BIPM}$ (1) are summarized in Table 1.

	Residuals with respect to Circular T	
Different scenarios	*Mean*	*Sigma*
GSTB-V1 Nominal processing	0.8 ns	3.5 ns
E-PTS Sensor Station	43 ns	1.7 ns
TWSTFT Link with USNO	-10 ns	1.8 ns

Table 1: Residuals of different scenarios with respect to $GGTO_{BIPM}$ from Circular T data (1).

Apart from mean values, looking at the results reported in the table above, it stands out that the residuals obtained following the E-PTS Sensor station and USNO TWSTFT link scenarios are those characterized by the better standard deviation.

Moreover, the Allan deviation of the different estimates of GGTO (corresponding to the three scenarios outlined before) has been computed over the whole analysis period, in order to assess a medium long term frequency stability analysis (Figure 10).

As it comes out from Figure 10, at one day the best estimate of GGTO is given by the E-PTS Sensor Station scenario, where the uncertainty is mainly driven by the capability to estimate GPS time from SIS. While at long term the all three scenarios converge to the same stability,

which is about $1 \cdot 10^{-14}$ for $\tau = 16$ days, which can be due to the instabilities of GPS Time and EGST time scales.

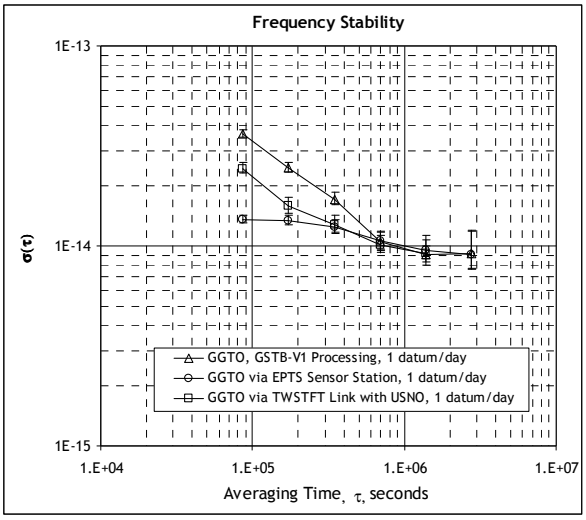


Figure 10: Allan deviation of the three different estimates of GGTO. The stability is computed over the whole experimentation period, from MJD 53069 (05/03/2004) to MJD 53369 (31/12/2004).

To check the impact of the satellite systems reference time scale instability, the Allan deviation of the considered GGTO estimates was computed over the first three months of GSTB-V1 experiment, during which EGST was characterized by a very good behaviour.

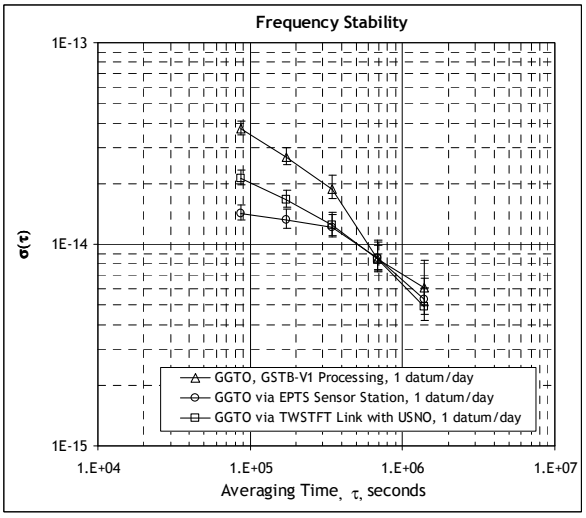


Figure 11: Allan deviation of the three different estimates of GGTO. The stability is computed over the first three months of the experiment, from MJD 53088 (24/03/2004) to MJD 53171 (15/06/2004), considered as the best period of EGST stability.

As it can be noticed in Figure 11, the three scenarios still converge to the same stability at long term, but to a much lower level corresponding to $\sigma_y(\tau) = 6 \cdot 10^{-15}$ for $\tau = 16$ days which seems to be ascribed by the better stability of EGST in this best run period.

As an additional analysis, the short term stability of GGTO (as estimated by the E-PTS Sensor Station scenario) has been evaluated, in order to assess the potential of a "near-real time" estimation.

The 1 Hz raw measures performed by the geodetic GPS receiver were processed according to the P3 procedure [9]; then, after an outliers identification and removal, an average of the different observation at the same epoch was performed, resulting in one datum per 16 minutes.

The Allan deviation of these 16 minutes data is reported in Figure 12, together with the Allan deviation of the $GGTO_{E\text{-}PTS\ SS}$ daily estimate (already reported in Figure 10).

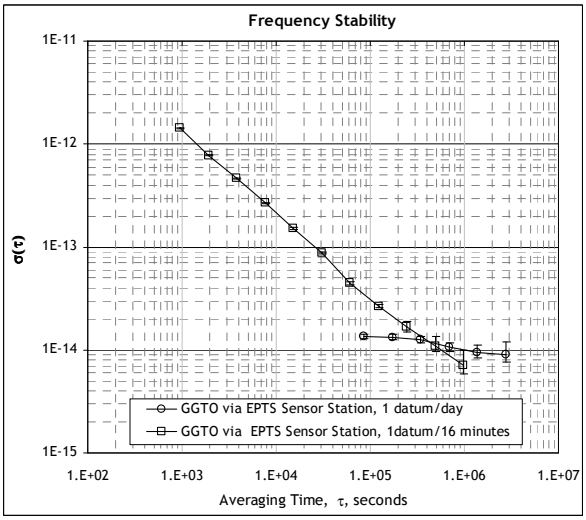


Figure 12: Short and medium-long term stability of GGTO for the E-PTS Sensor Station scenario: 1 datum/day (symbol o), 1 datum/16 minutes (symbol □).

As it stand out from this plot, even the 16 minutes GGTO estimates completely fulfil the requirement, asking for a frequency stability better than 8×10^{-14} over 1 day.

6. CONCLUSIONS

Different approaches to evaluate GGTO have been taken into account in this paper.

Taking advantage of the fact that EGST was a real-time physical signal available at IEN during year 2004, possible methods to estimate GGTO have been experienced, using additional data and evaluation at IEN, also with the contribution of USNO data.

GGTO was determined following three different scenarios:

- GSTB-V1 nominal processing,
- E-PTS Sensor Station,
- TWSTFT link with USNO.

In addition, as indirect validation, the BIPM Circular T data were used as reference, to check the accuracy on the different GGTO determination.

The mean and standard deviation values of the residuals of the three GGTO estimates with respect to BIPM Circular T data are reported. In this paper, the BIPM data were assumed to be the ultimate reference, nevertheless the BIPM data are also contaminated by uncertainty which is globally indicated around 10 ns in BIPM Circular T. Therefore the uncertainties values here reported should take care that this is the level of uncertainty which seems feasible today.

The first experimental results are discussed, and a preliminary evaluation of the uncertainty relating to each scenario is carried out.

Furthermore, a comparison among the three scenarios, in terms of frequency stability, has been performed.

Finally, a short term frequency stability analysis has been conducted for the E-PTS Sensor Station scenario, dealing about the potential of a “near-real time” estimation.

After an accurate analysis of these first experimental results, it seems possible to suppose that the best estimation of GGTO seems to be obtained using a scenario similar to the one based on the E-PTS Sensor Station , where the uncertainty is mainly driven by the capability to estimate GPS time from SIS. Further evaluations are still in progress.

ACKNOWLEDGEMENTS

The authors wish to thank the European Space Agency for granting the use of the GSTB-V1 data for these additional evaluations.

The authors would also like to express their gratitude to USNO for providing the USNO GPS measurements data required for the GGTO determination studies. A special thanks to Edward Powers (USNO) for kind support and helpful advices.

REFERENCES

[1] J. Hahn, E. Powers, “GPS/Galileo Time offset, Preliminary Interface document”, ESA-Navstar GPS Program Office, Issue 1, 20 March 2003.

[2] A. Moudrak., J. Hammesfahr, A. Konovaltsev,C. Bruyninx, P. Defraigne et al., “Determination of GPS-Galileo Time Offset to Support System Interoperability”, in Proceedings of 8th European Navigation Conference GNSS 2004, Rotterdam, The Netherlands, 17-19 May 2004.

[3] A. Moudrak, A. Konovaltsev, J. Furthner, A. Hornbostel, J. Hammesfahr et al., “GPS Galileo Time Offset: How It Affects Positioning Accuracy and How to Cope with It”, in Proceedings of ION GNSS 2004, Long Beach, CA, USA, 21-24 September, 2004.

[4] M. Falcone, F. Amarillo-Fernandez, J. Hahn, and H. Favin-Leveque, “The GALILEO System Test Bed for Early Prototyping and Experimentation of GALILEO Algorithms”, in Proceedings of the ION GPS 2002 Meeting, Portland, USA, September 2002.

[5] F. Cordara, R. Costa, L. Lorini, D. Orgiazzi, V. Pettiti, I. Sesia, P. Tavella, P. Elia, M. Mascarello, M. Falcone, J. Hahn, "Experimental Galileo System Time (E-GST): one year of real-time experiment", in Proceedings of the 36th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Washington DC, USA, December 2004.

[6] F. Cordara, R. Costa, L. Lorini, D. Orgiazzi, V. Pettiti, I. Sesia, P. Tavella, P. Elia, M. Mascarello, M. Falcone, J. Hahn, "Galileo System Test Bed V1: results on the Experimental Galileo System Time", to appear in Proceedings of 19th European Frequency and Time Forum (EFTF), Besançon, France, 21-24 March 2005.

[7] R. Costa, D. Orgiazzi, V. Pettiti, I. Sesia, P. Tavella, "Performance Comparison and Stability Characterization of Timing and Geodetic GPS Receivers at IEN", in Proceedings of 18th European Frequency and Time Forum (EFTF), Guilford, UK, April 2004.

[8] W. Lewandowski, and L. Tisserand, “Determination of the differential time corrections for GPS time equipment located at the OP, NPL, IEN, PTB, and VSL”, Rapport BIPM-2004/05, 2004.

[9] P. Defraigne and G. Petit, "Time Transfer to TAI Using Geodetic Receivers", in Metrologia (40), August 2003, pp. 184-188.

[10] G. Petit, Z. Jiang, P. Moussay, J. White, E. Powers, G. Dudle, P. Uhrich, "Progress in the Calibration of ‘Geodetic Like’ GPS Receivers for Accurate Time Comparisons", in Proceedings of 15th European Frequency and Time Forum (EFTF), Neuchatel, France, 2001, pp. 164-166.

[11] E. Powers, USNO, private communication 2005.

Frequency biases calibration of GLONASS P-code time receivers

A. Foks[1], W. Lewandowski[2], J. Nawrocki[3]

[1]Space Research Center, Warsaw, Poland, afoks@cbk.waw.pl
[2]Bureau International des Poids et Mesures, Sèvres, wlewandowski@bipm.org
[3]Astrogeodynamical Observatory, SRC, Borowiec, Poland, nawrocki@cbk.poznan.pl

Abstract

Although not finished and not so widely known as the GPS, the Russian Navigation Satellite System (GLONASS) provides interesting opportunities for international time metrology. Its P-code signal shows an outstanding performance. It especially allows calibration of GLONASS time receivers with an uncertainty below 1 ns.

Our paper reports latest exercise of differential calibration of GLONASS P-code time receivers located at the AOS, BIPM, VSL and USNO, by a means of a portable receiver. Determination of receiver time biases for different GLONASS frequencies is detailed. Removing these biases, reaching sometimes 10 ns, allows time transfer with an accuracy around 1 ns. It is also shown that use GLONASS P-code is improving significantly the stability of time transfer.

1 Introduction

Over the last 50 years the accuracy of atomic clocks has improved by an order of magnitude every seven years. Today they reach a stability in frequency of 1 part in 10^{16}. The stability of the international reference time scales, International Atomic Time (TAI) and Coordinated Universal Time (UTC), is presently about 10^{15} over a period of a few weeks. Meanwhile the standard uncertainty of GPS C/A-code common-view, currently the most widely used method of time transfer (and the main method providing the data for the calculation of TAI and UTC), is still of a few nanoseconds.

To enable comparisons of increasingly stable clocks it is important to improve time transfer methods. One of the new approaches to time and frequency comparisons consists of using GLONASS (Global Navigation Satellite System) P-code. Use of GLONASS in standard CGGTTS common-view mode was introduced into time metrology in 1996 by W. Lewandowski et al.[1, 2]. Studies of this approach have been continued by J. Azoubib, W. Lewandowski and J. Nawrocki in 2000 and 2001 [3, 4].

2 Advantages of GLONASS

The GPS and GLONASS systems share basically the same concept. A substantial difference between GPS and GLONASS is in the signal structure. GPS uses Code Division Multiple Access (CDMA): every satellite transmits the same two carriers modulated by PRN-codes particular to each satellite. GLONASS uses Frequency Division Multiple Access (FDMA): two individual carrier frequencies are assigned to each satellite, but the PRN-codes are the same for all satellites.

		GLONASS	GPS
operator		Military Space Forces (Russia)	Department of Defense (USA)
nominal number of satellites		24*	24
orbit altitude		19 100 km	20 200 km
orbital planes		3	6
orbital inclination		64.8°	55°
period of revolution		~ 11 h 16 min.	~ 11 h 58 min.
ground track repetition		8 sidereal days	1 sidereal day
reference frame		PZ-90	WGS-84
time reference		UTC(SU)	UTC(USNO)
ephemeris representation		position, velocity and acceleration	orbital parameters
signal separation		Frequency Division Multiple Access (FDMA)	Code Division Multiple Access (CDMA)
carrier frequencies	L1	1602.00 – 1614.94 MHz	1575.42 MHz
	L2	1246.00 – 1256.06 MHz	1227.60 MHz
	L1/L2	9 / 7	77 / 60
code chip rate	C/A	0.511 MHz	1.023 MHz
	P	5.113 MHz	10.23 MHz
code length (bits)	C/A	511	1023
	P	$5.113 \cdot 10^6$	$2.3547 \cdot 10^{14}$
navigation message length		2 min. 30 s	12 min. 30 s

*The present number of GLONASS satellites is below its nominal value.

Table 1. GPS vs. GLONASS.

Hence the GPS carrier frequencies f_{L1} and f_{L2} are invariable

$$f_{L1(GPS)} = 1575.42 \text{ MHz}$$
$$f_{L2(GPS)} = 1227.60 \text{ MHz}$$

and the individual GLONASS carrier frequencies f_{L1}^{k} and f_{L2}^{k} are defined by:

$$f_{L1(GLONASS)}^{k} = 1602.0 \text{ MHz} + 0.5625k \text{ MHz}$$
$$f_{L2(GLONASS)}^{k} = 1246.0 \text{ MHz} + 0.4375k \text{ MHz}$$

where k is the carrier number assigned to the specific satellite.

Both GPS and GLONASS have freely accessible C/A-code that modulates L1 only. Like the GPS precision code, the GLONASS P-code modulates both carriers, but unlike GPS it is also freely accessible.

The GLONASS P-code has two main advantages for high-precision time transfer.

Firstly, the GLONASS P-code modulation onto both L1 and L2 carrier frequencies allows high-precision measurements of ionospheric delays.

Secondly, the GLONASS P-code chip rate is one-tenth that of the GLONASS C/A-code and one-fifth that of the GPS C/A-code. This means that GLONASS P-code pseudo-range measurements are much more precise than GPS or GLONASS C/A-code measurements.

3 Determination of GLONASS frequency biases

The differential calibration of a GLONASS time receiver is realized by a one-site comparison with a portable receiver through common views [5, 6]. Both receivers are connected to the same clock and the antennas are separated by at most a few metres.

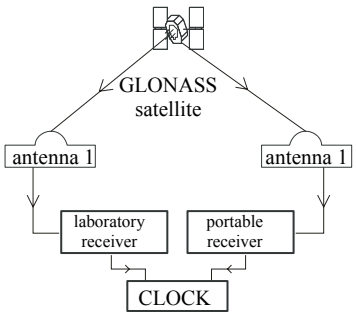


Figure 1. Scheme of differential calibration.

Such a comparison at short distance eliminates the common-clock error and errors resulting from satellite broadcast ephemerides, ionospheric and tropospheric delays, and antenna coordinates; only the noise of the receivers is observed.

The bias for the frequency k with reference frequency No. 6 is given by

$$B_k = \left(\overline{dt_i}\right)_6 - \left(\overline{dt_i}\right)_k$$

where $\left(\overline{dt_i}\right)_k$ is the mean value of the GLONASS P-code common views using the frequency k over the whole interval of computation.

The most recent differential calibrations of GLONASS equipment were made in the late 1990s.

We detail measurements that took place at the Bureau International des Poids et Mesures (BIPM), the Astrogeodynamical Observatory of Space Research Centre, Polish Academy of Sciences (AOS, Borowiec), Royal Observatory of Belgium (ORB, Brussel) and the Nederland Meetinstituut, Van Swinden Laboratory (NMi- VSL, Delft). Receiver No. 0017 (BIPM A) was chosen as reference; it had been compared with portable BIPM receiver No. 0019 in 1997 [5, 6].

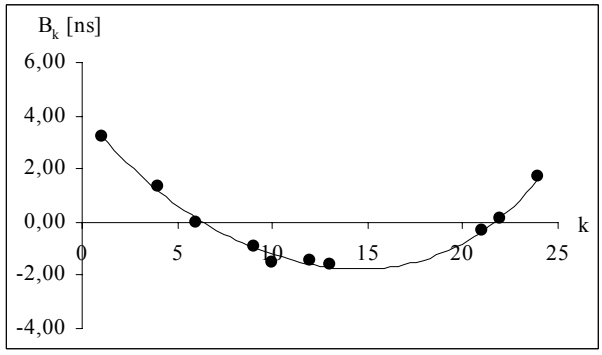
Lab.	Local receiver ser. No.	Portable receiver ser. No.	Period	Number of common views
BIPM	0017 (BIPM A)	0019 (ROVER)	11/09 – 23/10/97	899 (L1)
BIPM	0025 (BIPM D)	0017 (BIPM A)	31/08 – 02/10/98	739 (L1) 698 (L2)
AOS	0030 (AOS)	0017 (BIPM A)	11/01 – 24/01/99	922 (L1)
ORB	0022 (ORB)	0017 (BIPM A)	02/02 – 12/02/99	412 (L1) 398 (L2)
VSL	0018 (VSL 18)	0017 (BIPM A)	15/02 – 27/02/99	807 (L1) 783 (L2)

Table 2. Some calibrations of GLONASS P-code receivers made in the 1990s by the BIPM

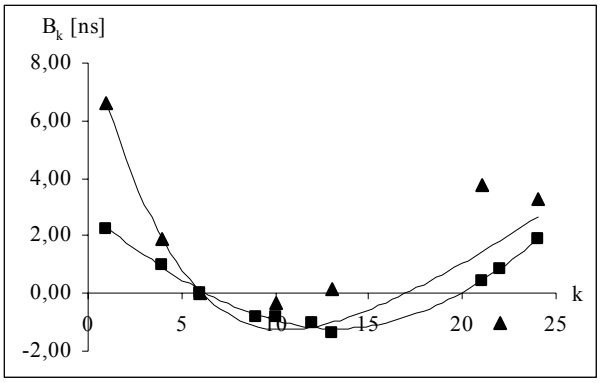
The GLONASS P-code frequency biases are shown in Table 3 and Figure 2. It can be seen that they differ not only for every GLONASS frequency, but also for each receiver.

f	BOR	BIPM D		ORB		VSL	
		L1	L2	L1	L2	L1	L2
1	3.20	2.26	6.60	4.08	0.49	-0.59	-1.30
2	-	-	-	-	-	-0.89	-0.62
4	1.36	0.98	1.90	0.17	-0.14	-0.93	-0.16
6	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	-	-	-	-0.30	-0.24	0.51	0.13
8	-	-	-	0.97	0.45	0.92	0.11
9	-0.91	-0.81	5.84	2.10	1.08	1.60	0.37
10	-1.50	-0.81	-0.33	2.53	1.28	1.25	-0.09
12	-1.47	-1.05	-3.39	3.35	2.21	1.26	-0.78
13	-1.60	-1.36	0.17	3.72	2.21	1.00	-1.07
21	-0.33	0.42	3.77	4.02	0.68	3.61	-3.76
22	0.12	0.87	-1.06	3.76	0.69	3.67	-3.55
24	1.69	1.86	3.30	5.24	0.99	3.10	-4.25

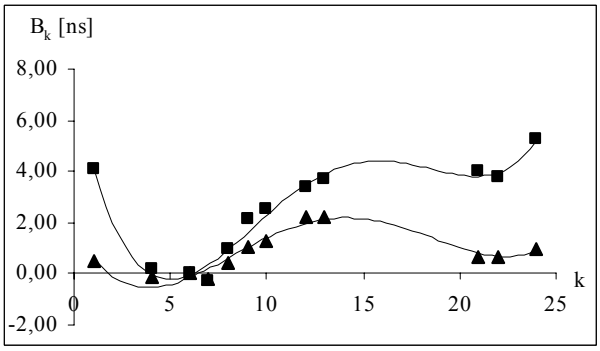
Table 3. GLONASS P-code frequency biases in ns with respect to frequency No. 6.



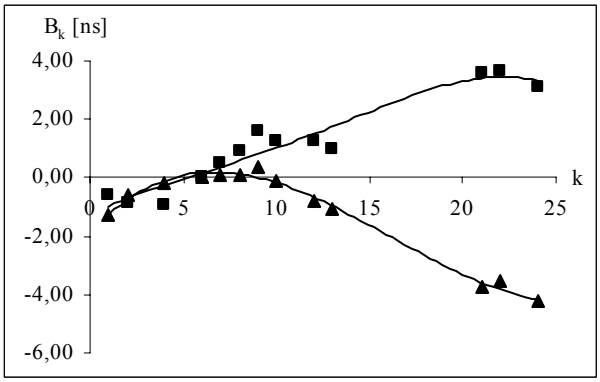
a) receiver No 0030 (AOS) L1



b) receiver No 0025 (BIPM D) ■ L1, ▲ L2



c) receiver No 0022 (ORB) ■ L1, ▲ L2



d) receiver No 0018 (VSL) ■ L1, ▲ L2

Figure 2. GLONASS P-code frequency biases with respect to frequency No. 6

Figure 3 shows an example of one-site comparisons of two receivers. The GLONASS C/A code multi-channel data are noisier then the GPS data. However, the GLONASS P-code single-channel comparison with frequency biases removed shows outstanding performance. Corresponding time deviations are shown in Figure 4.

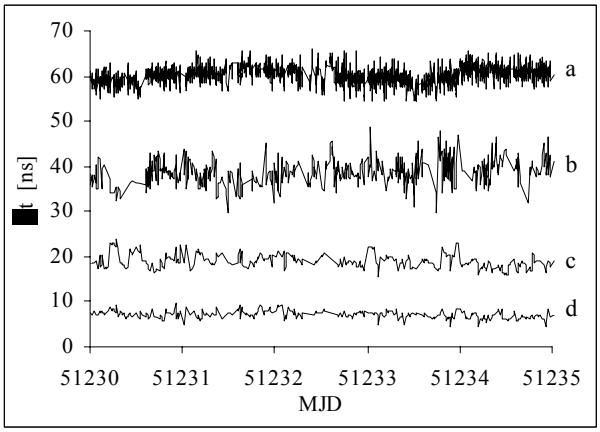


Figure 3. One-site comparisons at the VSL
- a. GPS C/A-code multi-channel;
- b. GLONASS C/A-code multi-channel;
- c. GLONASS P-code single-channel;
- d. GLONASS P-code single-channel with frequency biases removed.

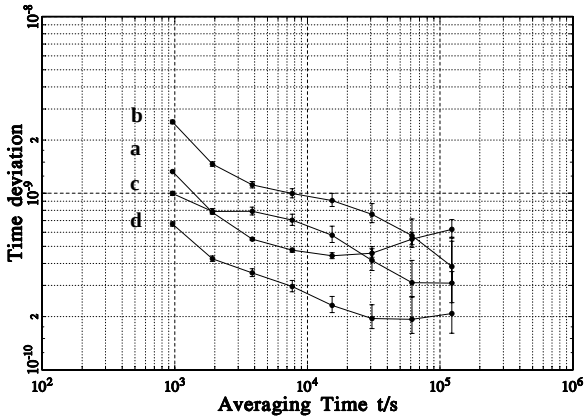


Figure 4. Time deviations for one-site comparisons:
a. GPS C/A-code, b. GLONASS C/A-code,
c. GLONASS P-code (L1) without corrections for GLONASS frequency biases, d. GLONASS P-code (L1) with corrections for GLONASS frequency biases.

4 Application of calibrated GLONASS frequency biases

We applied the frequency biases obtained from the data from January and February 1999 to the time transfer between AOS and NMi- VSL made in February 1999. Using the measurements with broadcast ephemerides and modeled ionospheric delays, the one-day root-mean-square (RMS) of GLONASS P-code common view after corrections for frequency biases is about 3.7 ns, which is the same as for GPS.

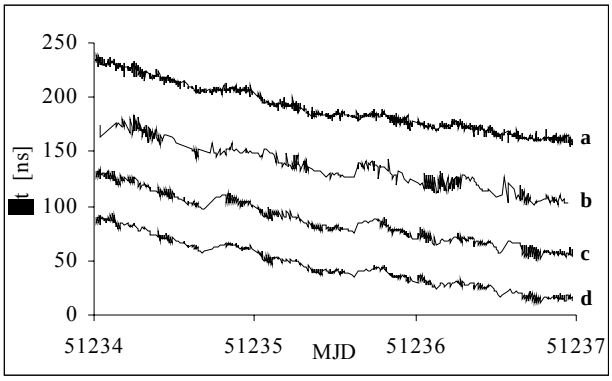


Figure 5. UTC(AOS)-UTC(VSL) compared by:
a. GPS C/A-code, b. GLONASS C/A-code,
c. GLONASS P-code (L1) without corrections for GLONASS frequency biases, d. GLONASS P-code (L1) with corrections for GLONASS frequency biases (with broadcast ephemerides and modelled ionospheric delay).

5 Conclusions

The use of GLONASS P-code promises further improvement in the stability of time transfer. Therefore it is very important to repeat calibrations of GLONASS receivers in order to determine the frequency biases of those currently in operation.

After corrections for ionospheric delay, derived from IGS maps or P-code double-frequency measurements, and for precise ephemerides, clock-comparison uncertainties of less than 1 ns can be obtained [3, 4].

A new generation of GLONASS time receivers will allow in the near future an operational use of GLONASS constellation for international timekeeping.

REFERENCES

[1] W. Lewandowski, J. Azoubib, A.G. Gevorkyan, P.P. Bogdanov, J. Danaher, G. de Jong, J. Hahn, 1996, *First Results from GLONASS Common-View Time Comparisons Realized According to the BIPM International Schedule*, Proc. 28th PTTI, 357-366.

[2] W. Lewandowski, J. Azoubib, G. de Jong, J. Nawrocki, 1997, J. Danaher, *A New Approach to International Satellite Time and Frequency Comparisons*: *'All-in-View' Multi-Channel GPS + GLONASS Observations*, Proc. ION-GPS, 1085-1091.

[3] J. Azoubib and W. Lewandowski, 2000, *Test of GLONASS precise-code time transfer,* Metrologia, **37**, 55-59.

[4] W. Lewandowski, J. Nawrocki, J. Azoubib, 2001, *First use of IGEX precise ephemerides for intercontinental GLONASS P-code time transfer*, Journal of Geodesy, **75**, 620-625.

[5] J. Azoubib, G. de Jong, W. Lewandowski, 1997, *Differential time corrections for multi-channel GPS and GLONASS time equipment located at 3S Navigation, BIPM and VSL Part 1*, Rapport BIPM-97/6.

[6] A. Foks, W. Lewandowski, J. Nawrocki, "Latest calibration of GLONASS P-code time receivers", *Proc. 36th PTTI,* 2004, Washington DC (in press).

NCXO : NEW DEVELOPMENTS FOR AVIONIC APPLICATIONS

Philippe VUILLEMIN, Philippe de JOFFREY, Bernard WOLCOFF*

AR Electronique - Espace Valentin Nord – B.P. 3113 – F 25047 BESANCON CEDEX
* AR Electronique - 15 Rue d'Estienne d'Orves – F 78500 SARTROUVILLE
Tél. : (00.33.3) 81.88.66.77 – Fax : (00.33.3) 81.88.33.55 Email : are@ar-e.com

1- Abstract

Recent progress in electronic technology (integration, low power, low cost…) has allowed the design of TCXOs (Temperature Compensated Xtal Oscillator) with different frequency vs temperature digital compensation systems. These new developments allow to achieve very high frequency vs temperature stability results under hardened environmental conditions, compatible with industrial manufacturing processes and competitive costs. Associated with the design of very low g sensitivity resonators, we introduce NCXOs (Digital Temperature Compensated Xtal Oscillators) devoted to military airborne application.
We present here the classical digital method of frequency vs temperature compensation with DAC (Digital to Analogue Converter) and uC (Microcontroller), as well as a new approach based on the use of a DP (Digital Potentiometer) and a uC.
We compare practical results in the frequency range 50 to 60 MHz with these two designs against basic OCXOs performances (temperature stability / warm-up, consumption / phase noise (inert and under vibrations)).

2- State of the art

With analogue TCXO, we achieve typical performances of frequency stability in extended temperature range between +/-0,5ppm and +/-1ppm. Their consumption is about 10mA. If the equipment needs better frequency stability vs temperature, we generally use a miniature OCXO. It offers better frequency stability in same extended temperature range, typically between +/-.005ppm to +/-0.1ppm, but the power consumption is twenty time higher in steady state at ambient temperature.
The design of new requirements such a radiocom or navigation systems requires increased level of performance for both stability and low consumption.
As an answer to these requirements, we think that a devoted design of digital TCXO – so called NCXO – can be an interesting solution, offering typical stability of +/-0,1ppm over wide temperature range, low consumption and quite instantaneous warm-up. The specific harsh environmental conditions (vibration or shock) are taken into account by the use of hard mounted, low g sensitivity crystals, one of our state of the art internal know how.

3- General principle

On TCXO, the resonator follows the temperature fluctuations, and the induced frequency deviations are corrected by an action on oscillator's frequency. The system is an open loop process.

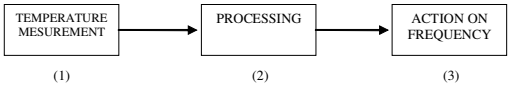


Fig.1 : principle of temperature compensation

The principle is to apply a frequency correction opposite to the free frequency vs temperature deviation of the crystal resonator. By superposition of these two effects we obtain a stable frequency vs temperature depending of the quality of the compensation system.
Fig.2 illustrates in different scales the effects generated by temperature fluctuations.

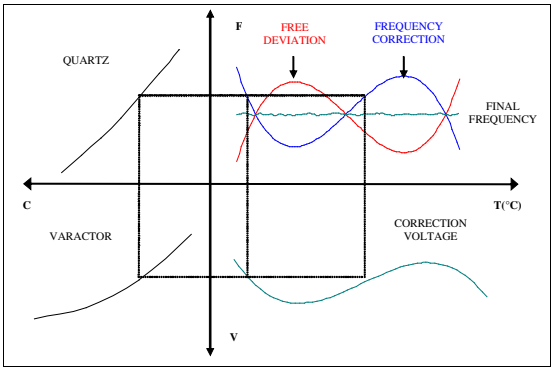


Fig.2

The compensation is obtained either by analogue or digital system.
The analogue compensation is generally made by using a thermistor network or analogue cells multiplier. In the two cases, we build a voltage vs temperature law applied on a varactor. It generates the necessary deviation on frequency to compensate the natural frequency deviation of the crystal resonator.
The digital compensation is achieved by building a voltage correction with numerical circuits.
We now present the two TCXOs that we have studied with two types of digital compensation.

4- NCXO principle

4- 1-Digital compensation with DAC

In order to detect temperature fluctuations, we use a thermal sensor which integrates the analogue sensor with AD converter. The processing is assured by a microcontroller which contains in its memory the digital data corrections. These data are used in a sophisticated algorithm. The digital corrections are then applied on a DA converter connected to the varactor.
Fig.3 shows the DAC NCXO version.

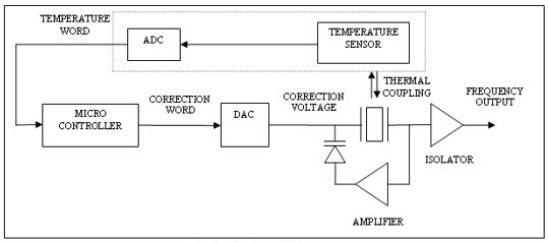


Fig.3 : DAC version

4-2- Digital correction with DP (Digital Potentiometer)

The principle is identical, we just use in this **original method a Digital Potentiometer**. This digital potentiometer is configurated in voltage divider for generating the necessary correction voltage applied on the varactor.
Fig.4 represents the NCXO DP version.

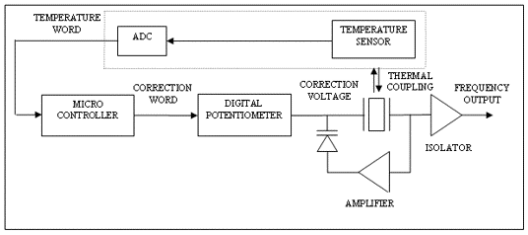
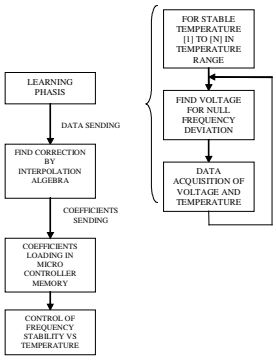


Fig.4 : DP version

Fig.5 gives the different manufacturing flow chart of the NCXO. They are the same for both NCXOs versions.



5- Characterization

5-1- Frequency stability vs temperature results

Fig.6, 7 show frequency stability in [-40, +85]°C for the two NCXOs versions. Measurements have been made continuously at slopes of 0,5°C/min and 1°C/min. The obtained results are +/-0,1ppm. These performances are confirmed on a sample batch of 5 pieces of each NCXOs version, showing a tight reproductibility.

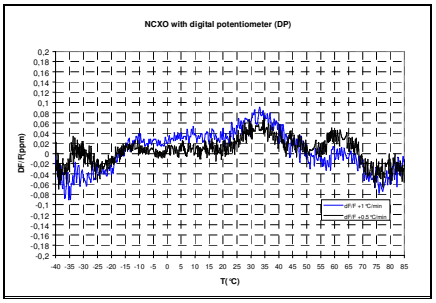


Fig.6 : DP version

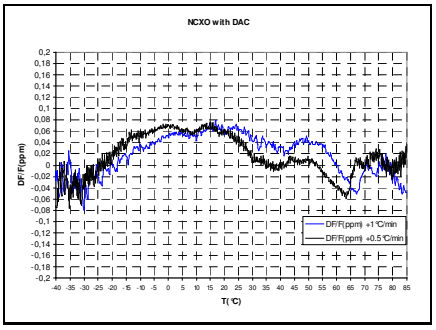


Fig.7 : DAC version

Fig.8, 9 concern the warm-up measurements made at : -40°C, +25°C, +85°C. The results are very interesting : the frequency stability of +/-0,1ppm is achieved quite instantaneously after power over the whole temperature range.

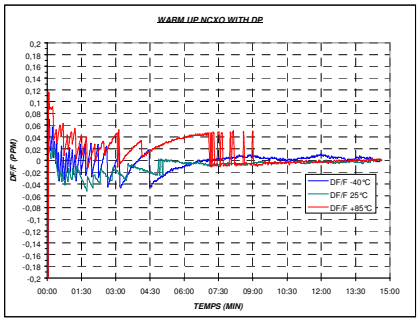


Fig.8 : DP version

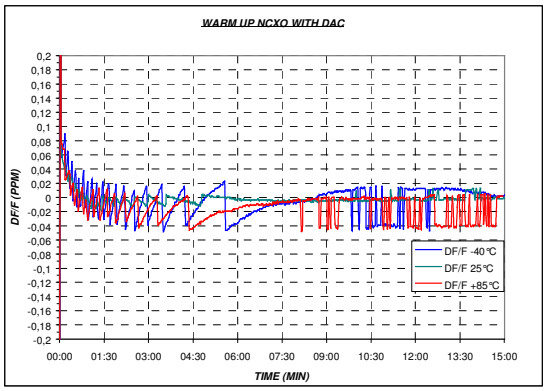


Fig.9 : DAC version

5-2- Dynamic response to fast temperature step

We have done two interesting measurements which consist on applying a step of +65°C (from ambient temperature to +85°C) in 5 minutes and the same measurement for temperature step of -40°C (from ambient temperature to -40°C) and observing the frequency response.
The results show that the frequency stability stays in less than +/- 0,2ppm (see fig.10) .This shows the excellent thermal coupling between the temperature sensor and the crystal resonator.

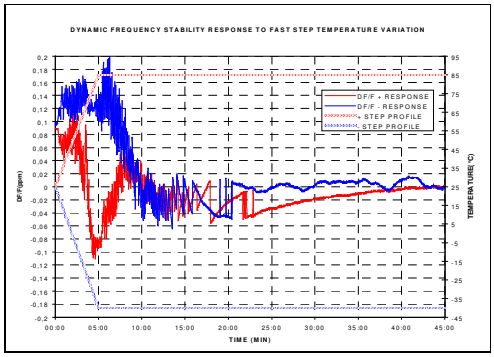
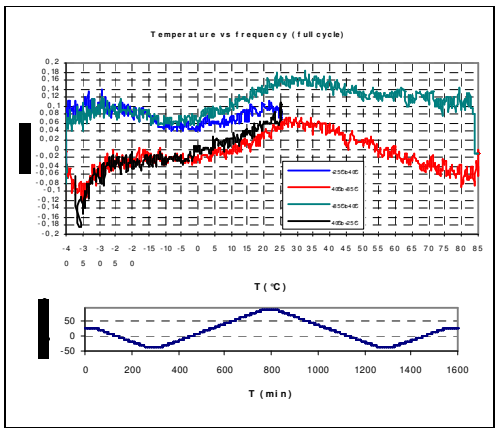


Fig.10

5-3- Hysteresis measurement

Fig.11 illustrates the NCXO frequency behaviour under complete temperature cycle (-25/-40/+85/-40/+25°C). The overall frequency hysteresis is lower than 2.10-7.



5-4- Phase noise measurements

Fig.12 & 13 show the phase noise measurements made on the two NCXOs versions.
For each version, we made a measurement in inert condition and 3 measurements under vibrations (X, Y, Z axes), random excitation profile MIL STD202, meth.214 cond. I, test cond. A. This allows to determine the g sensitivity of the oscillator in the three axes. The results are very interesting, thanks to the quality and the internal know-how around the manufacturing and development of crystal resonators that we have : we obtain g sensitivity better than 2.10^{-10}/g over each axis.
The phase noise measurements have been made at ambient temperature **with steady state of temperature compensation**.

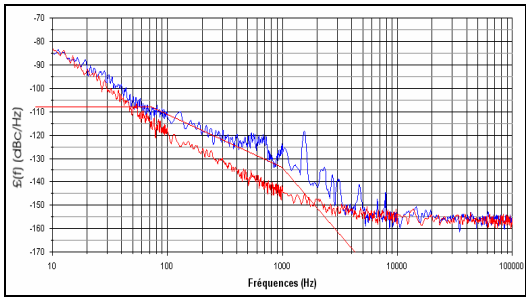


Fig.12 : DAC version

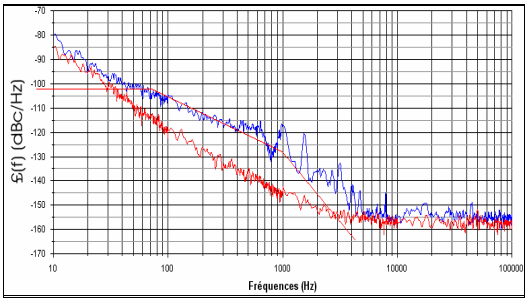


Fig.13 : DP version

5-5- Short term stability

In order to approach as much as possible the real practical environmental conditions, the short term stability measurements have been performed over the whole temperature range, with a 1°C/min slope.
The Allan variance has been characterized for integration times from 100ms to 10s.
Both miniature OCXO, 1.10-7 stability range, analogic TCXO, 1.10-6 class and NCXO, 1.10-7 class have been tested (see chart 1).
The results obtained for both type of NCXO are worst than OCXO and analogue TCXO for integration time < 1s. This is a consequence of the finite step correction. For integrated time of 10s, we achieve similar results than with analogue TCXO.

CHARACTERISTICS	*NCXO (WITH DAC)*	*NCXO (WITH DP)*	*OCXO*	*ANALOG TCXO (thermistor network)*
FREQUENCY (MHz)	56	56	56	56
CRYSTAL RESONATOR	AT TO5	AT TO5	SC TO8	AT TO5
OSCILLATOR TYPE	CLAPP	CLAPP	CLAPP	CLAPP
OUTPUT TYPE	SINE WAVE	SINE WAVE	SINE WAVE	SINE WAVE
PACKAGE (mm)	27 X 36 X 10	27 X 36 X 10	27 X 36 X 10	27 X 36 X 10
SUPPLY VOLTAGE (V)	5	5	5	12
CONSUMPTION (mA) AT 25°C	**8.5**	**8.5**	180 (steady state)	8.5
ΔF/F (-40 to 85°C)	**$\pm1.10^{-7}$**	**$\pm1.10^{-7}$**	$\pm1.10^{-8}$ to $\pm1.10^{-7}$	$\pm1.10^{-6}$
SUPPLY SENSITIVITY (±5%)	$\pm1.5.10^{-8}$	$\pm1.5.10^{-8}$	$\pm5.10^{-9}$	$\pm1.5.10^{-8}$
LOAD SENSITIVITY (50Ω ±5%)	$\pm1.5.10^{-8}$	$\pm1.5.10^{-8}$	$\pm1.10^{-9}$	$\pm1.5.10^{-8}$
AGEING 1 YEAR	**5.10^{-8} to 1.10^{-7}**		5.10^{-8} to 2.10^{-7}	5.10-8
OVERALL STABILITY OVER 10 YEARS	**5.10^{-7} to 8.10^{-7}**		5.10^{-7} to 1.10^{-6}	1,5.10-6
WARM UP	**10^{-7} AFTER 3s IN WHOLE TEMPERATURE RANGE**	**10^{-7} AFTER 3s IN WHOLE TEMPERATURE RANGE**	-40°C 2.7W 2.10^{-7} AFTER 4mn 25°C 2.5W 2.10^{-7} AFTER 1mn	-40°C 2.10-7 AFTER 3mn30 25°C 2.10-.7 AFTER 1mn30
£(F) (dBc/Hz) 10Hz 100Hz 1kHz NOISE FLOOR	-85 -120 -145 -155	-85 -120 -145 -155	-100 -130 -150 -160	-90 -120 -145 -155
ALLAN VARIANCE [-40 85°C] (dynamic conditions (1)) 0.1s 1s 10s	$4.10^{-10}<\sigma y(\tau)<7.10^{-9}$ $2.10^{-9}<\sigma y(\tau)<1.5.10^{-8}$ $3.10^{-10}<\sigma y(\tau)<2.10^{-8}$	$4.10^{-10}<\sigma y(\tau)<7.10^{-9}$ $2.10^{-9}<\sigma y(\tau)<1.5.10^{-8}$ $3.10^{-10}<\sigma y(\tau)<2.10^{-8}$	$6.10^{-12}<\sigma y(\tau)<8.10^{-12}$ $1.5.10^{-11}<\sigma y(\tau)<3.10^{-11}$ $1.5.10^{-10}<\sigma y(\tau)<4.10^{-10}$	$3.10^{-11}<\sigma y(\tau)<8.10^{-11}$ $4.10^{-11}<\sigma y(\tau)<4.10^{-10}$ $8.10^{-9}<\sigma y(\tau)<2.10^{-8}$

Note (1) : short term stability measurements are performed over [-40, +85]°C temperature range, slope +10°C/min.

Chart 1 : comparative measurements

Conclusion

We have designed two types of digital TCXOs. These two NCXOs versions present homogeneous and identical performances with good reproductibility.
They can offer quite similar frequency versus temperature stability levels as for miniature OCXOs (typically 0,1ppm) with significative advantages in terms of power consumption and warm-up time.
Additionally, due to their own design, very low aging rates can be achieved, as they avoid the need of permanently high ambient temperature as for OCXO design. This aspect has also a positive impact for long term reliability.

Perspective

Further development are in progress for both size and consumption reduction.
Additionally, complementary developments will be done for other professional applications, with higher stability requirements under more standard environmental conditions.

Glossary

TCXO : Temperature Compensated Xtal oscillator
NCXO : Numerical temperature Compensated Xtal oscillator
ADC : Analogue to Digital Converter
DAC : Digital to Analogue Converter
DP : Digital Potentiometer
UC : Microcontroller

NEW HIGH STABILITY 10 MHz OCXO FAMILY WITH UPGRADED SIZE VS PERFORMANCE RATIO

H. CHARTOGNE, P. RENOULT, E.GIRARDET, B. WOLCOFF*

AR Electronique - Espace Valentin Nord – B.P. 3113 – F 25047 BESANCON CEDEX
* AR Electronique - 15 Rue d'Estienne d'Orves – F 78500 SARTROUVILLE
Tél. : (00.33.3) 81.88.66.77 – Fax : (00.33.3) 81.88.33.55 Email : are@ar-e.com

1- Abstract

During the last decades, several generations of high stability OCXOs have been introduced on the market with constant progress mainly about the stability versus size ratio. This was first initialized by the use of SC cut crystals and followed by the electronic miniaturization (SMD components, asics).

AR Electronique has participated to this tendency with the PULSAR (50x40x20mm size), JUMBOSTAR (27x36x16mm size) and micrOSTAR series (26x26x12mm size). All these families offer top of the art level of both phase noise and short term stability in their class.
As an answer to the always more demanding requirements for both highest performances together with miniaturization, AR Electronique has recently developed a further generation of miniature 10 MHz OCXO, so called ***nanOSTAR***, which offers a new level of stability and phase noise versus size ratio. Frequency stability vs temperature of 5.10-9, aging <5.10-9/month, short term stability <8.10-13 (1s), phase noise <-108 dBc/Hz at 1 Hz, <-138 dBc/Hz at 10 Hz can be proposed on the ***nanOSTAR*** family, housed in a miniaturized case type 20x20x13mm.

This paper describes shortly the technological aspects of this project (crystal size miniaturization and optimization, thermal structure, low noise oscillator's electronic) together with electrical test results performed on the first pre-serial production batches, including phase noise, frequency stability versus environmental conditions, Allan Variance.

In the compact 10 MHz OCXO landscape, the ***nanOSTAR*** family will offer a real top level compromise between very high performance and miniaturization;
Further developments will be investigated, especially concerning hardened versions, low g sensitivity, devoted to avionic military applications.

2- High performance 10 MHz miniature OCXO's : new requirements

Recent developments of systems such as synchronisation networks, GPS receiver, instrumentation devices require the use of oscillators, providing high frequency stability, very low phase noise within a compact size.

Additional characteristics such as fast warm-up time, low retrace, as well as low consumption, low supply voltage can be necessary.

These requirements can be grouped together as follows:

[1] Frequency stability

Vs temperature Fcy slope/°C	5.10-9 ÷ 1.10-9 1.10-10 ÷ 0,5.10-10/°C
Aging 1 month 1 year over 10 years	5.10-9 ÷ 1.10-9 3.10-8 ÷ 1.10-8 3.10-7 ÷ 1.10-7
Short term 100ms to 10s (Allan variance)	1.10-11 ÷ 1.10-12
Fcy warm-up 5.10-8 Retrace (24h OFF, Ton+2h)	10 min ÷ 2 min 5.10-9 ÷ 1.10-9

[2] Phase noise

1 Hz	-100 ÷ -110 dBc/Hz
10 Hz	-130 ÷ -140 dBc/Hz
100 Hz	-145 ÷ -150 dBc/Hz
1 KHz	-150 ÷ -155 dBc/Hz

[3] Size / package

Dimensions	22 x 25 x 15 20 x 20 x 15
SMD	YES (option)
Supply voltage	12V ÷ 3V3 (option)

In order to summarize, both requirements [1] & [2], currently available on larger oscillator's families, are expected within sizes [3], typically 2 times to 4 times smaller than the previous ones.



Fig.1 : PULSAR, JUMBOSTAR, micrOSTAR, nanOSTAR families

3- Design and development

The directive idea for the design of the ***nanOSTAR*** family oscillator was to find the best compromise between highest characteristics level and size.
Both crystal resonator, oscillator and thermal structure design have been taken into account.

Crystal resonator
The current AR Electronique 10 MHz SC cut, 3rd overtone crystal, housed in a HC-53/U case is ideally suited for both frequency stability and spectral purity but definitively not compatible regarding the size.
A completely new resonator design has been launched with the following characteristics :
- enclosure type HC-49/U
- SC cut, 3rd overtone blank, diameter ~8mm, plano convex geometry
- typical motional parameters : L=1,2H, Q>1,2.10-6
Several manufacturing batches (typical size 80 to 100 units) have been tested on conventional oscillators, mainly regarding aging and phase noise characteristics (see below).

Electronic design
The main aspect has concerned the size miniaturization of both oscillator circuit (clapp version) and voltage regulator. Due to the very low phase noise targets, the regulated voltage filtering has been optimized. Additionally, a low supply voltage option (down to +3,3V) has been taken into account.

Oven structure
Despite the very small available volum, both oscillator, voltage regulator as well as oven control circuit are ovenized. The single oven design procures a very tight crystal temperature regulation (typically +/-0,1°C for 100°C external temperature variation), together with fast warm-up and very stable answer to dynamic ambient variations (temperature steps).
The typical oven temperature is +75 to +80°C. Reliability and experimental tests have been performed with higher value (up to +110°C).

Mechanical assembly / technology
The monolithic module, including crystal, oven, oscillator and oven control PC boards is housed in a 20 x 20 x 12,7 mm case, hermetically sealed.
A SMD version, including bended pins can be proposed. This design has been preferred as compared with 22 x 25 mm SMD case with PCB substrat, for both reliability and thermal discoupling aspects.
Designed for RoHS directive further compliance, the automatic assembly of such very high stability oscillator under high temperature reflow process (lead free requirement) is however not recommended.

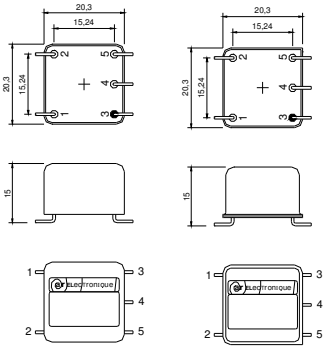


Fig.2 : nanOSTAR packages (2 versions)

4- Experimental results

All electrical characteristics have been characterized on several prototypes and pre-serial manufacturing batches. The most critical ones are illustrated below.

Aging behaviour
Each batch of HC-49U 10 MHz crystal has been fully tested in long term stability over significative periods (1 to 3 months, continuous working (see example fig.3). The overall results are very promising, as shown in fig. 4, 5. Average aging slopes of 1 to 2.10-9/month have been measured with a very tight distribution. The behaviour of the whole batches is very important as it qualifies all crystal manufacturing process efficiency.

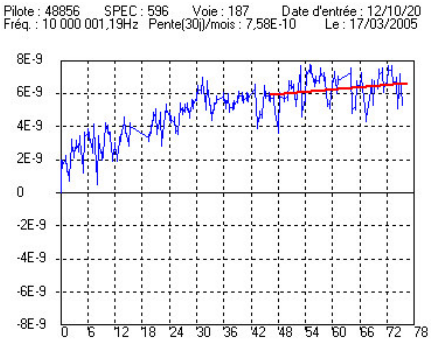


Fig. 3 : typical aging curve

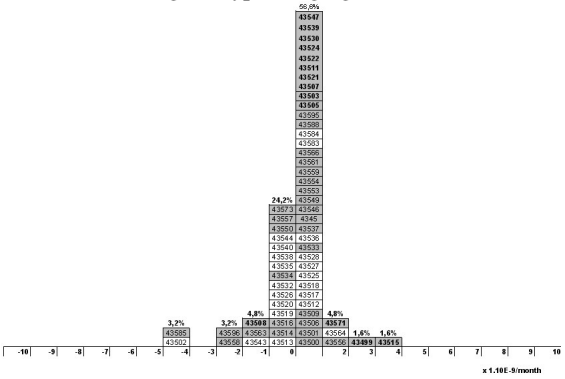


Fig.4 : aging distribution, batch n° 2 (62p)
(10-9/month), 45 ÷ 75 days first turn-on

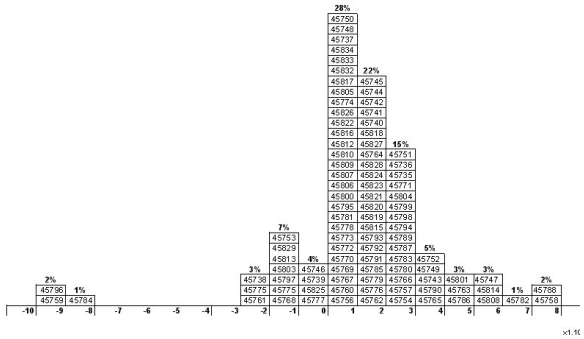


Fig.5 : aging distribution, batch n°3 (100p)
(10-9/month), 30 ÷ 45 days first turn-on

Frequency stability vs temperature

Both crystal resonator, oscillator and regulated voltage characteristics, as well as oven design and thermal control are involved in the oscillator's behaviour under temperature changes. The very compact structure allows to reach stabilities of a few 10-9 over [-20, +70]°C, with typical slope < 5.10-11/°C (see fig.6).

The dynamic behaviour can be a very critical parameter for some applications. A typical example is shown on fig.7, as the frequency response to instantaneous temperature steps +25/0/+70/+25°C.

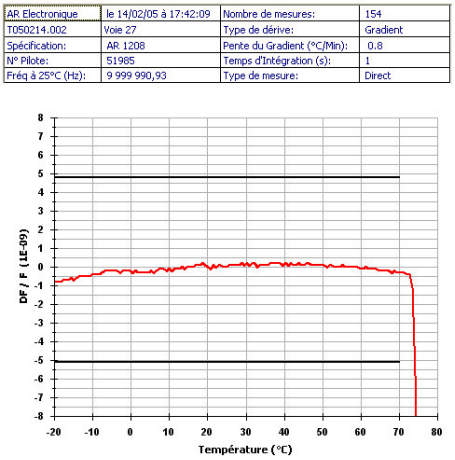


Fig.6 : frequency vs temperature
([-20, +70]°C, slope +0,8°C/min)

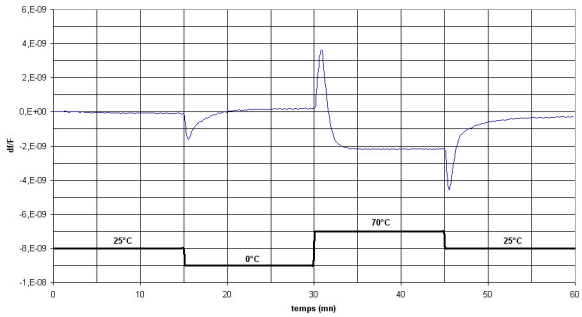


Fig.7 : frequency vs temperature steps

Warm-up

The frequency stabilization rate after turn-on, together with the power consumption have been tested on significative batches, over the temperature range. The results are summarized below. Curve 8 gives a typical example at ambient temperature.

Temperature [1]	0°C	+25°C	+50°C
Frequency [2]			
+/-5.10-8	120 ÷ 150s	90 ÷ 100s	60 ÷ 75s
+/-1.10-8	150 ÷ 180s	120 ÷ 135s	75 ÷ 90s
Consumption			
Warm-up	2,8W / 120s	2,8W / 90s	2,8W / 40s
Steady state [1]	1 ÷ 1,2W	0,7 ÷ 0,8W	0,5 ÷ 0,6W

Notes :

[1] : ambient temperature, not ventilated.

[2] : referred to stabilized frequency, 15 min after power on.

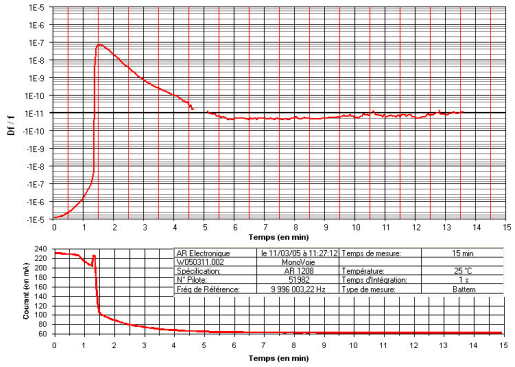


Fig.8 : warm-up (+25°C, frequency, current (+12V))

Frequency stability vs supply voltage

The oscillator's design has been done so that the frequency vs supply voltage sensitivity is typically 10 times lower than the temperature effect.

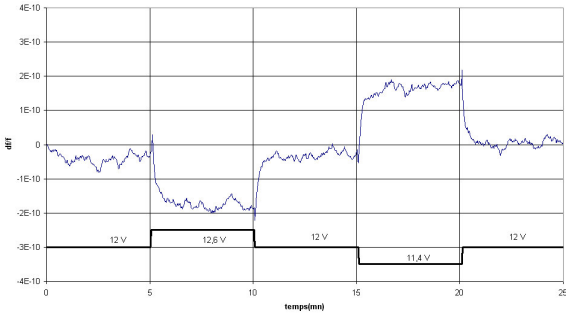


Fig.9 : frequency vs supply voltage

Time error vs elapsed time

A wide range of synchronisation applications requires the use of a 10 MHz, high stability source. The OCXO is phase locked on a primary source or GPS receiver. In case of PLL defect, the maintain of synchronized system is depending on the free run oscillator's stability, correlated to external environmental conditions.

The next table shows the typical behaviour of a ***nanOSTAR*** oscillator : a time error <10µs can be guaranteed after a period of 6 ÷ 10 hours, taking into account the following parameters :

Frequency offset when PLL defect appears	3.10-10 (3mV steps)	5 µs
Temperature effect (10°C) Supply voltage (1%)	5.10-10 1.10-10	1 µs
Aging (6 hours) (continuous working)	1.10-10	0,2 µs
Accumulated time error Over 6 hours		6,2 µs

Phase noise – Short term stability

These 2 criterions are the key points of a lot of new demanding applications. The oscillator's characterization versus time and frequency domains is tightly connected. For integration time lower than 1s, the short term stability, defined as Allan variance, is directly correlated to the phase noise characteristic. Above 1 to 10s, the thermal behaviour, the stabilization time and furthermore the aging are the main effects for short and medium term stability.

The following curves show typical experimental measurements obtained with the ***nanOSTAR*** OCXO, in both frequency and time domain.

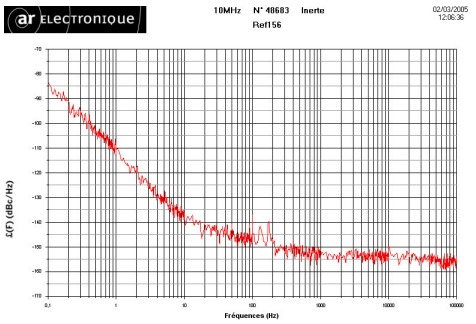


Fig.10 : phase noise (0,1 Hz to 100 KHz)

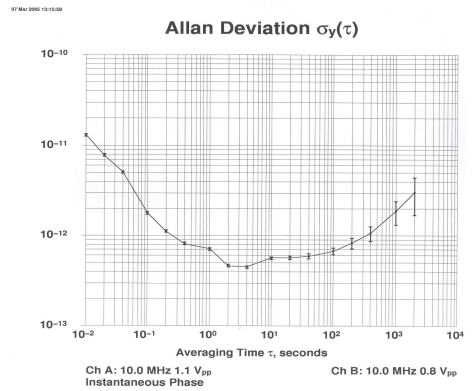


Fig.11. short term stability
(Allan variance, 10ms to 1000s)

5- Results analysis – further developments

With the 10 MHz OCXO ***nanOSTAR*** family, AR Electronique is able to offer a new level of stability and phase noise characteristics versus size.

Thanks to the new designed 10 MHz SC cut crystal, as well as oscillator and oven miniaturized functions, the previous top level characteristics of JUMBOSTAR (27 x 36 x 16) and micrOSTAR (26 x 26 x 15 mm) can be offered within a very compact 20 x 20 x 15 mm size :

- stability class 10-9 (temperature, aging/month)
- short term 1s to 100s < 1.10-12 (8 ÷ 5.10-13 typ.)
- phase noise < -105 dBc/Hz at 1 Hz (-108 typ.)
  < -135 dBc/Hz at 10 Hz (-138 typ.)

These overall performances, especially short term stability and phase noise, are not so far from U.S.O. characteristics, typically 25 to 50 times larger size.
This can be a real interest for new demanding systems where both top level stability and size requirements are critical.

Based on this concept, some complementary developments have been launched for compact OCXO's devoted to military airborne applications.
Their designs are centered on the use of miniaturized 10 MHz, SC cut, T0 enclosure, offering a very low g sensitivity (1 to 2. 10-10/g, worst axis).
The thermal concept is optimized for very fast warm-up time as well as extended temperature range. The mechanical structure is of course adapted to the very severe mechanical vibrations and shock requirements.

References

R.L. FILLER and J.R. VIG
The acceleration sensitivity of quartz crystal oscillators
Review proc.41st ASFC (1987)

J. RUTMAN and J. GROSLAMBERT
Characterization and measurement of frequency stability
EFTF (1990)

J. GROSLAMBERT, M. OLIVIER and J. UBERSFELD
Spectral and short term stability measurements
CPEM dig. (1974)

THE FRENCH TIME REFERENCE UTC(OP), AND THE CONNECTION OF THE EGNOS NETWORK TIME

Pierre UHRICH, David VALAT,
SYRTE, CNRS UMR 8630, Observatoire de Paris, 61 avenue de l'Observatoire, 75014 Paris, France.
Pierre.Uhrich@obspm.fr, David.Valat@obspm.fr

Michel BRUNET, Jean MARECHAL, Thierry BELTAN, Norbert SUARD,
CNES, 18 avenue Edouard-Belin, 31401 Toulouse Cedex 4, France.
Michel.Brunet@cnes.fr, Jean.Marechal@cnes.fr, Thierry.Beltan@cnes.fr, Norbert.Suard@cnes.fr

ABSTRACT

This paper describes in a first part the generation of UTC(OP), as realized in Observatoire de Paris (OP). The connection to UTC through the years is provided, including prediction issues for the steering of UTC(OP). A qualitative comparison with other UTC(k) is proposed. Future developments of an ensemble time scale are introduced.
The European Satellite Based Augmentation System EGNOS (for "European Geostationary Navigation Overlay Service") will provide to the users in the navigation message the time difference and the frequency between EGNOS Network Time (ENT) and UTC. For this aim, one EGNOS Earth station was installed in OP. In a second part, the paper details all the time aspects of EGNOS and describes the connection of ENT to UTC(OP), in order to insure the traceability of ENT to UTC.

1. INTRODUCTION

The international time reference UTC (Coordinated Universal Time) is by construction only available in deferred time and it is not a physical signal. Hence, for their needs, the National Measurement Institutes have to generate a real time physical realization of UTC, called UTC(k), k being the laboratory acronym. This is one of the basic missions of the SYRTE in Observatoire de Paris (OP). The paper describes in a first part the generation of UTC(OP), currently based on a single commercial Caesium standard selected for his very good long term stability. The relationship to UTC through the years is discussed, leading to prediction issues for the steering of UTC(OP). A qualitative comparison with other UTC(k) is proposed, and the current developments for a future UTC(OP) based on Hydrogen Masers and an ensemble of Caesium clocks are introduced.
The European Satellite Based Augmentation System, called EGNOS (European Geostationary Navigation Overlay Service), will provide to the users in Europe an augmentation of three pseudo-GPS signals plus the integrity information about the available GPS and GLONASS constellations [1, 2, 3]. These missions lead to some time requirements, described in a second part of the paper. EGNOS will also broadcast in its navigation message the time difference and the frequency between EGNOS Network Time (ENT) and UTC. This requires to connect ENT to at least one UTC(k) signal in Europe. For this aim, an EGNOS Earth station of the Ground Segment, so-called Ranging and Integrity Monitoring Station (RIMS), was installed in OP in July 2003. This connection is also described in detail.

2. UTC(OP)

2.1. Current UTC(OP) signal generation.

The UTC(OP) generation is described in the figure 1.

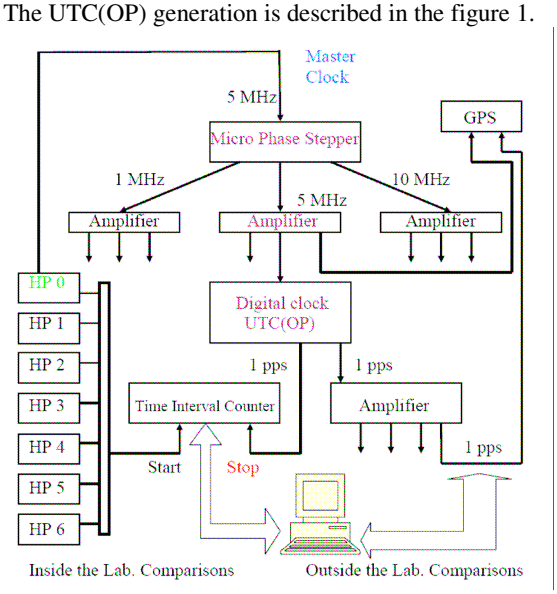


Figure 1. Current UTC(OP) generation.

There are currently seven commercial atomic Caesium standards in the SYRTE at OP, all of the same type: HP 5071 A High Performance Tube. One of the Caesium standard is selected for its very good long term stability. Its 5 MHz signal is sent to the input of a Micro Phase Stepper, which is used when

needed to steer the frequency of the signal by steps of 5 x 10^{-15}. The output signal is distributed to a Digital Clock, which generates a 1 pps (one pulse per second) signal. The physical reference point for UTC(OP), to which all delays are related, is the Stop input of a Time Interval Counter (TIC). Internally, all the clocks, including the one used for the UTC(OP) generation, are measured against this 1 pps through the TIC. Externally, the 1 pps is the reference signal for all the satellite comparisons to the outside, like GPS Common-Views or Two-Way Satellite Time and Frequency Transfer (TWSTFT).

2.2. Relationship between UTC and UTC(OP).

On the figure 2 are plotted the differences UTC-UTC(OP) from October 1995 to December 2004, as given by the Time Section of the BIPM after the computation of UTC with a five day sampling period. Over that period of time, there have been 4 changes of master clock and 43 frequency steering. One can clearly see that, since the beginning of that period of time, UTC(OP) follows:

$$-100 \text{ ns} < \text{UTC} - \text{UTC(OP)} < 100 \text{ ns}$$

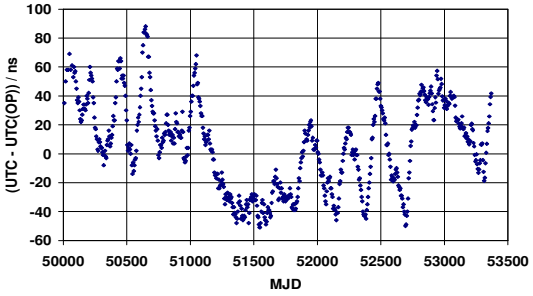


Figure 2. UTC-UTC(OP) from October 1995 to December 2004.

Even if UTC(OP) is not a free running clock, one can plot an Allan standard deviation as is done in the figure 3.

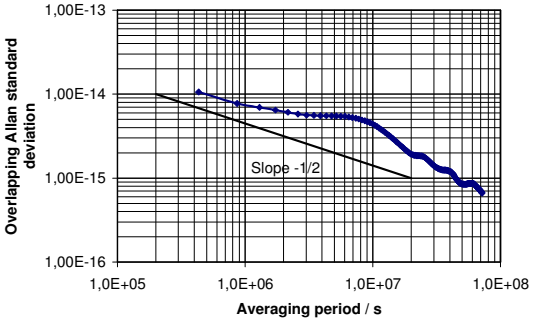


Figure 3. Overlapping Allan Standard Deviation of UTC-UTC(OP) [October 1995 to December 2004].

The trend is a white frequency noise modulation (slope -1/2), with the long term periodic behaviour clearly visible in the figure 2. This leads naturally to consider the frequency samples of UTC-UTC(OP), hence to compute an average frequency as the best frequency estimate. We obtain 2,3 x 10^{-17}, which is close to zero. As it was expected from all the procedures of generation and maintenance of UTC(OP), it shows that the best long term prediction of UTC-UTC(OP) is the last known value.

But the monthly average frequency of UTC-UTC(OP) can reach ± 1,5 x 10^{-14} [4], which might lead to a departure of UTC(OP) from UTC of about ± 38,9 ns over one month. As we try to keep UTC-UTC(OP) inside the ± 50 ns limits, a shorter term steering is required. This is why the current prediction of UTC-UTC(OP) is based on a linear model, with a least square average frequency computed back to the latest frequency steering date. The steering choice and amplitude are then made on the differences between UTC(OP) and other references: typically other daily available UTC(k) or other local OP Caesium clocks. This allows for a better prediction of the next frequency steering.

2.3. Qualitative comparison with other UTC(k).

Figure 4 shows UTC-UTC(k) for some selected laboratories during the year 2004.

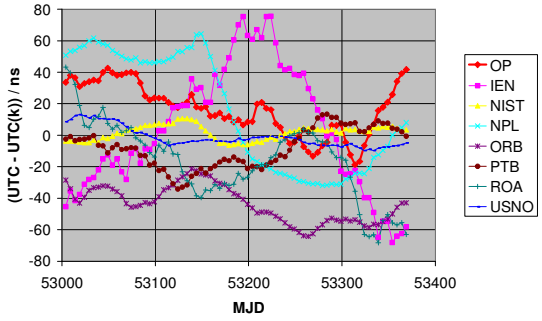


Figure 4. UTC-UTC(k) for some selected laboratories during the year 2004.

These UTC(k) are based on either one single Primary Frequency Standard (PTB, Germany), or a composite clock algorithm (NIST and USNO, United States of America), or one steered H-Maser (NPL, England), or an average of Caesium clocks (ROA, Spain), or one single Caesium clock (IEN, Italy, OP, France, and ORB, Belgium). The limits of these last three UTC(k) are most probably given by the instability of the Caesium standards. To improve the UTC(OP) generation appears to be possible by using like in NIST an ensemble of H-Masers for the short term stability, and an ensemble of Caesium standards for the long term stability.

2.4. Current developments for a future UTC(OP).

A new development was indeed recently undertaken in OP. The basic idea is to use the time scale algorithm AT1 developed by NIST, which is well documented [5], because the type and number of clocks foreseen at OP are similar. In particular, we have implemented the time scale algorithm developed in NPL a few years ago [6], with a specific handling of the in and out of clocks.

On the figure 5 are plotted the relative weights of the contributing clocks over almost 9 years backwards. There have been up to six Caesium standards and up to two H-Masers.

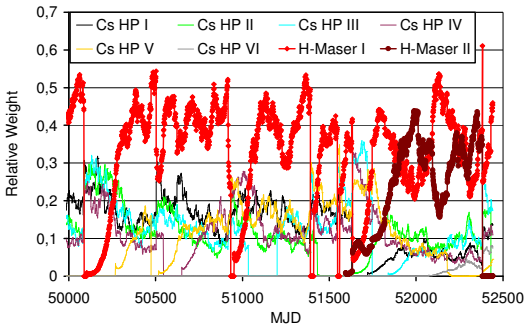


Figure 5. Relative weights of the contributing clocks over 9 years backwards.

The H-Maser relative weights are far above the Caesium clock weights, except when a problem occurs on the H-Masers, inducing a large jump in the other clock weights. Each time, the H-Masers relative weights go back above the Caesium weights after a relatively short period of time. What we now have to study is the influence of an upper limit to the relative weight on the stability of the time scale, in order to minimise the influence of such jumps.

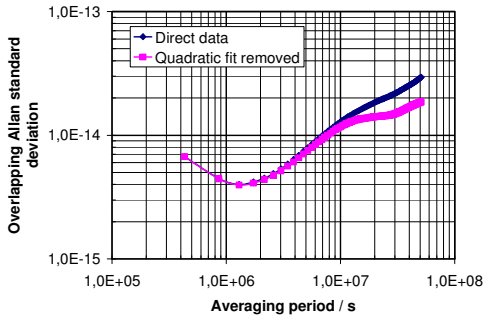


Figure 6. Current stability of the OP time scale under development with respect to UTC over 9 years.

The current stability with respect to UTC is shown on the figure 6. The upper curve is the stability computed directly from the time scale data, while the other curve is the stability when a quadratic fit has been removed. On this last curve, there is an apparent annual periodic term, which might be related to the H-Maser failures or to a seasonal effect. Over 30 d, the time stability of the time scale is about 7 ns, allowing for a potential monthly prediction of UTC-UTC(OP) inside a ± 21 ns limit.

Two new H-Masers have been purchased recently, and the inclusion of these units will improve the short term stability and bring more reliability to the time scale. There is no doubt the current UTC(OP) developments will be immediately useful. On the other hand, the BIPM Time Section is currently evaluating the uncertainties on the UTC-UTC(k).

3. EGNOS

EGNOS, the European Satellite Based Augmentation Systems (SBAS), will provide GPS and GLONASS regional augmentation services in aviation, maritime and land users by using transponder on board geostationary (GEO) satellites [1, 2, 3]. The EGNOS Ground Segment consists of Ranging and Integrity monitoring Stations (RIMS), which are connected to a set of redundant control and processing facilities (CPF), in order to determine the integrity, ephemeris and clock differential corrections for each monitored satellite, to compute the ionospheric delays, and to generate the GEO satellite ephemeris. The GEO downlinks these data on the GPS L1 frequency with a modulation and a coding scheme similar to GPS.

3.1. Time aspects of the EGNOS requirements.

EGNOS being based on differential principles, all measurements and data are referred to an internal EGNOS Network Time (ENT) whose performance requirements were derived exclusively from navigation accuracy performance requirements. Derived from Signal in Space (SIS) specifications, the European Space Agency (ESA) requires ENT to be steered within 50 ns (5σ) of GPS Time. This not very stringent requirement is mostly specified to remain compatible with the maximum capacity of the message used to correct the GPS satellite clocks. It will allow the user to mix in its navigation solution GPS and EGNOS signals. For this “Ranging” mode, ESA has specified an User Equivalent Range Error (UERE) of 25 m (2σ). This UERE encompass:

- the ‘ENT - GPS Time’ offset error;
- the ‘ENT - GEO Time’ transfer error (GEO Time is the equivalent time scale on-board the GEO);
- the ‘GEO - User’ range error.

The ENT - GEO Time transfer error is specified to be less than 10ns (3σ), after offset and frequency corrections provided in EGNOS GEO message #9.

This particular requirement is of special interest for a precise time broadcast function in real time. The table 1 provides the EGNOS time requirements.

Table 1. EGNOS time requirements [ESA].

1	(ENT - GPS time) offset ≤ 50 ns (5σ)
2	(GEO Time - GPS time) offset ≤ 50 ns (5σ)
3	(GEO Time - ENT) accuracy ≤ 10 ns (3σ)
4	(ENT - UTC(OP)) accuracy ≤ 10 ns (3σ)

As required in SIS specification, EGNOS will provide in message # 12 the time difference and the frequency between ENT and UTC. But, in discrepancy with the SIS requirement, the real time offset will not be provided with respect to UTC, which is a deferred-time paper time scale, but with respect to a physical clock signal of a European National Metrology Institute, and by taking into account the prediction of the difference between this UTC(k) and UTC.
EGNOS system presently recommends that ENT shall be connected to UTC as accurately as possible, because the UTC-UTC(k) time difference is not under EGNOS control and not available in real time. According to the performance analysis of UTC(OP) given in the chapter 2 above, the ENT – UTC accuracy will be better than 100 ns (3 σ).

3.2. RIMS clock synchronisation, ENT generation.

RIMS clock synchronisation is performed using the *composite-clock* technique [7]. In the composite-clock technique, ENT is defined as the implicit ensemble mean of all RIMS clocks and the synchronisation process generates estimates of the offset and frequency of each RIMS clock relative to it. These estimates can then be used to reference all RIMS pseudo-range measurements to ENT. This synchronisation process is necessary in order to allow simultaneously observed pseudo-range measurements from multiple RIMS to be combined in the function which estimates satellite clock errors.
The composite clock algorithm is computed by means of a Kalman filter. The filter measurement data comprises a linearly independent set of common-view observations with minimum a priori variances. A common-view observation for a pair of RIMS is formed by subtracting simultaneously observed pseudo-range measurements to a common satellite. Before subtraction, the pseudo-ranges are pre-processed to remove the RIMS antenna to satellite antenna geometric ranges and to reduce unwanted errors, such as multipath delays and thermal noise. The resulting common-view observations represent direct measurements of the RIMS clock offsets plus residual errors [8].
The relative weighing of clocks used within the filter is dependent upon several factors, but is largely determined by the process noise models associated with each of the clocks. These models characterise RIMS clock stability.
Since only clock difference measurements are available to the filter, the filter model has unobservable components that cause secular growth trends in the state error covariance matrices. Special measures must be employed to remove these trends, otherwise they will eventually lead to numerical instability within the filter. It can be shown that, once the secular growth trends due to the unobservable components of the system model have been removed from the covariance matrix, the resulting covariance describes the errors in the filter state estimates relative to the implicit ensemble mean, *not* to the ideal time-scale.

3.3. Steering of ENT to GPS Time.

In order to limit the dynamic range of the satellite clock corrections, thereby reducing the size of the wide area differential (WAD) clock messages and improving the efficiency of the message dissemination process, it is necessary to steer ENT to the GPS Time. Steering of ENT to the GPS Time is performed using a second order, low-pass digital filter. The steering input signal is an instantaneous estimate of the 'ENT - GPS Time' offset. This is computed from the estimated satellite clock offsets from ENT and the GPS broadcast satellite clock corrections, which are estimates of the satellite clock offsets from the GPS Time.
The EGNOS system baseline brings confidence that the 'ENT - GPS Time' steering requirement of < 50 ns will be achieved. Different scenario simulations are consistent with other findings [9].

3.4. Satellite clock corrections.

Satellite clock corrections and correction rates are computed using all available pseudo-range and Doppler measurements from the pre-processing function. The measurements are referenced to ENT using the RIMS synchronisation parameters and then collected into groups by satellite. For each satellite, weighted least-squares estimates of its clock's offset and rate-of-change with respect to ENT are computed. The offsets are later separated by low-pass filtering, similar to that used for the ENT steering function, into the slow and fast components which comprise the satellite clock correction messages. The rate-of-change estimates are used to project the corrections forward in time by the expected system latency.

3.5. Broadcast of ENT through the GEO satellites.

The satellite fast and slow clock corrections are disseminated to EGNOS users via the GEO satellites in separate messages. The EGNOS user receiver decodes these messages, rebuilds the combined satellite clock offsets from ENT for each satellite it tracks, and applies them to its pseudo-ranges, together with other WAD corrections provided by EGNOS. In this way, ENT replaces the GPS (or GLONASS) Time in the receiver's navigation solution. Hence, not only does the receiver compute an improved estimate of position, because of the WAD corrections, but it also computes an estimate of its internal receiver clock offset from ENT.

Once available on the ground, ENT has to be accurately transferred on board the GEO satellites. Indeed, the GEO Time is defined at the output of the GEO payload, precisely at the L1 antenna centre of phase. This function is ensured by the so-called Long-Loop, a servo-control mechanism, based on the almost symmetry between the up and down links from the transmitting Earth station to the GEO satellite.

The time difference between ENT and GPS Time is not transmitted in EGNOS messages.

3.6. Relationship of ENT and UTC.

To synchronise ENT and UTC, ESA decided to put a special RIMS called "RIMS-UTC" in OP. The functional diagram of ENT – UTC computation and broadcasts is given in the figure 7.

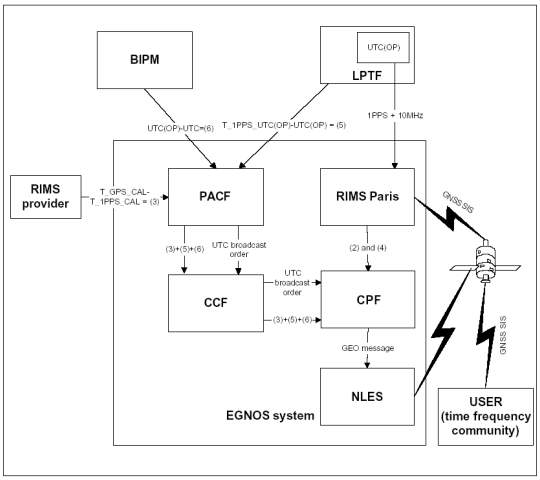


Figure 7. Functional diagram of ENT-UTC computation and broadcast [10].

The difference ENT – UTC is computed by CPF within EGNOS system using the following formula :

ENT – UTC = [ENT – UTC(OP)]
+ [UTC(OP) – UTC]

where ENT – UTC(OP) is determined using GPS common view technique with an uncertainty of less than 3 ns (2 σ), and UTC(OP) – UTC is computed by CPF in agreement with the OP extrapolation prediction method given in the chapter 2 above.

3.7. Relationship between ENT and UTC(OP).

The RIMS-UTC is physically connected to UTC(OP) with interfaces at 10 MHz and at 1pps. CPF EGNOS RIMS synchronisation module estimates directly the time difference between ENT and UTC(OP) using precise GPS phase and code common view technique, with an uncertainty of less than 3 ns (2 σ). The block diagram of the RIMS-UTC is given in the figure 8.

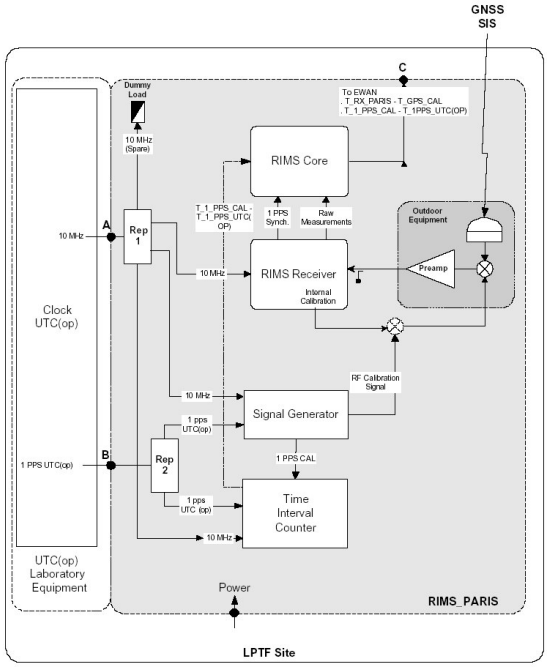


Figure 8. Block diagram of the RIMS-UTC [10].

In order to use the RIMS GPS multichannel receiver as it is, and to have an accurate synchronisation between the 1pps UTC(OP) and the internal time base of the RIMS receiver, it vas decided to use the method developed by CNES in the EURIDIS experiment [11, 12]. An RF calibration GPS-like signal synchronous to the 10 MHz and 1pps UTC(OP) is sent to the input of the receiver, and this signal is processed by the RIMS receiver simultaneously with the GPS and GEO signals.

The time equation of the RIMS-UTC is the following:

ENT – UTC(OP) = (ENT – T_RX-PARIS) (1)
+ (T_RX_PARIS – T_GPS_CAL) (2)
+ (T_GPS_CAL – T_1pps_CAL) (3)
+ (T_1pps_CAL – T_1pps_UTC(OP)) (4)
+ (T_1pps_UTC(OP) – UTC(OP)) (5)

where (1) is the offset between ENT and the internal RIMS Paris receiver (T_RX_Paris). It is estimated by CPF using all GPS raw measurements transmitted by all the RIMS and by the RIMS-UTC;
(2) is the result of the calibration code measurements, which correspond to the RIMS-UTC receiver measurements using the GPS like calibration signal. Measurements are sent every second to CPF;
(3) is the calibration offset: corresponds to the time offset between the RF calibration signal and the 1pps signal delivered by the calibrator generator;
(4) is the 1 pps measurement, measured by the time interval counter. Data are sent every second to CPF;
(5) is the time offset between the 1 pps_UTC(OP) taken at the input of the counter and the 1pps of UTC(OP) clock. This value is determined by OP.
Note that this method allows to use a GPS navigation multichannel receiver for time transfer. GPS, GEO and Calibration signals are received simultaneously in the RIMS receiver, and the internal delay of the navigation receiver is eliminated.
In order to obtain an accurate synchronization between ENT and UTC(OP), terms (3) to (5) were determined precisely in factory tests. Propagation delays of preamplifier, cables, splitter were measured accurately using precise counter, digital scope and microwave network analyser. The error budget for the calibration equipment terms (3) to (5) is estimated to 1,8 ns (1σ).

4. CONCLUSION

The inclusion of H-Masers in the generation of UTC(OP), together with the composite ensemble of the Caesium standards, will represent in a near future a major step forward. One expects that the short term stability, over a few days, will be given by one H-Maser, while the long term stability will be given by the Caesium clock ensemble. The reliability of such a clock signal would be insured by the use of more than one H-Maser and, as currently, of a few commercial Caesium standards. A future step might be the inclusion of the laboratory Primary Frequency Standards into the time scale computation. These developments will lead to a better UTC-UTC(OP) Key Comparison in the frame of the Mutual Recognition Agreement, and will improve the broadcast signal ENT-UTC.
The Network Time of EGNOS navigation system is connected to UTC(OP) and consequently to UTC. The time difference ENT-UTC will be broadcast in the navigation messages with an accuracy better than 100ns (3 σ), which will allow a direct use of EGNOS by the time user community. This established link between the European SBAS and UTC prepares the future connection between the Galileo Time and UTC: the Galileo Time Service Provider prototype project will develop and operate a composite-clock involving several European UTC(k) in order to provide the Galileo Time connection to UTC.

ACKNOWLEDGEMENTS

The authors acknowledge the EGNOS ESA Team, the EGNOS CNES team, and the EGNOS industrial teams for their cooperation.

REFERENCES

[1] J. Benedicto et al., EGNOS : the European Satellite Based Augmentation to GPS and GLONASS, Proc. of the GNSS98, October 98, Toulouse, France, § II-O-01, pp. 1-9.

[2] D. Flament et al., EGNOS, the European Based Augmentation to GPS and GLONASS – mission and system architecture, Proc. of the GNSS98, October 98, Toulouse, France, § II-O-12, pp. 1-10.

[3] P. Fenton et al., The use of the Wide Area Augmentation System (WAAS) as a time transfer system, ION NTM 2000, 26-28 January 2000, Anaheim, CA.

[4] P. Uhrich et al. , Le point sur la référence de temps française UTC(OP) et raccordement du système européen de radionavigation par satellites EGNOS à UTC(OP), Bulletin du BNM, Vol. 2004-1, n°124, pp. 25-35.

[5] M. A. Weiss, D. W. Allan, T. K. Peppler, A Study of the NBS Time Scale Algorithm, IEEE Trans. on Instrumentation and Measurement, Vol. 38, n°2, April 1989, pp. 631-635.

[6] B. Rueff, J. A. Davis, E. Pardo, The development of a clock algorithm at NPL, Proc. of the 14th European Frequency and Time Forum (EFTF), 14-16 March 2000, Torino, pp. 398-402.

[7] K. R. Brown, The theory of the GPS composite clock, Proc. of the ION GPS 91, pp. 223-241.

[8] J-M. Pieplu et al., EGNOS algorithms performance status and experiment activities, Proc. of the GNSS98, October 98, Toulouse, France, § IX-O-14, pp. 1-11.

[9] S. Peck et al., WAAS network time performance and validation, Proc. of the ION GPS 97, pp. 1123-1131.

[10] B. Charlot, Alcatel Space, EGNOS, Interface Control Document for UTC interface, DRD 0112/0032, Iss. 2.A.

[11] Ph. Gouni, M. Brunet, H. Secrétan, P. Uhrich, Time and frequency aspects in EURIDIS, Proc. of the 11th EFTF, 4-7 March 1997, Neuchâtel, pp. 325-330.

[12] M. Brunet et al., The role of time and frequency in EGNOS, Proc. of the PTTI 98, Reston, pp. 469-482.

Using DSP Technique to Average ADC Samples for Improving the Frequency Measurement Resolution

Shang-Shian Chen, Chia-Shu Liao

National Standard Time and Frequency Lab., TL, Chunghwa Telecom Co. Ltd.
No. 12, Lane 551, Min-Tsu Road Sec. 5 Yang-Mei,
Taoyuan, Taiwan 326, R.O.C.
Tel: 886 3 4245245, Fax: 886 3 4245474
berry@cht.com.tw

Keywords: frequency stability, frequency measurement, phase measurement, DSP technique, ADC.

Abstract

Most present common counters can't be used to measure the frequency by using the gate time less than 0.01 second. In addition, the frequency measurement resolution of common counters is about 3×10^{-10} in 1 second gate. In fact, the frequency resolution is inversely proportional to the gate time. For improving the frequency resolution, using the high-speed digital signal processor (DSP) to average the signal samples is created. We use analog-to-digital converter (ADC) to sample the device under test (DUT), and the frequency standard is used to be the reference clock. The DSP averages every ADC's sample to represent one of the DUT's measurements in the time interval of the gate time. Because it is able to average the samples and record the measurements in the time interval of each gate time that is specialized from 1 microsecond to 1 second, the DSP can measure the continuous samples of the DUT. Additionally, the measurement noise can be reduced and the measurement resolution can be improved by averaging the samples. According to the measurements, the frequency difference between DUT and frequency standard can be evaluated. Since the resolution in our design is inversely proportional to not only the gate time but also square of the number of the samples for averaging, it is better than that of common counters when gate time is larger. Additionally, the frequency stability less than one second is able to be evaluated in our measurement system.

1 Introduction

Because the stability of the frequency standard, like the cesium and the rubidium clock, are about 2×10^{-11} to 5×10^{-12} at one-second interval. The common counter can't be used to measure the stability of them at equal to or less than one-second interval. For the reason, the high precision counter is designed to satisfy with the clocks. In fact, the DSP technique is a good way to be used to measure the frequency offset and phase difference of the DUT. The DSP can process mass data in very short time. So, we can use the ADC to sample the output of the DUT and then put the sample data into the DSP. We use the frequency standard to be the reference clock of the ADC, and then the frequency difference between the DUT and the frequency standard will be evaluated by the DSP. In the DSP, the measurement noise can be reduced and the measurement resolution can be improved by averaging the samples. This paper provides a brief introduction of the system architecture. Then, the noise in the system is considered. Finally, the self test of our system is presented.

2 System architecture

In our system, we use ADC to sample the DUT and save the data into FPGA (Field Programmable Gate Array) that is the buffer for sample data. The FPGA packet the sample data into the frame, and then the DSP can read the frame and clear it in time when the frame is full. We use the frequency standard to be the reference clock of the ADC, and then the frequency difference between the DUT and the frequency standard will be evaluated by the DSP. The figure 1 shows the block diagram of the system.

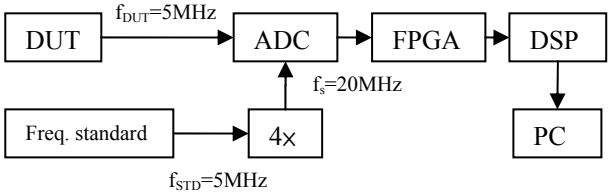


Figure 1. The block diagram of the System

The DSP will average the sample data and obtain the necessary information. We use the 4 time doubler to make 5MHz frequency standard to 20MHz for the sampling rate of the ADC. The sample data divides the phase of the DUT nearly every 90^0. The figure 2 shows the example about sample data. Despite the sample rate of the ADC in our system is 20MS/s, the speed of the DSP is not good enough to process every sample in time. So we set the minimum gate time is 1μs. If the DSP doesn't average the sample data, the sample data can be expressed below.

In the 1μs gate, the sample data is $S_{a,1\mu s}$、$S_{b,1\mu s}$、$S_{c,1\mu s}$、$S_{d,1\mu s}$:

$S_{a,1\mu s}=\{a_1, a_2, a_3,\ldots a_N\}$
$S_{b,1\mu s}=\{b_1, b_2, b_3,\ldots b_N\}$
$S_{c,1\mu s}=\{c_1, c_2, c_3,\ldots c_N\}$
$S_{d,1\mu s}=\{d_1, d_2, d_3,\ldots d_N\}$

In the 10μs gate, the sample data is $S_{a,10\mu s}$、$S_{b,10\mu s}$、$S_{c,10\mu s}$、$S_{d,10\mu s}$:

$S_{a,10\mu s}=\{a_1, a_{11}, a_{21},\ldots, a_{10\times N+1}\}$
$S_{b,10\mu s}=\{b_1, b_{11}, b_{21},\ldots, b_{10\times N+1}\}$
$S_{c,10\mu s}=\{c_1, c_{11}, c_{21},\ldots, c_{10\times N+1}\}$
$S_{d,10\mu s}=\{d_1, d_{11}, d_{21},\ldots, d_{10\times N+1}\}$

In the 1s gate, the sample data is $S_{a,1s}$、$S_{b,1s}$、$S_{c,1s}$、$S_{d,1s}$:

$S_{a,1s}=\{a_1, a_{10^6+1}, a_{2\times 10^6+1},\ldots, a_{N\times 10^6+1}\}$
$S_{b,1s}=\{b_1, b_{10^6+1}, b_{2\times 10^6+1},\ldots, b_{N\times 10^6+1}\}$
$S_{c,1s}=\{c_1, c_{10^6+1}, c_{2\times 10^6+1},\ldots, c_{N\times 10^6+1}\}$
$S_{d,1s}=\{d_1, d_{10^6+1}, d_{2\times 10^6+1},\ldots, d_{N\times 10^6+1}\}$

Transferring the sample data to the phase difference and the frequency offset.

Phase Difference:

$$P_d(i)=\tan^{-1}\left(\frac{S_{a,Gt}(i)-S_{c,Gt}(i)}{S_{b,Gt}(i)-S_{d,Gt}(i)}\right)\cdot\frac{1}{2\pi}\cdot\frac{4}{f_s}\cdot(-1) \quad \ldots\ldots(1)$$

Frequency offset:

$$y(i+1)=\frac{p_d(i+1)-p_d(i)}{Gt}\cdot(-1) \quad \ldots\ldots(2)$$

i is an integer and 1<i<N
Gt: Gate time
*f*s: sample frequency

The 1μs-gate resolution is about 10^6 times the resolution of 1s gate without averaging sample data. In our test, the 1s-gate resolution is about 2×10^{-10}, and it approximates to the common counter's resolution. If the DSP averages the sample data, the sample data can be expressed below.

In the 10μs gate, the sample data is $S_{a,10\mu s}$、$S_{b,10\mu s}$、$S_{c,10\mu s}$、$S_{d,10\mu s}$:

$$S_{a,10\mu s}=\{\sum_{n=1}^{10}a_n,\ \sum_{n=11}^{20}a_n,\ \sum_{n=21}^{30}a_n,\ldots,\ \sum_{n=10\cdot(N-1)+1}^{10\cdot N}a_n\}$$

$$S_{b,10\mu s}=\{\sum_{n=1}^{10}b_n,\ \sum_{n=11}^{20}b_n,\ \sum_{n=21}^{30}b_n,\ldots,\ \sum_{n=10\cdot(N-1)+1}^{10\cdot N}b_n\}$$

$$S_{c,10\mu s}=\{\sum_{n=1}^{10}c_n,\ \sum_{n=11}^{20}c_n,\ \sum_{n=21}^{30}c_n,\ldots,\ \sum_{n=10\cdot(N-1)+1}^{10\cdot N}c_n\}$$

$$S_{d,10\mu s}=\{\sum_{n=1}^{10}d_n,\ \sum_{n=11}^{20}d_n,\ \sum_{n=21}^{30}d_n,\ldots,\ \sum_{n=10\cdot(N-1)+1}^{10\cdot N}d_n\}$$

In the 1s gate, the sample data is $S_{a,1s}$、$S_{b,1s}$、$S_{c,1s}$、$S_{d,1s}$:

$$S_{a,1s}=\{\sum_{n=1}^{10^6}a_n,\ \sum_{n=10^6+1}^{2\times10^6}a_n,\ \sum_{n=2\times10^6+1}^{3\times10^6}a_n,\ldots,\ \sum_{n=(N-1)\times10^6+1}^{N\times10^6}a_n\}$$

$$S_{b,1s}=\{\sum_{n=1}^{10^6}b_n,\ \sum_{n=10^6+1}^{2\times10^6}b_n,\ \sum_{n=2\times10^6+1}^{3\times10^6}b_n,\ldots,\ \sum_{n=(N-1)\times10^6+1}^{N\times10^6}b_n\}$$

$$S_{c,1s}=\{\sum_{n=1}^{10^6}c_n,\ \sum_{n=10^6+1}^{2\times10^6}c_n,\ \sum_{n=2\times10^6+1}^{3\times10^6}c_n,\ldots,\ \sum_{n=(N-1)\times10^6+1}^{N\times10^6}c_n\}$$

$$S_{d,1s}=\{\sum_{n=1}^{10^6}d_n,\ \sum_{n=10^6+1}^{2\times10^6}d_n,\ \sum_{n=2\times10^6+1}^{3\times10^6}d_n,\ldots,\ \sum_{n=(N-1)\times10^6+1}^{N\times10^6}d_n\}$$

The phase difference and the frequency offset can be transferred from (1) and (2). The 1μs-gate resolution is 10^9 times the resolution of the 1s gate with averaging sample data. In our test, the 1s-gate resolution is about 2×10^{-13}, and it is more precise than the common counter's resolution.

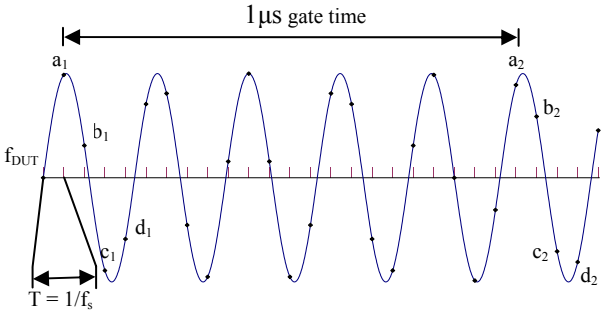


Figure 2. The example about sample data

3 System error

The system contains some errors, and the errors will influence the measurement resolution. The errors are the quantization error, the ADC sampling noise and the ADC time base error. Because we use the doubler that is the very low phase noise device to make frequency standard from 5MHz to 20MHz for ADC time base, the ADC time base error is very small. So, the major part of the error is the quantization error and the ADC sampling noise. In our system, we use the 8-bit ADC, and the quantization error can be estimated by using the computer simulation. The quantization error causes the frequency stability about 1.2×10^{-4} at the 1μs interval. In the other hand, we make the self test to find the sampling noise. When the DUT is the h-maser and the h-maser is also be the frequency standard, the frequency stability is 1.61×10^{-4} at the 1μs interval. So, we can estimate that the sampling noise causes the frequency stability about 4×10^{-5} at the 1μs interval. Additionally, the quantization error and the ADC sampling noise can be reduced by averaging mass sample data. In the self test, we can see the frequency stability is 1.93×10^{-13} at the 1s interval when the gate

time is 1s. So, the resolution is inversely proportional to the gate time and square of 10^6 samples.

4 Self-test result

We make the self test to find the resolution of every gate time from 1μs to 1s. In the self test, the DUT is the H-maser and the H-maser is also be the frequency standard. Because the stability of the H-maser, produced by Quartzlock, is less than 2×10^{-13} in 1s interval, the time base error is very small. We record 4000 samples per gate time and transfer the samples to phase difference and frequency offset. The figures from 4 to 17 show the phase difference and frequency stability in every gate time. From the figures, the system performance in every gate time is verified. The figure 3 shows the system resolution in every gate time.

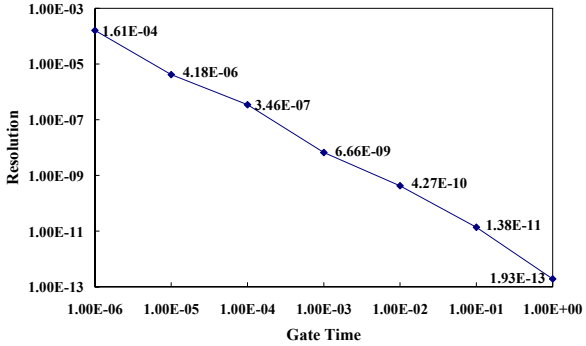


Figure 3. System resolution in every gate time.

5 Conclusions

We design a high precision counter for the phase and frequency measurement, and the resolution in our design is inversely proportional to the gate time and square of the number of the samples for averaging. In the self test, we can see the frequency stability is 1.93×10^{-13} at the 1s interval when the gate time is 1s. Despite the sample rate of the ADC in our system is 20MS/s, the speed of the DSP is not good enough to process every sample in time. So we set the minimum gate time is 1μs. In our test, we can see that processing more sample data in time can improve the system resolution. So, the higher speed DSP is very important for our system. The other problem is that the cache memory is also not sufficient to store much sample data. There are only 256K bytes of 133MHz synchronous burst static random-access memory (SBSRAM) in the DSP, and they can only save 4000 samples in the 1s gates. We will solve the problem in the future work.

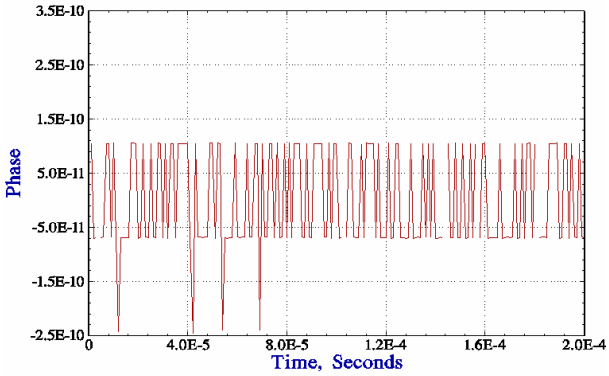


Figure 4. Phase difference in the 1μs gate.

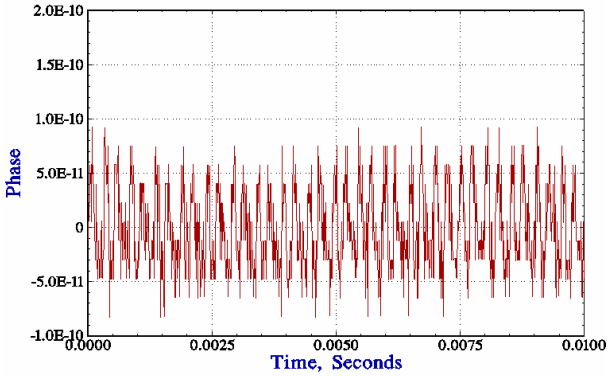


Figure 5. Phase difference in the 10μs gate.

Acknowledgements

We gratefully acknowledge the NBS (National Bureau of Standards) of the R.O.C. for supporting this project.

References

[1] Texas Instruments, "8-bit high-speed analog-to-digital converter, TLC5540", Technical documents, 2004.

[2] Texas Instruments, "TMS320C6201/6701 evaluation module technical reference", Technical documents, 1998.

[3] Xilinx, "XC4000E and XC4000X series field programmable gate arrays", Technical documents, 1999.

[4] Stanford Research Systems, "SR620 universal time interval counter", Technical documents, 1999.

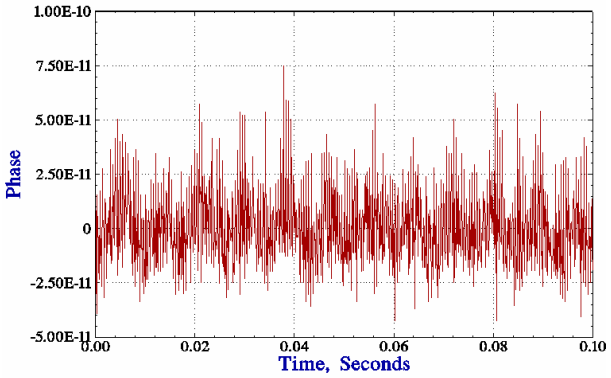


Figure 6. Phase difference in the 100µs gate.

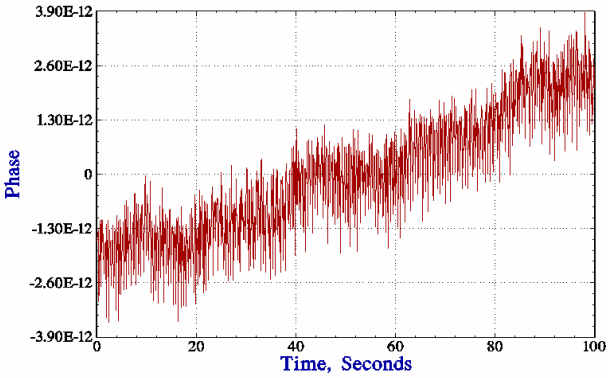


Figure 9. Phase difference in the 100ms gate.

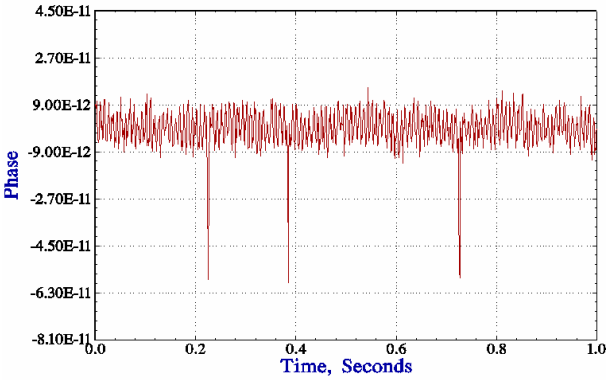


Figure 7. Phase difference in the 1ms gate.

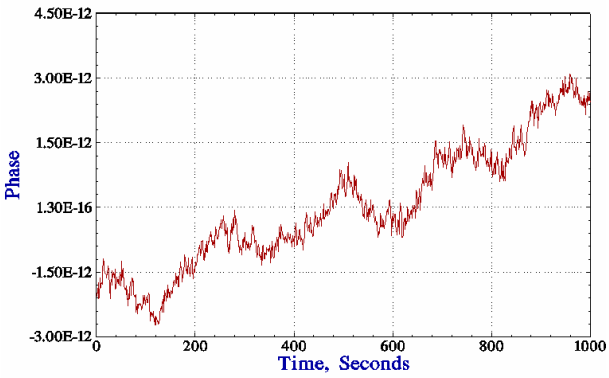


Figure 10. Phase difference in the 1s gate.

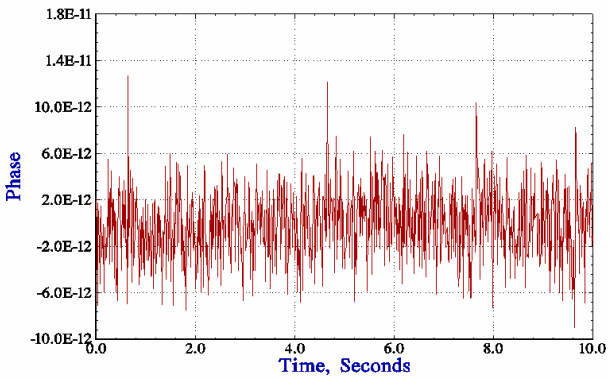


Figure 8. Phase difference in the 10ms gate.

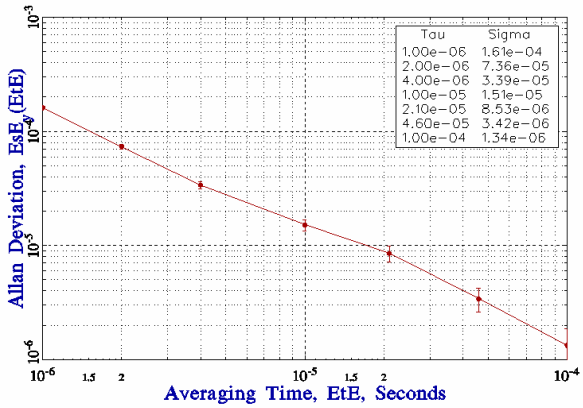


Figure 11. Frequency stability in the 1µs gate.

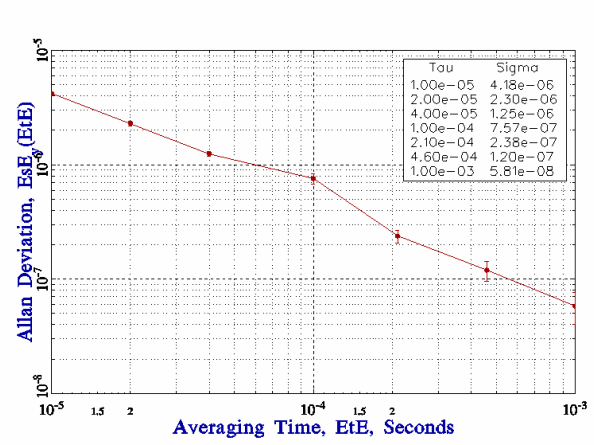


Figure12. Frequency stability in the 10µs gate.

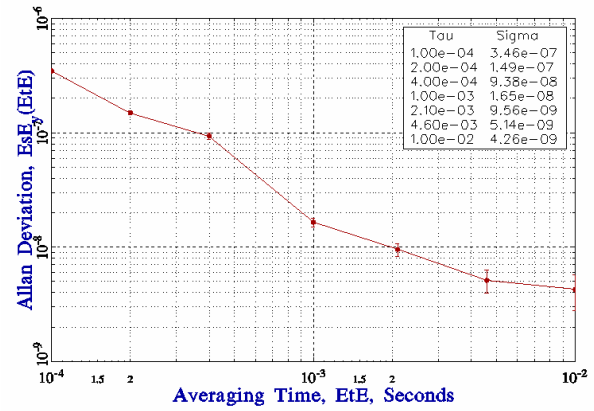


Figure 13. Frequency stability in the 100µs gate.

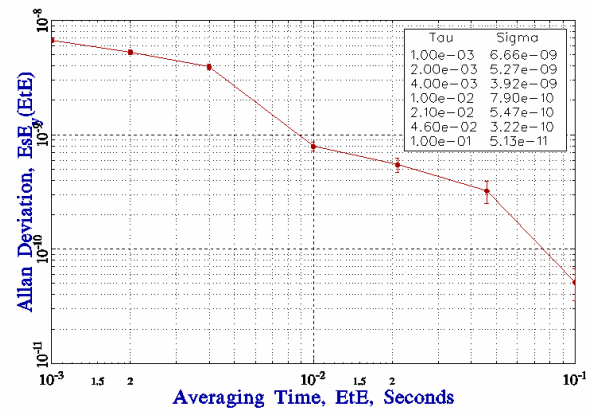


Figure 14. Frequency stability in the 1ms gate.

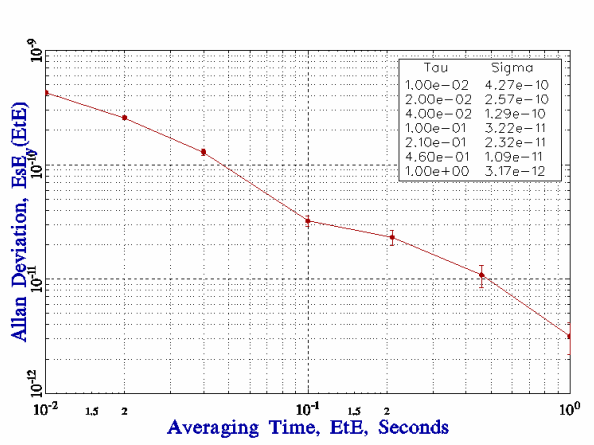


Figure 15. Frequency stability in the 10ms gate.

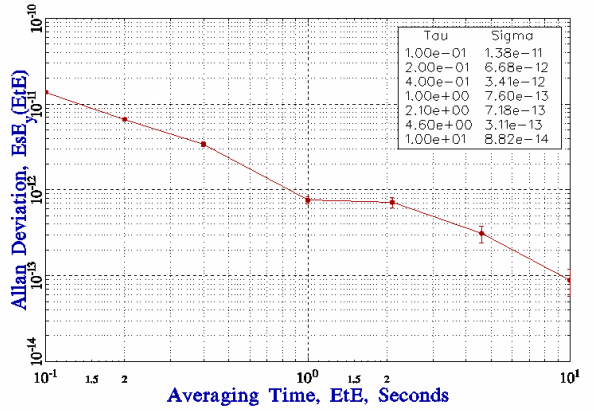


Figure 16. Frequency stability in the 100ms gate.

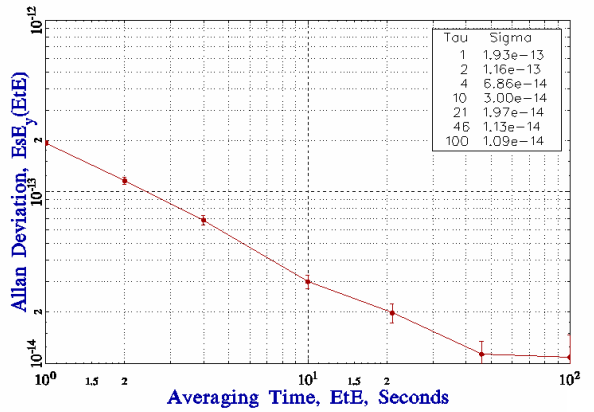


Figure 17. Frequency stability in the 1s gate.

DESIGN AND IMPLEMENTATION OF A STANDARD TIME TRACKING AND KEEPING SYSTEM FOR THE REAL-TIME CLOCK

C.C. LIN* and C.S. LIAO

National Standard Time & Frequency laboratory
Telecommunication Laboratories, Chunghwa Telecom Co., Ltd.
12, Lane 551, Min-Tsu Road Sec. 5, Yang-Mei, Taoyuan, Taiwan 326
Tel: 886-3-424-5471 Fax: 886-3-424-5474
E-mail: cchlin@chttl.com.tw

ABSTRACT

The purpose of this paper is to describe a development of an intelligent real-time clock (IRTC) based on system-on-programmable chip (SOPC) and single-chip microcomputer techniques. The whole project is conceptually composed of two parts.
The first part is concerned with tracking and processing standard time signal from the Taiwan's computer time system (TCTS) through the telephone line [1]. The algorithm of their operation was investigated and implemented on a MCS-51 single-chip in C.
Design and implementation of the timekeeping module on the field-programmable gate array (FPGA) device is the second part. We was using a finite-state machine (FSM) included state graphs and timing diagrams to describe the hardware operational processes, systematically. Some designs, such as offset adjustments, periodic phase modification, included in stuff/swallow counter were detailed to insure that the clock operations are always monotonic increasing and never run backwards. In addition, the practicable accuracy–keeping algorithm was also devised.

1. INTRODUCTION

Traditionally, a computer system has a real-time clock (RTC) for its timekeeping faculty, but which cannot synchronize automatically to external standard time and also have no a self-compensated mechanism for its time draft adjustment. Generally, a RTC clock usually consists of a hardware latch that is set by overflow of hardware counter. The clock requires a processor interrupt on every tick. These interrupt can cause significant overhead if the tick interval is small. Therefore, their time precision is always limited to the tick interval [3]. Aimed at the disadvantage of the RTC clock above, we designed an intelligent RTC clock to improve their performance called IRTC. Thus, special designs of hardware and software are required beside the traditional RTC. For example, we are necessary to design a time accumulator that has good resolution for time scalar to increasing adjustable precision and to design a tracing module that can trace an external standard to maintain system synchronism at fixed time interval, as well as, to design a self-compensated circuit to modify its system time drift. Also, the IRTC's time base was designed as a large range of useful realizations to suit different situations. That is, the oscillator can vary from the common and inexpensive, such as a quartz crystal, to the specialized precision of a hydrogen maser. Here, we used a common and inexpensive oscillator, quartz crystal, to satisfy our need because we can just only maintain the accuracy range in ±3 ms which limited by the accuracy of the external time reference. The external time reference is based on the TCTS server located at TL (telecom. laboratories) in Taiwan. The TCTS signal is available automatically over the telephone line. Besides, the operation of the stand-along timekeeping module can reach resolution to 1us when it is in free run situation.

2. STRUCTURE

A design sketch of our IRTC constructed entirely of functional blocks is shown in Fig. 1. Basically, it consists of two modules: the tracking module and the timekeeping module. The tracking module tracks the reference time signal derived from the TCTS system. The output of the tracking module splits two directions□one is added to the loop 1 (labeled a1 to a5) mixed with the output of the *time accumulator* at the node Σ1 which resulted in a value called time difference (reference clock - local clock)□The other one is added to the loop 2 (labeled b1 to b3) and mixed with the output of the *slop timer* at the node Σ2 which resulted in a value called drift rate that was due to frequency offset of the system oscillator. The controller calculated both values using the algorithm described in section 3 to characterize the performance of the system oscillator. The controller also uses this characterization to estimate optimal correction values feeding to the *programmable divider*. The programmable divider produces three types of pulse: *offset, offset-interval, and slop-interval*. Both offset-interval and slop-interval pulses flow in order, individually, to the stuff/swallow

counter. We work together both offset and offset-interval pulses timely to modify the time difference. Also the slop-interval pulse was used to modify the drift rate. The stuff/swallow counter deletes or stuffs a single pulse that belongs among the *pre-scalar's* pulse train. The time accumulators take the output of the pre-scalar's pulse train as its time-base to accumulate *time* and *date*. In other words, if we can control how quantity of pulses and their individual timing feed timely to the stuff/swallow counter, we will modify the accuracy of the system clock within prescribed range. The slop timer was in charge of accumulating the pluses (one microsecond per cycle) came from the 10MHz system oscillator divided by 10.

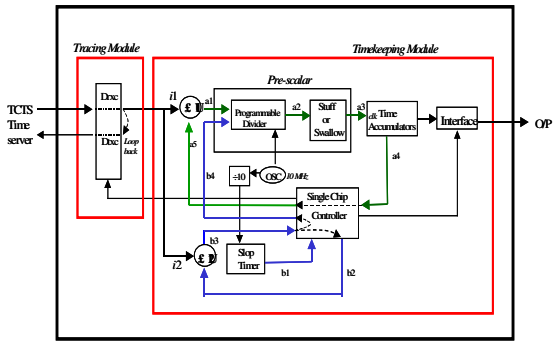


Fig. 1. Functional block diagram of our IRTC

3. RECALIBRATION INTERVAL

A time system of known drift rate can be kept within prescribed range of error with infrequent adjustments through a systematic approach. It usually preset to offsets in a fixed recalibration interval to keep the time system operating within a selected accuracy despite the oscillator's drift. We have known a basic equation of a time system that the total time error E at any time t can be expressed as formula (1) [2]. This drift must be known and must be constant.

$$E = E_0 + \left(\frac{f_0}{f_r} - 1\right)t + \frac{at^2}{2} \qquad (1)$$

Where:

E: Total time error at any time
E_0: Initial time error
f_0: Initial frequency at time $t = 0$
f_r: Reference frequency
a: Aging rate

We usually use crystal oscillator to implement our time system because it is inexpensive and easily to acquire. In our case, a common crystal oscillator is typical drift (aging rate) about 5ppm/year. This drift rate seems less affected about the accuracy of the system time if its recalibration interval is shorter than a day. That is, if we ignore the drift rate, we may calculate the recalibration interval T as follows:

If $t = T$, $E = -E_0$, according to formula (1), we obtain

$$-E_0 = E_0 + \left(\frac{f_0}{f_r} - 1\right)T + 0 \qquad (2)$$

$$0 = 2E_0 + \left(\frac{f_0}{f_r} - 1\right)T \qquad (3)$$

$$T = \frac{2E_0}{1 - \frac{f_0}{f_r}} \qquad (4)$$

For example, consider a time system to be maintained within ±3 ms. The signal that drives the clock is derived from a common crystal with a known nameplate frequency departure $22x10^{-6}$. Find the length of the recalibration interval T?

To solve this problem numerically we substitute these values to formula (4) as

$$T = \frac{3 \times 2 \times 10^{-3}}{22 \times 10^{-6}} = 272\,\text{sec} \approx 4.5\,\text{min} \qquad (5)$$

The result means if we want to keep the range of accuracy for a time system among the range of ± 3ms, the system time should be recalibrated every 4.5 minutes. We think this recalibrated interval is so impractical because frequent operation would waste a lot of system resource. However, our purpose is to design a system that can keep accuracy in the range of ±3 ms any time, but the recalibration time interval only need once per hour. To achieve this goal, we used a slop method and a highly dynamic algorithm [4] to calculate as follow:

$$M = \frac{e_{t2} - e_{t1}}{t2 - t1} \qquad (6)$$

$$-M_{(old+1)} = (1-\alpha)M_{(old)} + \alpha M_{(new)} \qquad (7)$$

Where:

M: Slop
$t2 - t1$: Recalibration interval
e_{t2}: The time at $t2$
e_{t1}: The time at $t1$

Total time error = $-M_{(old+1)} \times$ Elapsed time

In formula (6), the IRTC will maintain a variable M that is the best current estimate time drift rate, here, we called *slop* which used to compensate the system time drift when it is working in the free run situation. Once the tracking module completed a tracking process, we can get a new slop $M_{(new)}$ every hour and then

updated the value $M_{(old+1)}$ calculated in formula (7). The variable α is a smooth factor that determines a weight represented by the new M value. Through repeatedly testing, we obtained a typical value α that is equal to 0.7 in our experience.

4. THE PRE-SCALAR DESIGN

The pre-scalar module consists of programmable divider and stuff/swallow counter. We used finite-state machine (FSM) including state diagram and timing diagram which functional behavior can be represented using a finite number of states to design this module. The functions of this module was designed that need at least to meet the following specifications [4][5][6]:

1. The system time adjustment is always monotonic increasing and never run backwards.
2. The system time adjustment never changes the oscillator frequency but always uses the periodic phase modification.
3. The system must have a mechanism to advance (stuff) or retard (swallow) the phase of the system oscillator.
4. The offset adjustment must be spread over multiple tick intervals to avoid a big time variation in a short time interval.

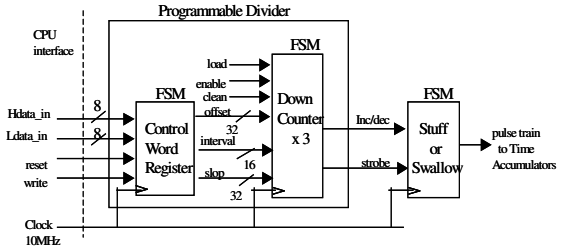


Fig. 2. The FSM block of pre-scalar module

In Fig. 2 the pre-scalar module includes a programmable divider and a stuff/swallow counter. The entire programmable divider is enclosed by the outer rectangle. The inner rectangle contains a control-word register and three down-counters: offset, interval, and slop.

4.1 THE PROGRAMMABLE DIVIDER

The control word registers latches three values: *offset-mapped*, *interval-mapped*, and *slop-mapped*. These values derived from the controller which run algorithms of the formula (6) and (7). Though five *write* actions, the control word register got three types of information: the time difference represented by offset (32 bit included 1 sign bit at MSB), the adjustment-interval of the time difference represented by offset-interval (16 bit), and the adjustment-interval of the time drift rate represented by slop-interval (32 bit included 1 sign bit at MSB). The state-transition diagram of the control word register is shown in Fig. 3. Three outputs of the control word register are all mapping to individual hardware down counter shown in Fig. 4 and the timing diagram are shown in Fig. 4-1. If we want to adjust the offset, then we should let the offset-interval counter is controlled by the offset counter. If offset-interval counter counts down to the zero, the offset counter would reduce one until the content is equal to zero. Once the offset-interval counter reduces to zero, then the slop-interval counter start working for the drift rate adjustment. In a word, when we adjust the time difference, we must feed a value to offset counter and offset-interval counter individually. Also when we adjust the drift rate, we must only feed a value to slop-interval counter. If only one of two interval counters reduced to zero, it would cause the stuff/swallow counter to add or delete a pulse in a normal series sequence produced by the system oscillator.

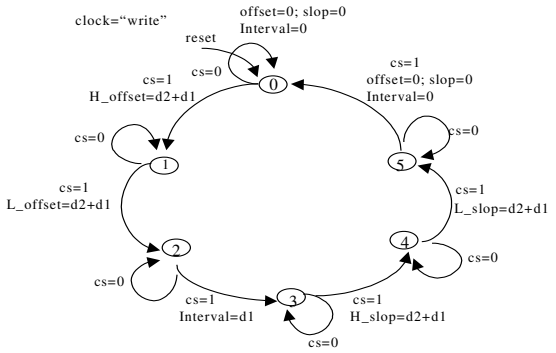


Fig. 3. State diagram of the control word register

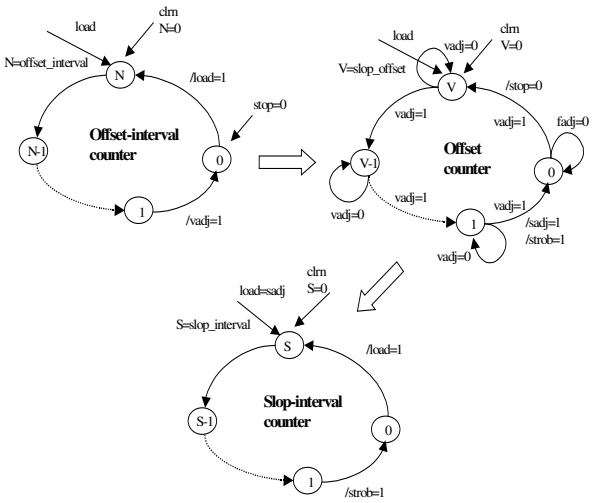


Fig. 4. State diagram of three down counters

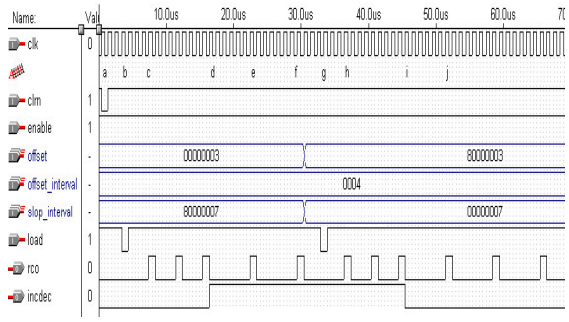


Fig. 4-1. Timing diagram of three down counters

4.2 THE STUFF/SWALLOW COUNTER

The pre-scalar also includes a state machine that is stuff/swallow counter. This counter is used to insert or delete the phase of the system oscillator by one cycle when system time need to correct. Every cycle is equal to 0.1 us. In Fig. 5, there is two control signals: one is labeled by the symbol *inc/dec*, the other is labeled by the symbol *strobe*. Whenever, if a reset signal active, it will transfer the machine to state *S0*. When the machine is in present state, it will change to next state if *strobe* is equal to *0* at the next active clock edge, such as *S0->S1; S1->S2; … S9->S0*. The machine will jump over two states at the next active clock edge when *inc/dec* is equal to *1*, then it will swallow one cycle at the output pulse train (speed time up). On the contrary, when *inc/dec* became equal to *0*, the machine will go back to the original state at the next active clock edge, then it will stuff one cycle at the output pulse train (slow time down). This stuff/swallow counter operates in a numerical decimal format. So we should consider 10 numerical statuses individually, such as, *0, 1, 2, 3, 4, 5, 6, 7, 8, 9*, because these statuses may individual appear randomly. The output pulse only appears in the two transition states that is *S8* transfers to *S9* or *S0*.

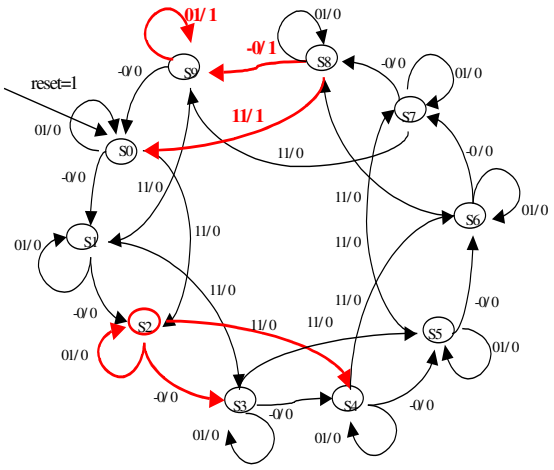


Fig. 5. State diagram of the stuff/swallow counter

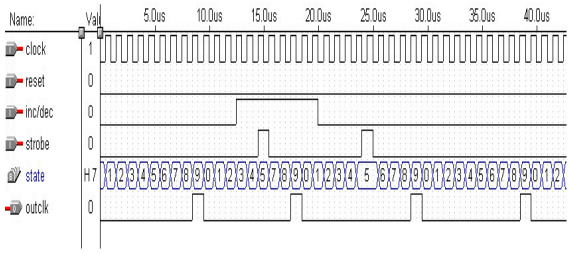


Fig. 5-1. Timing diagram of the stuff/swallow counter

In Fig. 5-1 shows the timing diagram. The node labeled *outclk* is the output pulse train for the *time accumulator* input. The *strobe* and *incdec* signal controls the system to increase or decrease a pulse to adjust the time error. Labeled *state* shows the count values at the inner register. When register count to the *9*, it will output a pulse at the *outclk* node. At the *state* node, we can see the *6* disappeared and *5* extended in Fig. 5-1.

5. THE PROGRAM DESIGN

The tracking module tracks the TCTS server located at TL in Taiwan. The TCTS is a time-server by using a loop-back mode (path delay is equal to 1/2 round-trip delay) providing modified time code over the telephone line. We mention the tracking and keeping process using the flow chart as following:

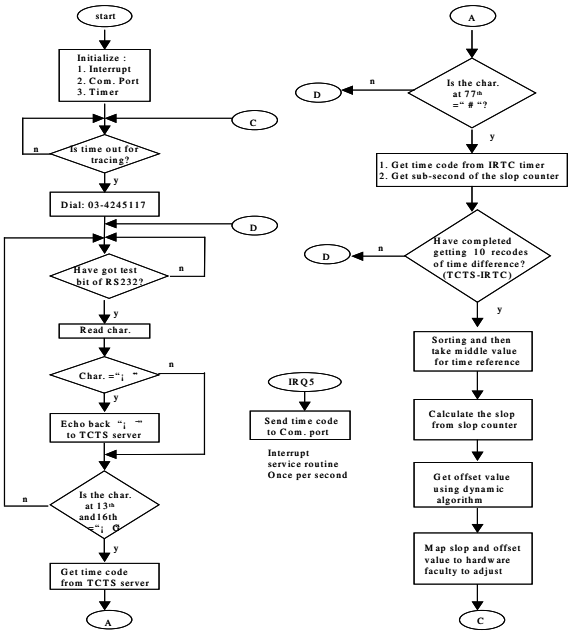


Fig. 6. Flowchart of the program

Fig. 6 shows the flowchart that compares both time codes between the TCTS's and the IRTC's to modify the IRTC's time difference and its drift rate. A few

things must be pointed out about this process. The time interval that tracking the TCTS server is once per hour. We always get 10 recodes of the time code from the TCTS. Among these data, we are sorting first and then find the middle value as a time reference for the foundation of the adjustment. The IRQ5 tells a controller interrupt at the every second to execute the service routine that output the IRTC's time code from the time accumulators for external applications.

6. MEASURED RESULTS

In Fig. 7 we collected the data of both accumulators, slop timer and time accumulator over 50 hours, and used these data to predict time deviation. Both clock input are all derived from the same oscillator 10 MHz which is the system time base. The slop timer is used to accumulate the system time error that produced due to the initial frequency offset when the system starts operation. The line labeled *R* represented time error in Fig. 7, its slop is approach to $-4.3419*10^{-6}$. Other one called time accumulator is the system practical output for the external applications. It was adjustment by the offset-counter and the slop-interval counter according to the value of the offset and the negative slop. The line labeled *B* represented the trend of the time accumulator's data once per hour, which has been modified seems closing to the x-axis. If we only magnify the line scalar labeled *B*, we can get Fig. 8 in which specify the maximum time error range between ±3 ms, and the trend of the time error closed to the line of *0* gradually.

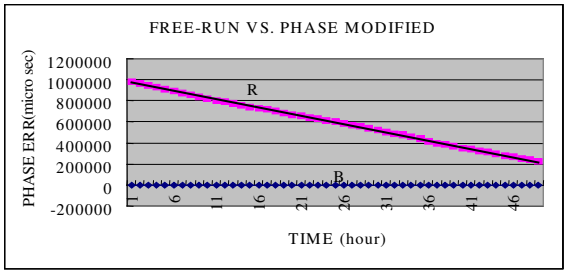


Fig. 7. Free-run vs. phase modified

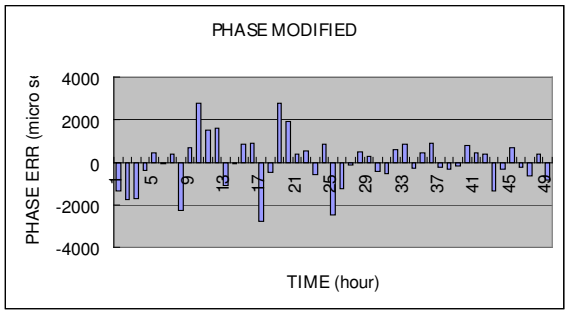


Fig. 8. Magnify the line labeled "B" in Fig. 7

7. CONCLUSIONS

In this paper, two main parts of the IRTC were described. The time tracking module intended to synchronize an external source of standard time using the loop-back compensated-delay theorem, delay=1/2 round-trip delay. The time keeping module, presented here, intended to keep standard time when system time oscillator works in the free run phase. There corrected the phase (time drift) by occasional vernier adjustments according to the timekeeping formula:

$$-M_{(old+1)} = (1-\alpha)M_{(old)} + \alpha M_{(new)}$$

In summary, both modules integrated can compensate systematic time offsets and drift rate, as well as various types of random effects like temperature variation. To analyze these data in section 6 above, we verified this IRTC's performance can keep the accuracy within ±3 ms whenever. This accuracy depends on external time reference. But if only a timekeeping module worked alone, it could have 1us resolution.

REFERENCES

[1] C. C. Lin, C. S. Liao, "Design and Implementation of 4 Ports Computer Time Service System in Taiwan", 13th European and Time Forum & IEEE Frequency Control Symposium, April 1999.

[2] "Timekeeping & Frequency Calibration", Application Note 52-2, HP Co., 1975.

[3] David L. Mills, "Modelling and Analysis of Computer Network Clocks", Electrical Engineering Department., University of Delaware Technical Report 92-5-2, May 1992.

[4] Andrew S. Tanenbaum, "Computer Networks", 4th edition, pp 550-552, (Prentice-Hall, Upper Saddle River New Jsersey, 2003).

[5] Michael D. Ciletti, " Advanced Digital Design with the Verilog HDL", pp 69-142, (Prentice-Hall, Upper Saddle River, New Jsersey, 2003).

[6] http://www.altera.com

DESIGNING AND INVESTIGATION OF DIGITAL MULTI-CHANNEL TEMPERATURE CONTROL SYSTEM WITH HIGH PERFORMANCE FOR HYDROGEN TIME AND FREQUENCY STANDARD

S.Yu. Medvedev, A.Yu. Pavlenko, P.N. Smirnov

"VREMYA-CH" Joint Stock Company
Osharskaya st. 67, 603105 Nizhni Novgorod, Russia

ABSTRACT

Nowadays the necessity to improve the output frequency parameters of hydrogen standards has put greater emphasis on the requirements to the system of temperature stabilization in the area of SHF resonance cavity and storage bulb. This paper is aimed to provide theoretical calculation and experimental verification of the maximum possible temperature stability of the object (resonance cavity with the storage bulb) when using in the digital multi-channel temperature stability system the principle of the heating element control within impulse mode according to pseudorandom principle. In addition to high precision temperature maintenance the advantages of the above mentioned system are the low power consumption and possible remote PC control.

As follows from the theory and subsequent experimental confirmation the worked out temperature stability system doesn't lead to the rise of constant or variable magnetic fields at the frequency of the possible frequency transitions in the area of storage bulb. Besides, using this system also allows not to make worse the spectrum of output reference signal, which seems impossible with the usual standard impulse control.

INTRODUCTION

In this paper we take into consideration only the shifts dependent on temperature, such as the second order Doppler effect [1]:

$$\frac{\Delta f}{f} \le -1.4 \cdot 10^{-13} / {}^{\circ}C \qquad (1)$$

and the shift caused by collisions of hydrogen atoms with coating of storage bulb (the so-called wall shift) [1]:

$$\frac{\Delta f}{f} \le -1 \cdot 10^{-12} / {}^{\circ}C \qquad (2)$$

In the present paper we don't take into account the changing of dimensions and, consequently, the resonance frequency of resonator due to temperature fluctuations. This is assumed as all the measurements are performed with the switched on automatic cavity tuning system (ACT).

Allan variance [2] in case of locked loop of ACT system is characteristic of short-time frequency instability of hydrogen discriminator. This characteristic is determined by both physical and design parameters of the instrument and is defined by the following equation [1]:

$$\sigma_y(\tau) = \sqrt{\frac{k_s kTF}{2\beta_1\beta_2}} \frac{(1+\beta_1+\beta_2)}{2Q_0} \frac{(1+S_0-\alpha)^2}{\alpha\sqrt{S_0}(1+S_0)} \tau^{-\frac{1}{2}}, \qquad (3)$$

where β_1 and β_2 - input and output coupling factors of resonance cavity, Q_0 - quality factor of the atomic unsaturated line, S_0 - the saturation factor of the atomic transition at resonance, k=1.38·10^{-23} J/K - Boltsman's constant, T - temperature of the cavity, F - noise figure of the receiver, $k_S=S_0/P_{in}$, where P_{in} is the interrogation power at the cavity loop input and α - oscillation parameter. We assume the following values: β_1=0.1, β_2=0.1, T=313K, Q_0=1.18·10^9, P_{in}=6.31·10^{-12}W, F=1.78 and k_S=1.8·10^{11}, as measured experimentally. The Allan variance for these parameters is:

$$\sigma_y(\tau) = 1.315 \cdot 10^{-13} \cdot \frac{(1+S_0-\alpha)^2}{\alpha\sqrt{S_0}(1+S_0)} \tau^{-\frac{1}{2}} \qquad (4)$$

or

$$\sigma_y(\tau) = A \cdot \tau^{-\frac{1}{2}} \qquad (5)$$

In the table 1 we present coefficient A dependent on the oscillation parameter α and on the saturation factor of the atomic transition S_0 .

Table 1. Dependence A on α and S_0

A		S_0			
		0.5	1.0	1.5	2.0
α	0.2	1.07·10^{-12}	1.08·10^{-12}	1.16·10^{-12}	1.24·10^{-12}
	0.3	6.06·10^{-13}	6.44·10^{-13}	7.05·10^{-13}	7.66·10^{-13}
	0.4	3.82·10^{-13}	4.28·10^{-13}	4.82·10^{-13}	5.33·10^{-13}
	0.5	2.52·10^{-13}	3.01·10^{-13}	3.50·10^{-13}	3.94·10^{-13}
	0.6	1.70·10^{-13}	2.19·10^{-13}	2.63·10^{-13}	3.03·10^{-13}
	0.7	1.15·10^{-13}	1.62·10^{-13}	2.02·10^{-13}	2.38·10^{-13}

For the values α=0.45 and S_0=1.2, measured for hydrogen frequency standard VCH-1006, we get:

$$\sigma_y(\tau) = 3.71 \times 10^{-13} \tau^{-\frac{1}{2}} \qquad (6)$$

The diagram of this dependence is shown in picture 1. It should be mentioned that this is the best value of frequency stability, which can be achieved in the given design of the physical part of passive hydrogen time and frequency standard. However, it is possible to achieve

the value of relative frequency instability up to 10^{-13} when α and S_0 parameters are accordingly optimized. Thus, all other effects, leading to worse output frequency stability must be half of this value.

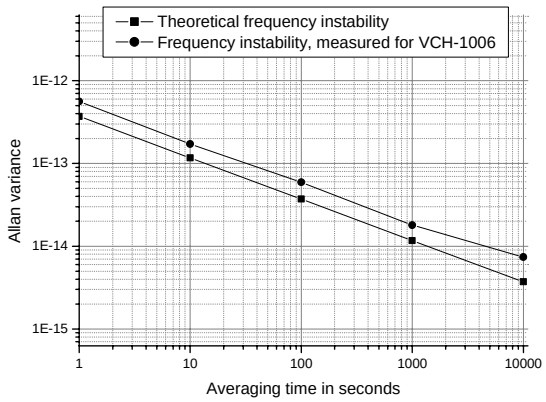


Figure 1. Theoretical frequency instability and the best practical frequency instability for VCH-1006

Taking into consideration the total influence of temperature on the frequency of quantum transition (expressions (1) and (2)) of order $-1.1\cdot10^{-12}$/ºC and equation (6) for the maximum achieved frequency stability of hydrogen discriminator, we can get the following requirement for temperature control system:

$$\sigma_T(\tau) \approx 10^{-1\,^\circ}C \cdot \tau^{-\frac{1}{2}} \qquad (7)$$

Here $\sigma_T(\tau)$ is temperature instability of resonance cavity of quantum hydrogen discriminator for temperature averaging time τ.

THEORETICAL PREDICTION

To solve the task determined by the cause (7) it is supposed to use the microprocessor system of digital temperature control (figure 2), which is based on the method of switch-heater control of object according to pseudorandom principle.

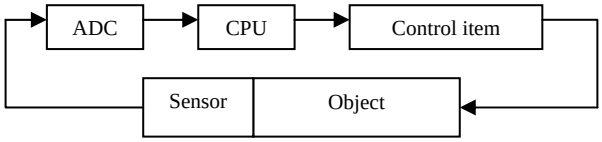


Figure 2. Block diagram of digital system of temperature control

This method allows to minimize the influence of interference noise of switch-heater on the object, by which we imply resonance cavity and storage bulb of hydrogen frequency standard.

What is of main interest in this scheme is the principle of heating element control in an impulse mode according to pseudorandom principle. The controlling element represents the electronic switch, which delivers voltage to the heater according to pseudorandom principle and in equal fixed periods of time. The output value of PI-regulator [3] (error signal) is compared with pseudorandom value, generated by microprocessor system. If the output value of PI-regulator is higher than the pseudorandom value at a separate time interval, the switch is turned on. This process is repeated with frequency ~10 kHz. The function of distribution of pseudorandom value has the linear characteristic along all the interval of possible values. That is why the heating power per unit of time will be proportional to output value of PI-regulator. This principle not only increases the efficiency factor due to the use of switch heater device, but also diminishes the interference noise due to more equal distribution of radiated energy along the spectrum achieved due to pseudorandom nature of the heater current impulses.

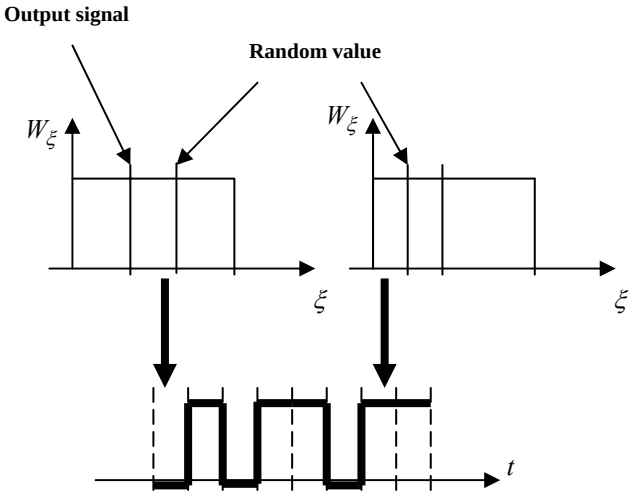


Figure 3. Working principle of temperature control system

Operation of discrete PI-regulator, used in the digital temperature control system, can be denoted by the following discrete time equation:

$$X(n) = k_1 P(n) + k_2 I(n), \text{ where} \qquad (8.1)$$

$$P(n) = aV(n) \qquad (8.2)$$

$$I(n) = I(n-1) + V(n) \qquad (8.3)$$

$$V(n) = T(n) - T^* \qquad (8.4)$$

Here T^* is the required value of the regulated parameter and $T(n)$ is its present value within the point of time number n. $V(n)$ means difference between the present and required values, $X(n)$ is the output value of the regulator, $P(n)$ and $I(n)$ are proportional and integral correspondingly.

Our task is to estimate the theoretically maximum possible temperature stability, that can be achieved when we use impulse nature of power in the above described digital system temperature control.

For calculations we assume the following:

1. Power impulses represent the process with the power spectral density of the following type (figure 4):

$$W(f) = A \cdot \left(\frac{\sin(10^{-4}\pi \cdot f)}{10^{-4}\pi \cdot f} \right)^2 \qquad (9)$$

We can estimate factor A from normalizing condition:

$$A \cdot \int_0^{\infty} \left(\frac{\sin(10^{-4}\pi \cdot f)}{10^{-4}\pi \cdot f} \right)^2 df = \frac{1}{2}\frac{U^2}{R} \qquad (10)$$

U=20V - supply voltage, R=100Ohm- resistance of heating element. Thus, A=4·10^{-4}.

2. Depending on the suggested analogy between the processes performed in thermal and electrical systems, the object of regulation (resonator with a storage bulb) is seen as a low-frequency filter with cutoff frequency defined by the thermal time constant. The time constant of the object τ_1, estimated theoretically, equals to 32319.6 seconds, what corresponds to cutoff frequency of low-frequency filter (11):

$$f_g = \frac{1}{2\pi\tau_1} \approx 4.92 \cdot 10^{-6}\, Hz \qquad (11)$$

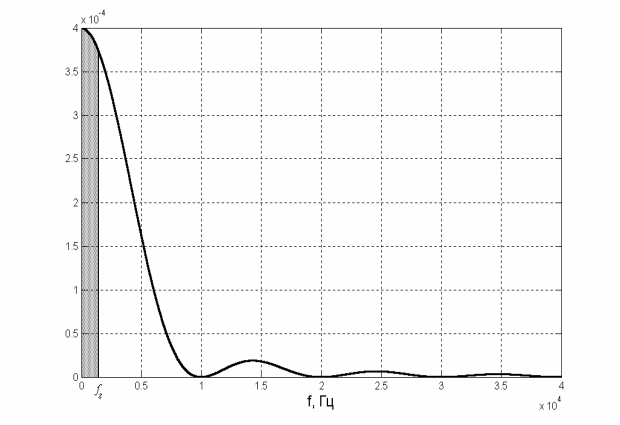


Figure 4. Spectrum of telegraph process

Dependence of the temperature of the regulated object on the supplied power is linear (and we take interest in minor fluctuations) and is defined by the equation:

$$T = 45°C \cdot \frac{W(t)}{W_{\max}} + 25°C \qquad (12)$$

Temperature fluctuations are defined by the value of power, which hit the band pass of the low-frequency filter.

After the pass of the low-frequency filter we get:

$$W_1 = 4 \cdot 10^{-4} \cdot \int_0^{4.92\cdot10^{-6}} \left(\frac{\sin(10^{-4}\pi \cdot f)}{10^{-4}\pi \cdot f} \right)^2 df = 1.88 \cdot 10^{-9} \qquad (13)$$

Consequently temperature variation is:

$$\sigma_T = 45°C \cdot \sigma_U = 45°C \cdot \sqrt{W_1} \approx 1.96 \cdot 10^{-4}°C\,, \qquad (14)$$

which is far less than the maximum allowable level, defined by the condition (7). Temperature variation σ_T in the expression (14) coincides with temperature stability $\sigma_T(\tau)$ in the expression (7) for short τ. This enables to assume that the suggested digital temperature control system (based on the method of switch-heater control of the object according to pseudorandom principle) may be applicable in the construction of quantum frequency standards on basis of hydrogen generator.

EXPERIMENTAL VERIFICATION

In support of theoretical results several practical measurements were carried out with the help of digital temperature control system.

At the first stage we defined the thermal time constant of quantum discriminator. For that after the rise in temperature inside the resonator to a certain value, we broke the heating element circuit. Using the values, obtained with the help of digital temperature control system, we can plot a diagram showing dependence of temperature of the regulated object on time. This diagram is given in figure 5.

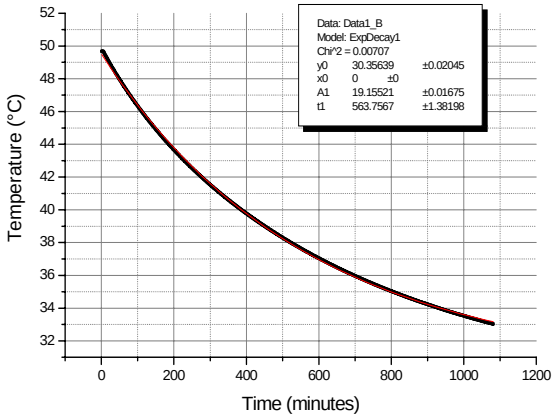


Figure 5. Defining temperature time constant of the discriminator

After approximation of this diagram by the first order exponential function (refer to the picture) we numerically found the time constant τ, which makes 536 minutes. This result corresponds to the thermal time constant, estimated theoretically in our previous work.

As the result of the experiments we obtained the so called curves of temperature control (dependence of temperature of the regulated object on time) for different values of integrator averaging time. The values of dependence are shown in picture 6 (for all the diagrams the time between points is 1 minutes, and temperature averaging time in program equals to 30 seconds).

The diagram, given in figure 6, show that, when averaging time is increased, the number of relaxation oscillations is diminished. But the time, required for amount the precise value of resonance cavity temperature, is increased.

Figure 7 presents the dependence of power, supplied to the heating element (the standing axis reflects the number of impulses per unit time), on time.

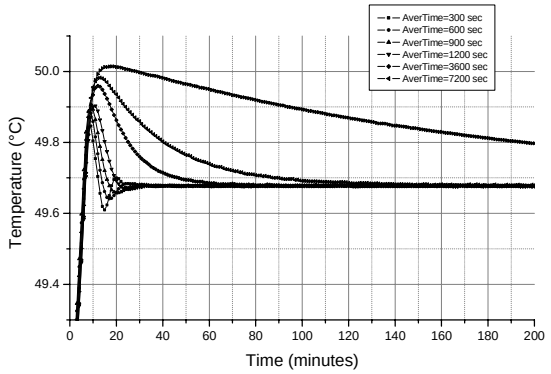


Figure 6. Dependence of temperature of the regulated object on time

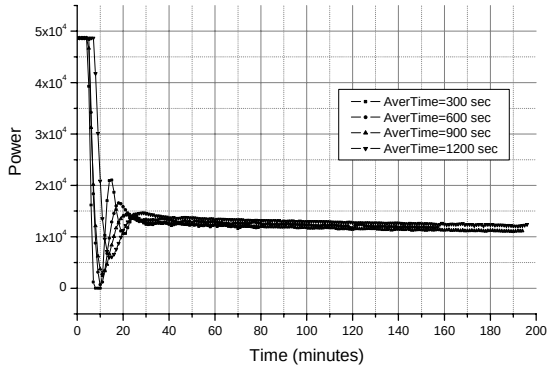


Figure 7. Dependence of the heater power on time

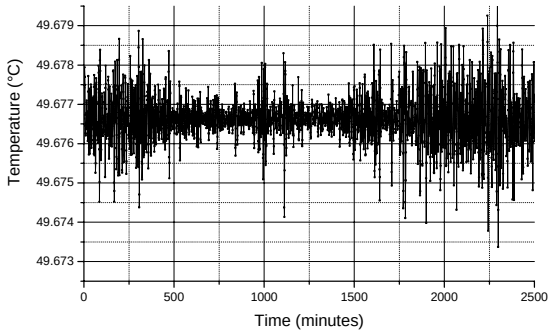
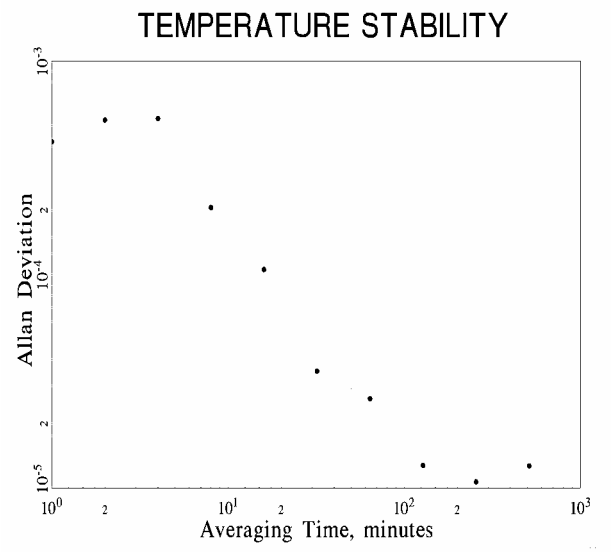


Figure 8. Temperature fluctuation of the object of regulation

Figure 8 represents dependence of temperature of the regulated object on time directly with temperature retaining. Data output period equals to 1 minute, temperatures averaging time is 54 seconds and integrator DSP averaging time is 300 seconds. As seen from the picture, the maximum temperature deviations of the regulated object (resonator with a storage bulb) from the average value do not exceed $3 \cdot 10^{-4}$ ºC. Using the obtained results we may plot the dependence of temperature instability (Alan variance [2]) on the averaging time. The diagram, representing this dependence, is shown in picture 10.



Picture 9. Dependence of temperature instability on averaging time

The diagram with Allan variance dependence on the averaging time shows that short-time temperature instability of the regulated object is not worse than 10^{-3} ºC when averaging time is 60 seconds and more. This corresponds to instability of frequency shifts of the quantum standard not more than

$$1.5 \cdot 10^{-14} \cdot \tau^{-\frac{1}{2}}. \qquad (15)$$

It should be mentioned that this kind of experimentally obtained dependence $\sigma(\tau)$ is not characteristic of the noise process, in other words, it is not a straight line in logarithmic system of reference (figure 1). For large averaging times this is, probably, connected with external temperature fluctuations. For the averaging times less than 3 minutes “non-noise” character of dependence $\sigma(\tau)$ is probably due to imperfection of the software used for generation of the random numbers, necessary for thermostat operation. The stated facts require further investigations.

REFERENCES

[1] C. Audoin, B. Guinot. The Measurement of Time. Cambridge: Cambridge University Press, 2001.

[2] D.B. Sullivan, D.W. Allan, D.A. Howe. NIST Technical Note, p. 1337, March, 1990.

[3] E.N. Bazarov, N.A. Demidov, E.T. Zhukov. Quartz and quantum frequency measures. USSR, Ministry of Defence Press, 1989.

TIME AND FREQUENCY METROLOGY ACCREDITED LABORATORIES IN BESANÇON

P. Salzenstein[1], F. Meyer[2], E. Tisserand[2], G. Martin[1], J. Petetin[2], N. Franquet[1], F. Lardet-Vieudrin[1], O. Lavedrine[2] and M. Vincent[2]

[1] FEMTO-ST Institute / LPMO & LCEP Dpt., CNRS UMR 6174
LNE associate laboratory No 2.13
32 avenue de l'Observatoire, 25044 Besançon Cedex
Phone: +33 (0)3 81 85 39 99, Fax: +33 (0)3 81 85 39 98, e-mail : patrice.salzenstein@lpmo.edu

[2] Observatoire de Besançon, CNRS UMR 6091,
LNE associate laboratory No 2.06
41bis, avenue de l'Observatoire, BP1615, F25010 Besançon, FRANCE
Phone: +33 (0)3 81 66 69 27, Fax: +33 (0)3 81 66 69 44, e-mail : francois.meyer@obs-besancon.fr

ABSTRACT

To organize traceability of any measurement to the national standards for calibration applications, Besançon is well equipped in the whole time and frequency domain. The LPMO department of the FEMTO-ST institute works on short term stability, spectral density of phase noise and ultra-stable oscillators. The observatory of Besançon is mainly involved in oscillator stability, chronometry, time scales and time and frequency transfer by GPS. Both laboratories are ISO 17025 accredited and associated to the Bureau National de Métrologie (BNM) – recently became the Laboratoire National de Métrologie et d'Essais (LNE).
By developing inter-laboratory time and frequency transfer by optical fiber, the laboratories are about to set up a local time scale of which long, middle and short term stability will be ensured by the observatory's cesium clocks, FEMTO-ST's hydrogen-maser and quartz or cryogenic oscillators. The current field of accreditation and the objectives of the two laboratories are presented in this paper.

KEYWORDS

Frequency stability, Phase noise, Time scale, Accreditation

INTRODUCTION

In France, for historical reasons, the time and frequency domain is geographically located in Paris and Besançon. Paris laboratories tradition in time measuring came from the observatories of Paris and Meudon when time references were an astronomers affair. In the past, the development of watches industries and laboratories in Besançon explains the needs that lead to the development of metrological expertise in the time and frequency domain.

The main public laboratories in this domain are the following :

LNE-SYRTE previously called LPTF (Laboratoire Primaire du Temps et des Fréquences), is the French primary laboratory (National Metrological Institute) in charge for the French standards of time and frequency.

The observatory of Besançon, and its LNE associate laboratory (# 2.06) is the main actor in charge for the dissemination of the national frequency reference for both academic and industrial laboratories, ensuring the traceability of their local references to the national standards.

The two following departments of FEMTO-ST institute, created in 2004 :

- LPMO (Laboratoire de Physique et Métrologie des Oscillateurs), and its LNE associate laboratory (# 2.13) that maintain and develop standards for short term frequency stability

- LCEP (Laboratoire de Chronométrie Electronique et Piézoélectricité) department that owns an Hydrogen Maser, standard for middle term frequency stability

The aim of this paper is to present the abilities of the accredited laboratory in the time and frequency

domain and some of the solutions that are under development for increasing their calibration capabilities.

1. SHORT TERM FREQUENCY STABILITY ACCREDITED LABORATORY

Created in 1958 by Professors A. Kastler (Nobel Prize of physics) and J. Ubersfeld, the Laboratoire de l'Horloge Atomique, Besançon's section, had from this time a frequency-time vocation. This laboratory took the name of Laboratoire de Physique et Métrologie des Oscillateurs in 1974 (LPMO). Finaly it joined FEMTO-ST at the beginning of the year 2004 and then became one of the departments of this CNRS laboratory. For several years the LPMO has been associated to the University and the Institute of Micro-technologies of Franche-Comté (respectively UFC and IMFC). Since 1973, LPMO is accredited for short term frequency stability calibrations. LPMO is able to deliver COFRAC (Comité Français d'Accréditation) calibration certificates, conform to ISO 17025. That relates to short term frequency stability in time domain and phase noise measurements in spectral domain [1].

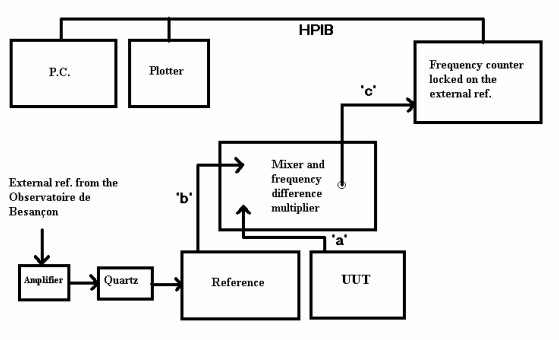


Fig. 1. Principle of the short term frequency stability calibration.

Despite the laboratory is able to fully characterize signals for several frequencies, it is officially accredited for characterizing Allan variance of 1 MHz-1 GHz frequency sources for integration time (τ) in the 0.1 s - 100 s range.

The best frequency stability that can be characterized under COFRAC accreditation is about 3.10^{-13} to 5.10^{-12} depending on carrier frequency and integration time.

The calibration is realized by comparison between the Unity Under Test (UUT) and a quartz reference locked on a Cesium. A frequency difference multiplier is then used to obtain a signal that can be analyzed by a frequency counter. Principle of this calibration and a picture of the bench are given on the figures 1 and 2.



Fig. 2. Frequency stability calibration bench.

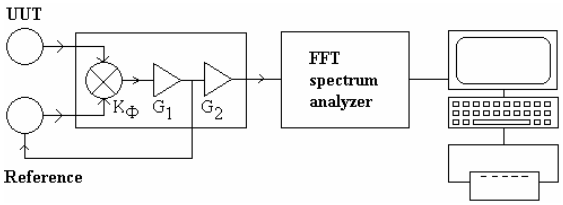


Fig. 3. Principle of the phase noise calibration.

The LNE associate laboratory can calibrate any frequency sources between 5 MHz and 18 GHz in terms of phase noise, and the laboratory is well equipped to characterize frequency signals up to 40 GHz. Despite it is possible to give the phase noise for Fourier frequencies between 0.01 Hz and several MHz, the official calibration capabilities concerns the Fourier frequencies between 1 Hz and 100 kHz from the carrier.



Fig. 4. Phase noise bench.

The noise floor of this bench is between – 172 dBc/Hz and –55 dBc/Hz depending on the carrier

and the Fourier frequency. The principle of this calibration consists in locking a reference standard on the UUT. The beat note is then filtered and amplified to a Fast Fourier Transform (FFT) spectrum analyzer. Refer to figure 4 for photographs of the bench.

2. TIME INTERVAL, TIME STABILITY, FREQUENCY, FREQUENCY STABILITY ACCREDITED LABORATORY

In 1871, the town council of Besançon expressed their wish to create a chronometric observatory to support the Besançon clock making industry. After 7 years of debates and negociation, a ministerial decree dated 11th of March, 1878, paved the way for an astronomical, chronometric and weather observatory in Besançon. In 1882, a convention is signed between the town of Besançon, the department of Doubs and the State, by which the city is committed to providing the ground and the buildings, the State to buying the instruments and paying the personnel, and the department to subsidizing a weather facility [2].

The observatory of Besançon has been one of the first accredited French laboratories, as it was recognized by the accreditation authorities in 1971. As written above, it is an LNE associate laboratory and ensures the traceability of the performed calibrations to the national standards; that relates to time interval, time stability, long term frequency and frequency stability measurements.

Three atomic clocks HP5071A-001 (1350172, 1350536 and 1351321) are operated at the observatory in a temperature/humidity controlled room. Those clocks participate to the Atomic International Time (TAI) and to the French Atomic Time TA(F) via 3 GPS receivers (2 Sercel NRT, 1 Ashtech Z12T).

The observatory delivers calibration certificates in time and frequency extending from simple wristwatches to atomic clocks. For frequency measurements, the laboratory can calibrate sources between 0.001 Hz and 1.3 GHz. For time interval measurements, it is in the 0.001s – 10,000 s range. For frequency stability, it is accredited for characterizing Allan variance for integration time in the 1h – 24 h range. The best frequency stability at 24 h that can be characterized under COFRAC accreditation is 2.10^{-14} after 30 days.

The observatory of Besançon is also involved in the development of instrumentation for time and frequency transfer by GPS [3] and realizes also specific developments in the time and frequency domain, for astronomical observatories, such as, for example, the time tagging board for the Auger observatory [4].

3. INTER-LABORATORY TIME AND FREQUENCY TRANSFER BY OPTICAL FIBER

All the available frequencies in the two time and frequency laboratories of Besançon (an hydrogen maser at the LCEP dpt. of FEMTO-ST, cryogenic resonator oscillators at the LPMO dpt of FEMTO-ST and cesium clocks at the LAOB) constitute the state of the art in term of time and frequency references. An optical inter-laboratory link of metrological quality makes it possible to share these frequency references between the three sites. Characterizing the optical link stability and developing a system to correct the phase fluctuations due to optical fibers, we wish to be able to transfer a signal stability of relative frequency of 2.10^{-15} for a few weeks [5].

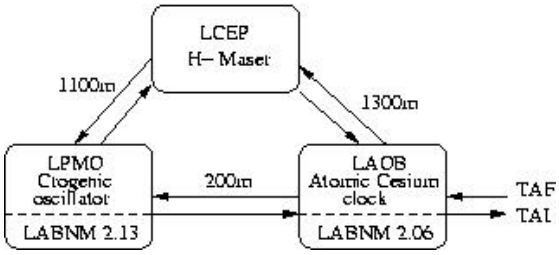


Fig. 5. Optical link between the three laboratories.

For short term frequency stability, ultra high stability signals are now available at the LPMO from cryogenic resonator - oscillator in X band. it transfers better than 10^{-14} short term frequency in the 10-100s range.

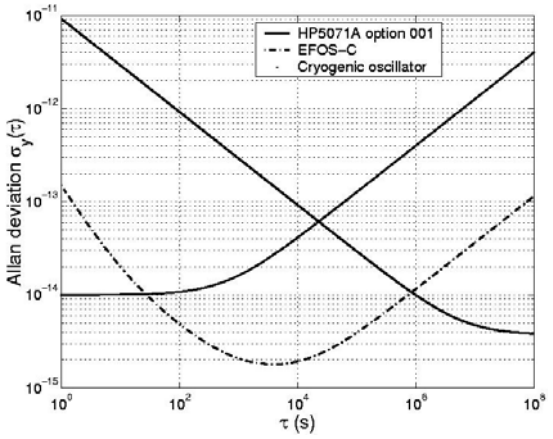


Fig. 6. Comparison of the stability of each reference in terms of Allan deviation $\sigma_Y(\tau)$.

Best middle term stability is obtained using an hydrogen maser at the LCEP around 2.10^{-15} at

1000s. Atomic cesium clocks allows long term stability at the OB around 4.10^{-14} for 10 days. Moreover, the three cesium clocks at OB participate to both French atomic timescale TAF and international atomic timescale TAI.
It is possible to transfer the reference signals without degrading the frequency stability of the three different references.

FURTHER WORKS AND CONCLUSIONS

The inter-laboratory time and frequency transfer by optical fiber will soon provide a local time and frequency scale in the laboratories.
In the next future, ultra high stability signals from cryogenic resonator - oscillator in X band will be able to provide a reference signal that would lead the LPMO dept to considerably increase its accredited capabilities in delivering short term frequency stability calibration certificate.
The observatory of Besançon has developed GPS "melting pot" (all in view) solutions to help the French time and frequency laboratories in ensuring their traceability to the national frequency references. It is also concerned in enhancing its own links to the national references and is currently working on establishing an accurate multi-channel common view GPS link with the primary laboratory (BNM-Syrte).
Both time and frequency accredited laboratories in Besançon plan to found together a new structure in order to constitute a new laboratory that will be able to fully characterize frequency sources from the short to the long term.

REFERENCES

[1] P. Salzenstein, G. Martin, N. Franquet, F. Lardet-Vieudrin and V. Giordano, Proposal for Calibration Accreditation in the Time & Frequency Domain, accepted by the COFRAC - French organism for accreditation, Besançon, France, september 2004

[2] Observatory of Besançon http://www.obs-besancon.fr/

[3] F. Meyer, N. Gautherot, Frequency transfer using low cost GPS receivers, Accessing national frequency references, in these proceedings.

[4] F. Meyer, P. A. Tuckey, F. Vernotte. Time tagging for the Auger observatory, Proceedings of the 16th European Frequency and Time Forum, St Petersburg, Russia, march 2002.

[5] J. Saalaoui, P. Salzenstein, F. Lardet-Vieudrin, S. Galliou, F. Vernotte and V. Giordano, Inter-laboratory time and frequency transfer by optical fiber, Proceedings of the VI International Conference for Young Researchers Wave Electronics and its Applications in Information and telecommunications Systems, St-Petersburg, Russia, 7-11 september 2003, A2 pp 5-8

Accessing national frequency references using low cost GPS receivers: SYREF, example of an integrated solution.

François Meyer, Nicolas Gautherot

Laboratoire d'Astrophysique de l'Observatoire de Besançon - UMR CNRS 6091
41 bis avenue de l'observatoire, BP 1615, 25010 Besançon Cedex, France
E-mail: fmeyer@obs-besancon.fr

1 Abstract

The frequency dissemination of the french frequency reference using the 162 kHz carrier transmitted by TDF's Allouis facility has experienced some problems during last years because of degraded broadcasting conditions, impacting the reliability of the system.
That lead the COFRAC to recommend that the accredited laboratories find another source of traceability to the national frequency reference.
Last generation GPS modules by Motorola can be the core of a cost effective solution to the problem; we present the results of frequency comparisons using such modules and an integrated solution that addresses the particular issue of accessing the national frequency reference TA(F).
This work is supported by LNME (Laboratoire National Métrologie Essais).

2 Introduction

One of the simplest way to access the french national frequency reference realized by TA(F) has long been the use of the carrier frequency of France Inter as broadcast by TDF from the Allouis facility.
This carrier gives access to the TA(F) frequency with a stability around 1.5×10^{-12} for a one day averaging time.
Unfortunately, technical modifications of the broadcast have implemented and tested for the past few years and the quality of reception of the 162 kHz carrier has been undergoing problems ; and the situation is not expected to improve in the future for technical reasons, relating to the progressive migration to digital broadcast.
Among the users of this service (for years and until recently this service has been provided out of any contractual frame), accredited laboratories are in an uncomfortable situation, where the reliability of their only source of traceability to the national reference is put into question.

3 The M12+ oncore

For several years now, the Observatory of Besancon has been operating GPS links to national time and frequency references, for both academic and industrial laboratories.

Successive versions of Motorola's Oncore receivers (VP, UT+) have been used (in common view [1], for the VP, in all-in-view mode for the UT+) [2] [3] [4] [5].

The last generation of these modules is the M12+. See [6] for an evaluation of the performance of the unit.

As for the previous Oncore modules, the M12+ 1 PPS output must be corrected for the negative sawtooth ; this correction is a prediction of the local oscillator drift effect on the next 1 PPS occurrence; this correction is generally performed at the software level during data post processing.

4 The counter, divider and hosting board

We have designed a board including a 1 GHz time interval counter, a frequency divider and hosting the M12+ module. The board is powered through a PCI standard connector but is not connected to the PCI bus.

Also, we have included on the board a hardware solution to correct the M12+ 1PPS output for the negative sawtooth on-the-fly. This corrected output hence becomes a clean realisation of GPS time with a short term jitter of less than 3 ns.

5 The Syref unit

Two prototype boards have been embedded in two 19", 4 units, industrial PC racks (including keyboard, mousepad and VGA LCD screen). One has been set up at BNM-Syrte in Paris and the other one at the Besançon observatory.
These have been used to compare Cs172 at OB with UTC(OP) via GPS in all-in-view (AIV) mode ; in this mode, all visible satellites are tracked by the receiver that computes its best estimate of GPS time, realized by its 1 PPS output. Thus our data are time differences between local 1 PPSs (UTC(OP) and Cs172(OB)) and GPS time as realized by the M12+ GPS module.
The results obtained by the prototypes have been compared to those obtained via common view (CV) comparisons using time dedicated receivers (Sercel NRT2 at OB, AO-TTR5 at Syrte).
Figure 2 shows the results of individual 13 minute tracks. On figure 3, a frequency offset of 8.8ns/d has been removed from both series. Figure 4 shows the difference between both series.
Figure 5 allows for the comparison of the Allan deviation values computed on CV and AIV series; it shows a noise level of the same order, though a bit higher for the M12 units. Those are still able to reveal instabilities of the order of 10^{-13} for a one day averaging time.

6 Unit setup and operations

Basic system setup consists in plugging the GPS antenna, the local reference frequency, booting and setting antenna coordinates.
The counter can be triggered either directly by a local 1 PPS or from a local 5 or 10 MHz frequency through the divider included on the board.
The system follows the BIPM scheduled CV dates to perform 13 minute tracks. As a first monitoring level and in real time (at 13 minute interval) the user can access the plot of its local frequency reference versus GPS Time frequency.
The SYREF unit produces CGGTTS format files where satellite dependent fields are filled with '9'. These data are transfered on a daily basis to the obseratoire de Besançon for processing.
Automated data transfer to the Besançon observatory can be implemented either via a modem or above TCP/IP.
Routine operations at Besançon include daily automatic data collections from the SYREF units, semi-automatic monitoring of those and a weekly computation of the results against UTC(OP). TA(F) is made available at the beginning of each month; at this time, the frequency offset of the local reference versus TA(F) is computed with its uncertainty, and a calibration certificate can be emitted and transmitted by regular mail. That procedure does not require any intervention from the user.

7 Conclusion

Addressing the problem of ensuring the traceability of a local frequency source to a remote frequency reference, we have presented an integrated, fully automated solution built around a Motorola M12+ unit, consisting in an originally designed counter and controlling board embedded in a 4 unit rackable industrial computer. Frequency transfer may can be performed with a stability of 10^{-13} after 1 day.

Acknowledgements

This work is supported by LNME (Laboratoire National Métrologie Essais).

References

[1] Allan, D.W., and C. Thomas, 1994, Metrologia 31, pp. 69-79.

[2] R. Giffard, **Estimation of GPS ionospheric delay usinf L1 code and carrier phase observables** in Proceedings of the 31st Precise Time and Time Interval (PTTI), December 1999, Dana Point, USA, pp. 405-416.

[3] J. Nawrocki, W. Lewandowski and J. Azoubib, **GPS multi-channel common-view time transfer using Motorola Oncore receiver with CCDS standards**in Proceedings of the 12th European Time and Frequency Forum (EFTF), March 1998, Warsawa, Poland, pp. 510-515.

[4] W. Lewandowski, Philippe Moussay, Pascal Guerin, Francois Meyer and Michel Vincent, **Testing Motorola Oncore GPS receiver for time metrology** in Proceedings of the 11th European Time and Frequency Forum (EFTF), March 1997, Neuchatel, Switzerland, pp. 493-497.

[5] P. T. H. Fisk, R. Bruce Washington, Malcolm A. Lawn and Michael J. Wouters, **Time and frequency activities at the CSIRO National Measurement Laboratory, Sydney,**

Australia in Proceedings of the in Proceedings of the 31st Precise Time and Time Interval (PTTI), December 1999, Dana Point, USA, pp. 211-404.

[6] Hambly R.M., Clark, T.A, "Critical evaluation of the motorola m12+..." 2002, Proc. 34th Annual Precise Time and Time Interval (PTTI) Meeting, pp. 109-116



Figure 1: The Syref unit.

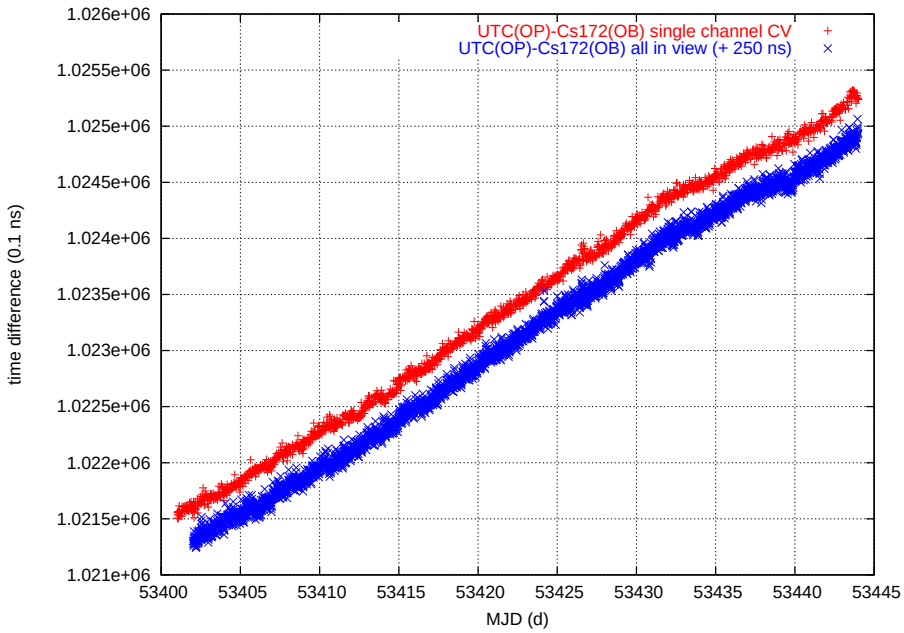


Figure 2: time series as a function of MJD of:

- $+$: *[UTC(OP) - Cs172(OB)]*$_{TTR5/NRT2}$:
  classical common view (hereafter referred to as $CV(OP-OB)$*) using an AO-TTR5 in Syrte and a Sercel NRT2 in Besançon*
- $\times$: *[UTC(OP) - Cs172(OB)]*$_{Syref/Syref}$:
  2 oncore M12+ SYREF units running in all-in-view (hereafter referred to as $AIV_{syref}(OP-OB)$*) mode*

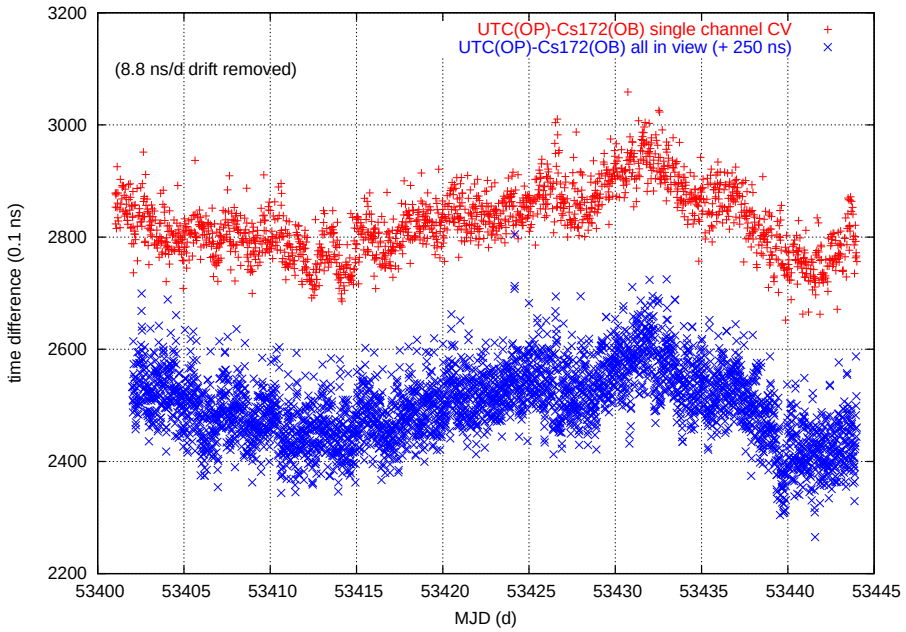


Figure 3: UTC(OP) - Cs172(OB) using both methods after removal of a $8.8ns/d$ drift.

- $+$: *$CV(OP-OB)$ series*
- $\times$: *$AIV_{syref}(OP-OB)$ series*

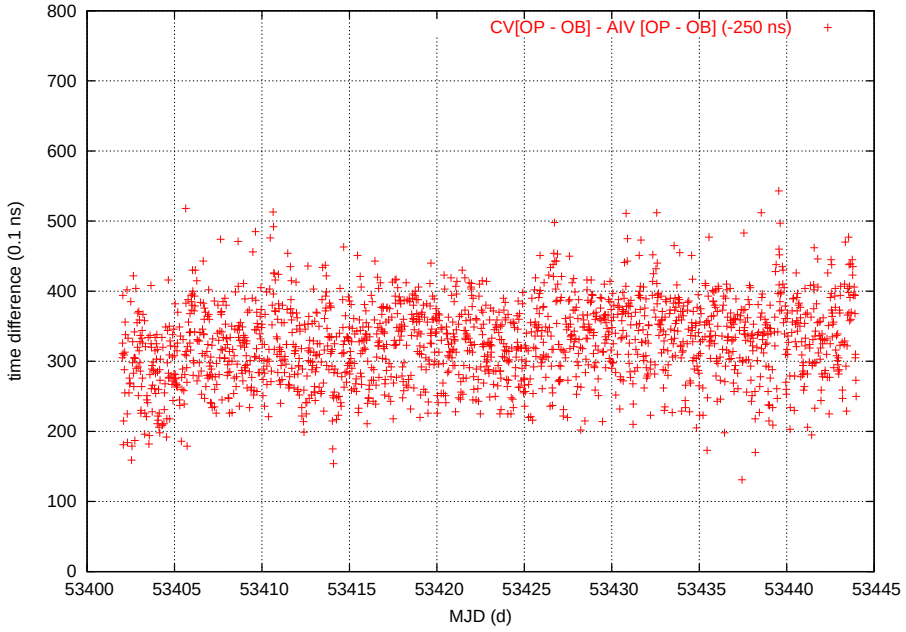


Figure 4: Difference between $CV(OP-OB)$ and AIV_{syref} series.

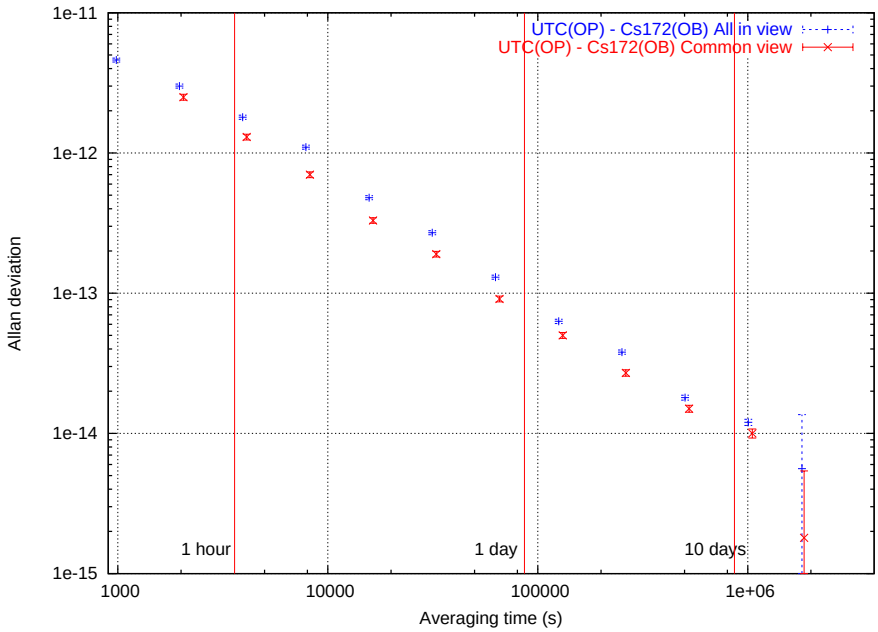


Figure 5: Allan deviation of both $CV(OP - OB)$ (lower) and AIV_{syref} (upper)

New frequency comparisons using GPS carrier phase at CNES

J. Delporte, F. Mercier, Ph. Guillemot

CNES – 18, avenue Edouard Belin – 31401 Toulouse Cedex 9

jerome.delporte@cnes.fr

ABSTRACT

The frequency transfer between distant clocks using GPS carrier phase is widely used throughout the world by many laboratories. However it suffers some limitations like discontinuities at day boundaries in case of using daily batches. A new software has been developed at CNES to overcome this limitation. It uses sp3 and clk IGS products and can compute for instance 10 days batches ensuring clock continuity over that period.

Frequency comparisons based upon GPS carrier phase between high quality frequency references on different baselines (very short, medium, transatlantic) are presented using this new software. These results show that stability in the range of $1\text{-}2.10^{-15}$ on continental baselines and $3\text{-}7.10^{-15}$ on transatlantic baselines can be reached with that technique.

1. INTRODUCTION

To compare distant ground clocks, satellite time or frequency transfers have been used for many years. In particular, the Global Positioning System (GPS) is very convenient and is widely used for instance by timing laboratories.

The classical technique is the so-called common-view where C/A code measurements on one single GPS satellite (chosen using a predefined schedule) are processed to provide the comparison of the two ground clocks. This is equivalent to a single difference code measurement residual at each epoch. Such a method provides a time transfer precision of a few nanoseconds, which is not sufficient for high performance atomic clocks.

High quality geodetic GPS receivers are able to handle several channels (so that several satellites can be tracked in the same time) and two frequencies (so that the ionospheric-free combination can be computed). So, using all the code measurements, an extension of the classical technique is possible in order to have more precise results by averaging on all satellites in common view of both stations and by cancelling out the ionosphere effects. Moreover, such receivers are able to record code and carrier phase observables. It is well known that the latter offers promising perspectives for accurate frequency transfer for integration times between several hours and several days [1-7].

Some receivers are also able to collect code and phase data from geostationnary (GEO) satellites that transmit GPS-like messages. Such satellites are very interesting thanks to their continuous observability. Unfortunately, they transmit on a single frequency (L1), which raises the problem of the ionospheric correction. Moreover, a precise orbit shall also be determined. Several authors [7,8] investigated the advantages/drawbacks of GEO w.r.t. GPS satellites for accurate frequency transfer and compared performances.

The clock solution computed by GPS carrier phase comparison may present discontinuities at day boundaries [5,6]. They are due to the discontinuities of the phase ambiguities at the day boundaries when processing daily batches. A possible solution to generate a continuous time transfer is to use the set of midnight clock parameters to link the computation batches [6].

In two previous papers [3,7], we presented results obtained at CNES using GPS carrier phase frequency transfer with two different GPS receivers and with two different clock resolution algorithms.

The new software, initially developed by CNES Precise Orbit Restitution Service for EGNOS ground stations positioning, is adapted for ground and space clocks identification. The CNES Microwave and Time/Frequency Department uses this software with GPS code and phase data coming from several GPS receivers connected to a Hydrogen Maser.

In this paper, we will first present the different receivers available at the CNES Time/Frequency laboratory. Then, the algorithm and software will be detailed. Last, some results of frequency comparisons on different baselines will be highlighted. The aim of our work is to investigate the performances of GPS frequency transfer and, in particular, to characterise our Hydrogen Maser.

2. HARDWARE DESCRIPTION

2.1. GPS Receivers

Several receivers are used in the CNES Time/Frequency laboratory. The table below summarises the observables of each receiver :

	Observables	GPS/GEO
Ashtech Z12-T	C1, P1, P2, L1, L2	GPS only
NovAtel OEM-3	C1, P2, L1, L2	GPS/GEO
NovAtel MPC	C1, P2, L1, L2	GPS/GEO
Septentrio PolaRX	C1, P1, P2, L1, L2	GPS/GEO

Table 1 : Observables of CNES receivers

They are time receivers, except the NovAtel OEM-3. Every receiver has their own omnidirectional antenna.

2.2. Frequency standard

These receivers are connected to our Hydrogen Maser (EFOS-16 from Neuchâtel Observatory). This active Maser has no automatic cavity tuning (ACT) in order to improve the short term stability. The drawback is that the long term stability is degraded.

In paragraph 4.2.1, we will try to characterize its stability by comparing GPS carrier phase frequency transfer to different local measurements.

3. SOFTWARE DESCRIPTION

The software used in this study was developed by CNES orbit restitution service for the positioning of GPS ground stations for EGNOS project.

This software uses a least squares formulation. It was initially limited to one day measurements [3] (use of IGS daily ephemerides and clocks and RINEX files). A new development has been recently conducted to be able to solve for longer durations, with also the ability to down sample the measurements and to allow phase continuity between daily files. This continuity is ensured by keeping the ambiguities from one day to the next one.

3.1 Architecture

The general architecture is described below. The basic design is to use elementary executables that are launched in a script, that is written by a specific front end, depending on the solution configuration.

There are three main parts :

- pre-processing and partial derivatives computation
- absolute positioning least squares filter : receiver antenna is positioned using IGS GPS ephemerides and clocks as inputs. The receiver clock relative to IGS time is also estimated
- relative positioning least square filter : one station is taken as reference and the second is positioned using single differences measurements. The station receiver clocks difference is also obtained.

The pre-processing and relative positioning parts are described hereafter.

3.2. Pre-processing and partial derivatives

The pre-processing is performed independently for each daily RINEX file, no orbit and station coordinates are needed. The continuity of phase measurements at midnight (change of RINEX file) is taken into account.

The measurement modelisation uses the measured emission time :

$$\begin{aligned} D &= c\left(t_{rec} + h_{rec} - \left(t_{emi} + h_{sat}\right)\right) \\ &= D_{geo}\left(t_{emi}\right) + c\left(h_{rec} - h_{sat}\right) \end{aligned}$$

where D is the modelled pseudo-distance, t_{emi} and t_{rec} are respectively the emission and reception times in GPS time, h_{sat} and h_{rec} are respectively the satellite and receiver clock offsets and $D_{geo}(t_{emi})$ is the geometrical path modelisation (including tropospheric delay).

A first estimation of satellite clock offset $\hat{h}_{sat}$ is given by the IGS files, allowing the computation of t_{emi} with a sufficient precision for the geometrical modelisations (for example, code pseudo-distance $P1$ is defined by $P1 = c\left(t^r_{rec} - t^e_{emi}\right)$ with t^r_{rec} the RINEX time of measurement, and $t^e_{emi} = t_{emi} + \hat{h}_{sat}$ the emitter time tag).

The modelisation takes into account precise coordinates of antenna phase center with solid earth tides effect, radial offset of satellite antenna, tropospheric propagation delays. The phase windup (phase rotation depending on elevation and azimuth evolution over a satellite pass) has not been implemented yet.

A file is generated for each station, containing code and phase measurements, modelisation and partial derivatives. These files can be used for example for direct positioning of a station, or as inputs for the single difference solution (relative positioning of two stations).

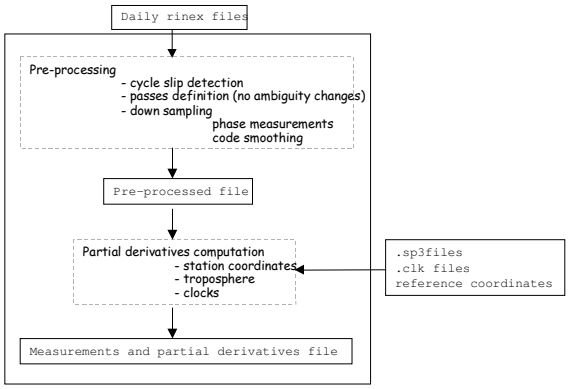


Figure 1 : Pre-processing and partial derivatives

For single station absolute positioning, it is necessary to use a precise value for the satellite clock offset because the modelisation is reduced to $D = D_{geo}(t_{emi}) + c(h_{rec} - \hat{h}_{sat})$. This is why the use of .clk files has been added to increase the precision.

For relative positioning, the use of interpolated satellite clock values is not a problem because the single difference eliminates the term h_{sat} in the expression of D.

3.3. Relative positioning

For relative positioning, the two files produced by the pre-processing are used (station A is the reference, station B is positioned). First, the relative measurements and partial derivatives are constructed (difference of measurements, partial derivatives for station B coordinates, partial derivatives for tropospheric delay of the two stations, partial derivative for difference of receiver clocks). Satellite clocks are eliminated by this process.

The tropospheric delay is modelled for each station as continuous segments of constant duration (typically two hours). Limitation of measurements in elevation can also be applied.

The filter handles two sets of parameters :

- global parameters : coordinates, tropospheric zenith delay, ambiguities
- clock parameters : difference of stations B and A at each epoch.

A first pass is performed for elimination of clock at each epoch (in this case, the normal matrix clock block is diagonal), and reduced normal matrices are produced and solved for global parameters.

A second pass allows the reconstruction of the clock vector : the computed global parameters are replaced in the equations and code and phase residuals are obtained.

The clock difference is the weighted mean of the phase and code residuals. In the post-processing computation, the clock is evaluated using only the phase measurements. Some residuals can again be eliminated for this clock evaluation (elevation limitation, outliers).

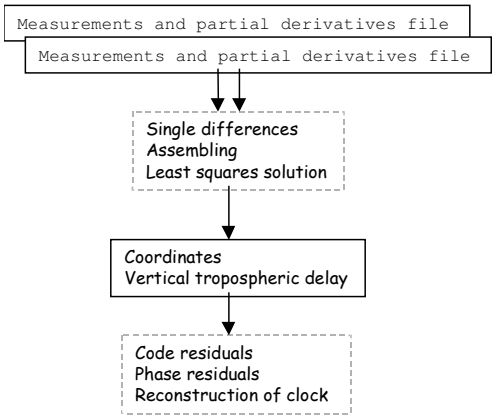


Figure 2 : Relative positioning filter

4. RESULTS

4.1. Very short baseline (CNES/CNES)

This experiment is performed using a common clock (Hydrogen Maser) for each receiver of our laboratory. The two receivers used in this experiment are the NovAtel OEM-3 and the NovAtel MPC. The baseline is about 1 m.

We processed 5 consecutive days (30 seconds RINEX data files). We compared clock solutions provided by our software with different elevation limitations. In Table 2 (column 2) are mentioned the mean number of satellites in common view as a function of the elevation limitation.

In this experiment, the clock solution provided by the software presents an artificial phase drift. This drift has been found to depend on the elevation limitation. The order or magnitude of this drift is some tens of ps/day. Such a phenomenon in common clock configuration has already been observed [2] but its origin is still unclear.

The Allan deviation of the clock solution for 4 different elevation limitations (10 to 40°) is given in Figure 3.

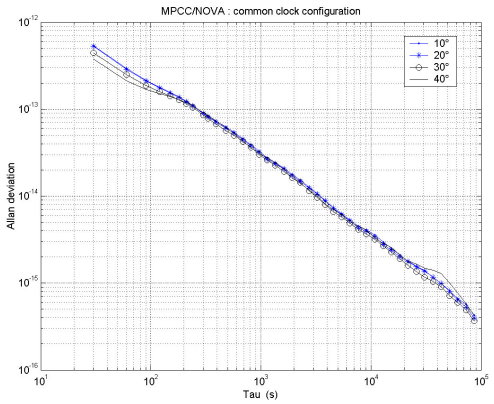


Figure 3 : Frequency stability (Allan deviation) in common clock configuration

There is no significant difference between these curves, except in the very short term.

4.2. Medium baselines (CNES/BRUS and BRUS/WSRT)

The map below presents the location of the stations used in 4.2 :

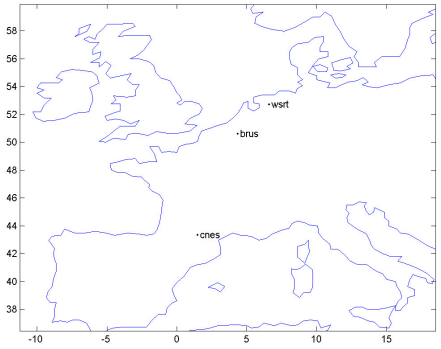


Figure 4 : Map of the GPS stations used in 4.2

4.2.1. CNES/BRUS

In this case, we compare our Hydrogen Maser (with the NovAtel OEM-3 receiver) to another H-Maser located in the IGS station BRUS (Brussels, Belgium). The receiver in this station is an Ashtech Z12-T. The baseline is about 800 km. We processed 8 consecutive days.

The computation of the clock solution has been performed with different elevation limitations (5°, 10°, 20° and 30°). Table 2 provides the mean number of satellites in common view for this experiment (column 3) and for the previous one (column 2).

Elevation limitation	Number of satellites for CNES/CNES baseline	Number of satellites for CNES/BRUS baseline
40°	3.4	3
30°	4.6	4.3
20°	5.4	5.5
10°	5.6	6.5
5°	5.6	6.6

Table 2 : Number of satellites in common view as a function of the elevation limitation on CNES/CNES and CNES/BRUS baseline

At high elevation, as expected, we have more satellites in common view in the very short baseline experiment. But, at low elevation, this is not the case because of the behavior of the receivers (loss of tracking, …).

In the figure below are presented the Allan deviation of the clock solution for 4 different elevation limitations :

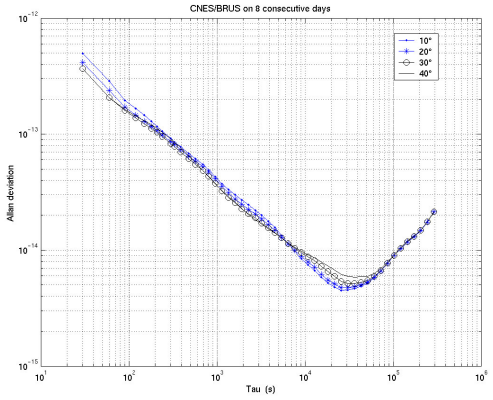


Figure 5 : Frequency stability (Allan deviation) on BRUS/CNES baseline for different elevation limitations

For clarity, we didn't plot the Allan deviation curve for a 5° elevation limitation that provides no improvement w.r.t. 10°. The transfer noise is much lower than what is usually reported for GPS P3 frequency transfer [4] on similar baselines.

With a high elevation limitation, only the "best" satellites are used for the computation, which should provide the best solution. However, the number of satellites used for the computation is then very low. As the clock solution is averaged on fewer satellites, the stability is degraded. So there is a trade-off to perform on that point. As we are more interested in the stability in the range of 10^4 seconds and more, we can derive from this curve that the best compromise for elevation limitation is 10° on such a baseline.

In Figure 6, we compare the stability obtained between our Hydrogen Maser and BRUS (with an elevation limitation of 10°, indicate by dark stars) to other characterisations of our H-Maser :

- its stability measured in 1996 at Neuchâtel Observatory by triangulation with 2 other masers (indicated by a blue line)
- the stability obtained by hourly comparison with a Caesium clock Agilent 5071A-001 (indicated by dark circles)
- the stability obtained by comparison with the cryogenic sapphire oscillator SOPHIE (from University of Western Australia) located 800 m away from our laboratory (indicated by blue points). This comparison is performed at 10 MHz through optical fibers.

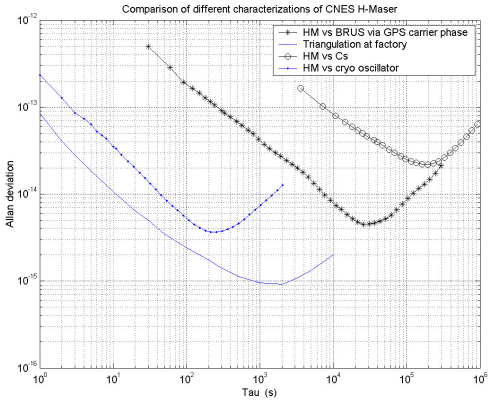


Figure 6 : Different characterisations of CNES Hydrogen Maser

The reference stability is deemed to be the one obtained by triangulation in 1996 at Neuchâtel Observatory, but it can't be obtained again as it would require two other H-Masers.

The comparison of our H-Maser with a cryogenic sapphire oscillator is likely to be degraded by the frequency dividers (100 MHz to 10 MHz). This point shall be investigated further. The drift observed after 300 seconds on this curve is due to the cryogenic sapphire oscillator.

The frequency comparison with BRUS by GPS carrier phase is obviously limited in the short term by the stability of the link. However, if we extend the curve obtained by triangulation, we obtain an excellent agreement with GPS carrier phase.

The local comparison with a Cs clock is of course affected by the stability of the Cs clock in the short and mid term. But, in the long-term, GPS carrier phase is also in very good agreement with it.

Therefore, we can conclude that our H-Maser can be characterised by GPS carrier phase frequency transfer at the level of 4.10^{-15} from 3.10^{4} s on.

4.2.2. BRUS/WSRT

It is also interesting to look into other baselines involving Hydrogen Masers with Automatic Cavity Tuning in order to investigate stability up to one day without being limited by the frequency drift.

We compared IGS stations BRUS and WSRT (Westerbork, The Netherlands) on 7 consecutive days. WSRT receiver is an AOA SWR-12 ACT and its frequency standard is a SigmaTau 2010 Hydrogen Maser. The baseline is about 280 km.

The mean number of satellites in common view is 6.7 for 10°, 5.7 for 20° and 4.4 for 30° elevation limitation. The best results have been obtained with a 20° elevation limitation but the differences between the three solutions are very small. The following figures displays the Allan deviation obtained in that case :

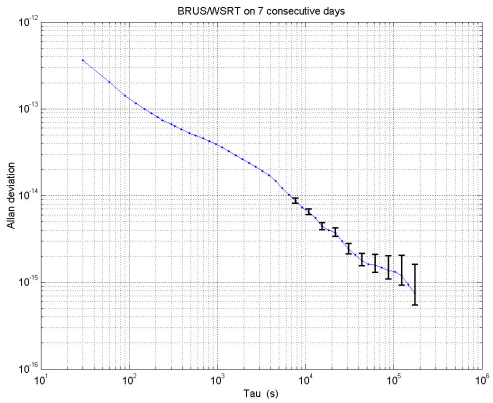


Figure 7 : Frequency stability (Allan deviation) on BRUS/WSRT baseline

This curve clearly demonstrates the capability of this technique to compare distant ground clocks at the level of $1\text{-}2.10^{-15}$ (Allan deviation) on one day integration time on baselines of a few hundreds of km.

4.3. Transatlantic baseline

Figure 8 shows the location of station ALGO (Algonquin, Canada) w.r.t. previously mentioned stations.

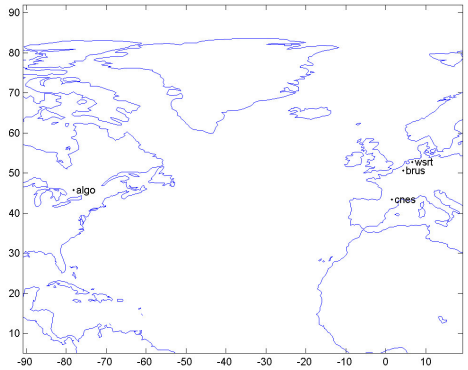


Figure 8 : Map of the GPS stations used in this paper

We compared the IGS stations BRUS and ALGO. ALGO is equipped with an AOA Benchmark ACT receiver connected to a Hydrogen Maser. The baseline is about 5600 km.

Due to data gaps in ALGO station, it has been possible to process a batch of only 4 days of data. The mean number of satellites in common view is 3.6 for 10° and 2.6 for 20° elevation limitation.

The following figure displays the Allan deviation obtained with an elevation limitation of 10° :

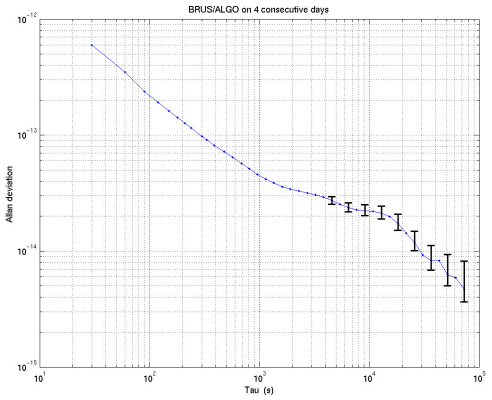


Figure 9 : Frequency stability (Allan deviation) on BRUS/ALGO baseline

On such a transatlantic baseline, stability in the range of 3-7. 10^{-15} on one day can be reached.

5. CONCLUSIONS

We presented in this paper the software developed at CNES Toulouse for GPS carrier phase frequency transfer. This software can handle batches of several days in order to produce a continuous clock solution over that period. For longer periods (more than 2 weeks roughly speaking), a sampling of the data can be easily performed to avoid too long a computation time.

We highlighted results on different baselines and clearly showed that frequency transfer at the level of 1-2.10^{-15} over one day is feasible for baselines of some hundreds of km. For transatlantic baselines, the performance of the frequency transfer is reduced down to 3-7.10^{-15} over one day, which may be due to the number of satellites in common view, the phase windup (that has not been implemented yet) or some lack of modelisation (satellite complete attitude, …).

ACKNOWLEDGEMENT

The authors would like to thank Jean-François Dutrey for his essential help in the local experimental set-up.

REFERENCES

[1] G. Petit, “Processing strategies for accurate frequency comparison using GPS carrier phase”, in Proceedings of EFTF-IEEE FCS Joint Meeting, 1999.

[2] C. Hackman, “New frequency comparisons using GPS carrier-phase time transfer”, in Proceedings of EFTF-IEEE FCS Joint Meeting, 2003.

[3] J. Delporte, “Accurate Frequency Transfer by GPS carrier phase at CNES”, in Proceedings of EFTF, 2002.

[4] P. Defraigne, “Time and Frequency Transfer using GNSS”, in Proceedings of IGS Workshop & Symposium, 2004.

[5] R. Dach, “Status report of the AIUB-METAS geodetic time transfer”, in Proceedings of EFTF, 2004.

[6] J. Ray, “IGS/BIPM pilot project : GPS carrier phase for time/frequency transfer and time scale formation”, Metrologia, vol. 40, pp. 270-288, 2003.

[7] J. Delporte, “Progress in Accurate Frequency Transfer by GPS and GEO carrier phase at CNES”, in Proceedings of EFTF, 2003.

[8] P. Fenton, “The use of the Wide Area Augmentation System (WAAS) as a Time Transfer System”, in Proceedings of ION NTM, 2000.

Session 4B / Compact clocks

CHIP-SCALE ATOMIC FREQUENCY REFERENCES: FABRICATION AND PERFORMANCE

J. Kitching[*], S. Knappe[*], L. Liew[†], J. Moreland[†], H. G. Robinson[*], P. Schwindt[*], V. Shah[*‡], V. Gerginov[**] and L. Hollberg[*]

[*] Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO 80305
[†] Electromagnetics Division, National Institute of Standards and Technology, Boulder, CO 80305
[‡] The University of Colorado, Boulder, CO 80309
[**] The University of Notre Dame, South Bend, IN 46556

Keywords: Atomic, clock, compact, microfabrication, wafer bonding.

Abstract

The physics package for a chip-scale atomic frequency reference was constructed and tested. The device has a total volume of 9.5 mm^3, dissipates 75 mW of electrical power at an ambient temperature of 45 °C and has a short-term fractional frequency instability of $2.4\times10^{-10}/\sqrt{\tau}$. Advanced cell fabrication techniques indicate a long-term instability near 10^{-11} may be feasible.

1 Introduction

While the long-term precision of atomic clocks is unsurpassed, the size and power required to run these devices has prevented their use in a variety of areas, particularly in those applications requiring portability or battery operation. The state of the art in compact commercial atomic frequency references are Rb vapor-cell devices with volumes near 100 cm^3 that operate on a few watts of power and cost about $1000.

Miniaturization based on microelectromechanical systems (MEMS) offers many of the same compelling advantages to atomic frequency references as it does to other large-scale technologies. In addition to small size, a corresponding improvement in the device power dissipation is gained because the heat lost to the environment via the device surface is smaller. MEMS could also enable high-volume, wafer-based production of atomic frequency references, which would substantially reduce cost. Such improvements would make atomic timekeeping useful in a variety of advanced applications where quartz references are now used.

The most recent work at NIST has focused on the development of the physics package. Previously, the smallest physics packages of atomic frequency references have been based on alkali vapor cells fabricated by means of traditional glass-blowing techniques and have had volumes of roughly 1 cm^3 [1]. In addition, the power required to run the physics package has been quite high (near 1 W), partly because of the need to heat the large cell to its operating temperature of a few tens of degrees Kelvin above ambient. Finally, the physics packages are expensive to fabricate in a commercial setting since components must be manufactured and assembled individually into complete devices. The microfabrication approach developed by NIST simultaneously addresses all of these issues and could enable the fabrication of small, low-power and inexpensive physics packages for use in new generations of compact atomic clocks.

2. Physics Package Design and Assembly

At the heart of the NIST microfabricated physics packages are alkali vapor cells fabricated by MEMS processing techniques [2,3]. The cells are made by sandwiching a wafer of silicon, with a hole etched in it, between two glass wafers. Alkali atoms, along with an appropriate buffer gas, are sealed in the volume between the glass wafers. Photographs of an early microfabricated vapor cell are shown in Figure 1 .

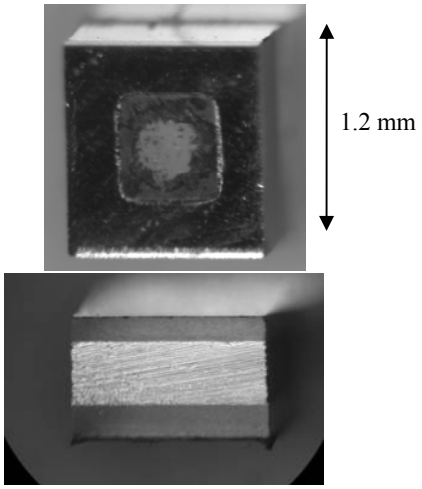


Figure 1. Photographs of a micromachined Cs vapor cell fabricated by anodic bonding.

Based on these cells, we have constructed an integrated physics package based on the design outlined in Ref. [2] and fully described in Ref. [4]. A schematic and photographs of the device are shown in Figure 2. A die containing a vertical-cavity surface-emitting laser (VCSEL) is bonded onto a

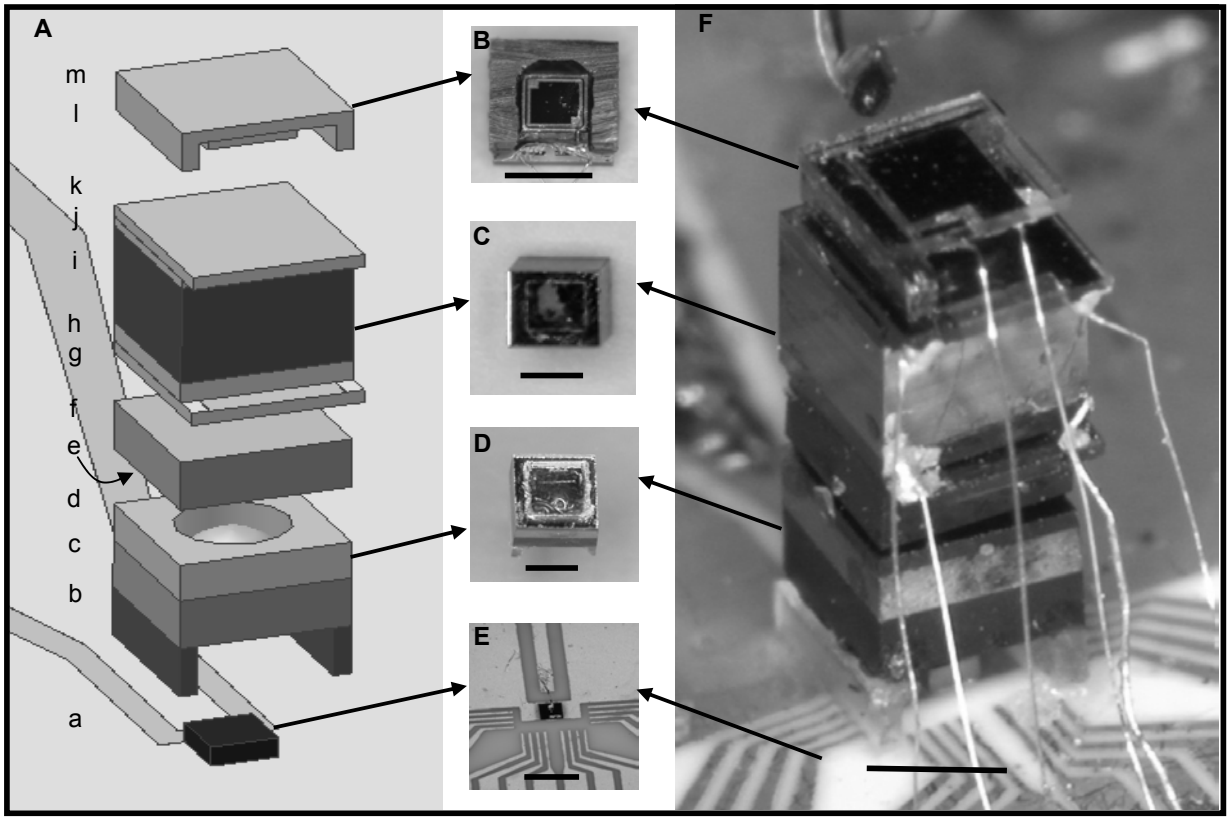


Figure 2. The microfabricated atomic clock physics package based on Cs atoms. (A) Schematic of assembly. Layers from bottom to top: a, Laser and baseplate; b, Glass (500 µm); c, ND filter (500 µm); d, Spacer (375 µm); e, Quartz (70 µm, not shown); f, ND filter (500 µm); g, Glass/ITO (125 µm/30 nm); h, Glass (200 µm); i, Si (1000 µm); j, Glass (200 µm); k, Glass/ITO (125 µm/30 nm); l, Si (375 µm); m, Glass (125µm). Total height, 4.2 mm, width and depth, 1.5 mm. Photographs (B), photodiode assembly, (C), cell assembly, (D), optics assembly and (E), laser assembly and (F), the full atomic frequency reference physics package realized as a microchip. The black lines in the photographs indicate 1 mm.

substrate patterned with gold (Figure 2A, layer a). The VCSEL is used because of its low power requirements (typically < 5 mW for most devices), high modulation efficiency, and availability of single-mode devices at the 852 nm optical transition in Cs and the 795 nm optical transition in Rb.

The light emitted by the VCSEL is conditioned by an optics assembly (Figure 2A, layers b-f) attached to the baseplate. This optics assembly attenuates and collimates the light and changes the light polarization from linear to circular. In this vertically integrated structure, glass spacers (Figure 2A, layer b) support the rest of the clock assembly over the laser and provide thermal isolation between the heated cell and the baseplate. Two neutral-density filters (Figure 2A, layers c and f), cut from a wafer of optically dense glass, attenuate the light power to roughly 10 µW, and the light is collimated by means of a commercially available microrefractive lens fabricated by inkjet deposition of optical epoxy (Figure 2A, layer d). The light polarization is altered with a piece of a quartz wafer, 80 µm thick, with its optical axis oriented at 45º to the direction of the initial polarization of the laser beam (Figure 2A, layer e, not shown).

The cell assembly (Figure 2A, layers g-k) is placed on top of the optics assembly. During operation of the clock, the cell is heated with planar heaters (Figure 2A, layers g and k), fabricated from a thin layer of transparent indium-tin oxide (ITO) deposited on a glass substrate and then wire-bonded to the substrate through gold contact strips. Heaters are placed above and below the cell to provide uniform heating when a current is passed through the ITO film. Because the heat enters the cell though the windows, these parts remain slightly warmer than the cell body, which prevents buildup of an opaque alkali metal film on the windows that might prevent light from entering or exiting the cell.

Finally a photodiode assembly (Figure 2A, layers l-m) is mounted onto the top of the structure to detect the light power transmitted through the cell.

In most cases, the layers are bonded together with optical epoxy. Gold wires of thickness 25 µm are bonded on one end to individual components before they are assembled and the other end is bonded to the baseplate after assembly. We anticipate that with further refinement of the design, all wire-bonding can be done after assembly of the entire structure.

Photographs of the fully assembled device and individual components are shown in Figure 2B-F.

3. Device Characterization

The physics package is operated by first heating the alkali vapor cell to about 80 ºC. At this temperature, vapor-phase Cs atoms are in equilibrium with the liquid phase at a density of about 10^{13} atoms/cm^3, which is sufficient for observation of the atomic resonance. Modulation near 4.6 GHz, one-half of the Cs ground-state hyperfine oscillation frequency, is applied to the laser. As the modulation frequency is swept around the frequency of the atomic resonance, a change in the DC optical power transmitted through the cell occurs, which is detected with the photodiode at the top of the structure. A plot of the photodiode signal as a function of the detuning of twice the modulation signal is shown in Figure 3. The resonance width of 7.1 kHz gives a Q-factor of 1.2×10^6.

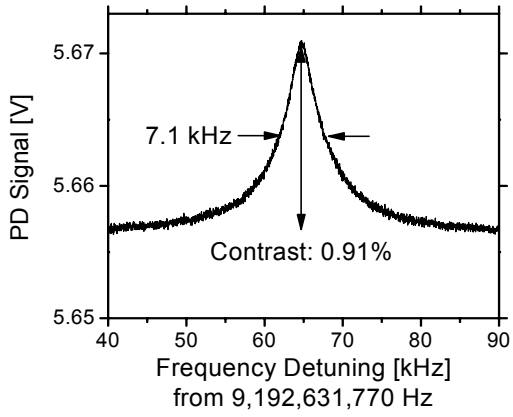


Figure 3. Atomic resonance signal from the Cs microfabricated atomic clock physics package shown in Figure 2.

By modulating the frequency of the local oscillator and using lock-in detection, an error signal can be generated from the resonance in Figure 3. When this error signal is fed back to the frequency-tuning port of the LO, the LO frequency can be stabilized onto the atomic resonance. The stabilized output frequency of the LO is then compared to a secondary, more stable oscillator to measure its frequency as a function of time; a plot of these data is shown in Figure 4, Trace A. When integrated for a time τ, the fractional frequency instability as measured by the Allan deviation is shown in Figure 5, Trace A. An instability of 2.5×10^{-10} at one second of integration is obtained, which averages down to a minimum of 4×10^{-11} near 100 s. A linear drift of -1.6×10^{-8}/day is also observed, which we attribute to chemical reactions occurring in the cell that change the pressure of the buffer gas and therefore also the frequency of the clock.

The short-term instability of the frequency reference is determined by the resonance width (Q-factor) and the noise present at the photodetector. The resonance width is dominated in this particular device by optical power-broadening, with secondary contributions due to collisions of alkali atoms with the walls of the cell and buffer-gas atoms. The noise is dominated by shot-noise in the conversion of optical power into electrons in the photodiode.

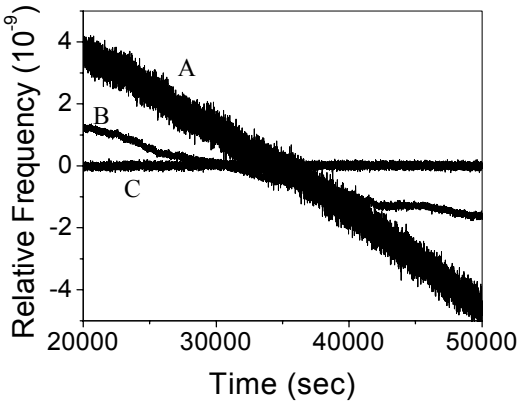


Figure 4. Output frequency of the microfabricated atomic clock physics packages and cells as a function of time. A: Physics package based on Cs (drift -2×10^{-8}/day). B: physics package based on ^{87}Rb (drift -5×10^{-9}/day). and C: cell based on ^{87}Rb fabricated using the beam-filling technique (drift << 10^{-10}/day).

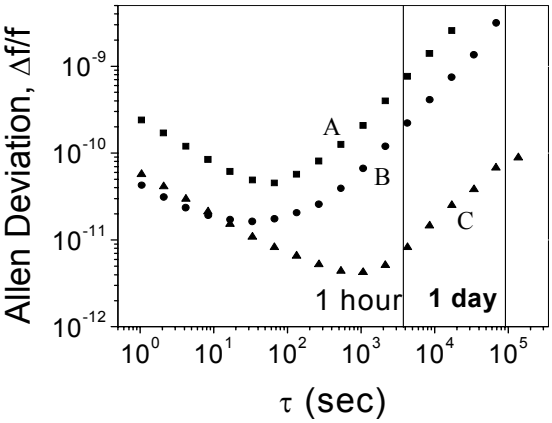


Figure 5. Fractional frequency instability of the microfabricated devices as a function of integration time. Squares indicate stability of Cs physics package, circles indicate stability of ^{87}Rb physics package, and triangles indicate stability of ^{87}Rb cell fabricated with the beam filling technique.

4. Power Dissipation

The electrical power required to run the physics package was 75 mW, and was dominated by the power required to heat the cell above the ambient temperature of 46 ºC set by the baseplate. The high baseplate temperature was required to tune the particular laser used in the device to the correct

wavelength and is not generally necessary; the frequency reference was in fact operated without baseplate heating but the short-term instability was somewhat higher. An advantage obtained through the use of a VCSEL is the extremely low power (< 2 mW) required to both run the laser and to modulate it with amplitude sufficient to excite the CPT resonance.

By modelling the heat flow in the structure, both analytically and with a finite-element computation, the heat loss channels could be roughly identified, as shown in Figure 6. We estimate that 30 mW is lost through the lower spacer unit, and that 24 mW is lost through the six gold wire bonds providing the electrical connections to the baseplate. The remainder is presumably lost through radiation and through conduction to the air surrounding the physics package. Our modelling suggests that the heat lost through these last two channels could be reduced to well below one milliwatt by use of vacuum packaging with a low-emissivity coating on its interior surface [2,5]. The remaining sources of heat loss, conduction through the lower spacer unit and the wire-bonds, can be addressed through more advanced thermal design of the spacer unit itself. For example, the use of a polymer material with a lower thermal conductivity and thin (10 μm) gold traces as part of the electrical path would reduce the heater power to below 12 mW needed to maintain a temperature difference of 50 K between the baseplate and the cell. The two key differences between the design of these chip-scale physics packages and current commercial frequency reference designs are the use of a laser, rather than a lamp, as the optical source and a smaller cell size combined with better thermal isolation.

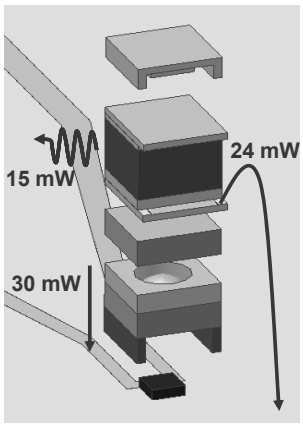


Figure 6. Estimated heat loss sources in the chip-scale atomic clock physics package.

As mentioned above, two other components are required, in addition to the physics package, to form a complete atomic frequency reference. The development of small low-power oscillators at gigahertz-frequencies with sufficient frequency stability to be able to be locked to the atomic resonance is being undertaken by a number of research groups. In addition, a digital control system based on a microprocessor is being developed at NIST and is expected to be able to implement the four feedback loops necessary to run the clock. The anticipated power budget of the complete frequency reference is outlined in Table 1.

Table 1. Anticipated power budget for fully integrated chip-scale atomic clock based on Cs atoms, operating at an ambient temperature near 25 °C.

Component	Power	Confidence
Heating of cell (modeling)	12 mW	moderate
Laser DC power (measured)	4 mW	high
Laser RF power (measured)	70 μW	high
Local oscillator (measured)	3 mW	high
Control circuit (estimate)	20 mW	moderate
B-field, temp. sensor, etc.	< 1 mW	high
Total:	**40 mW**	**moderate**

5. A Microfabricated Physics Package Based on ^{87}Rb

It is known that excitation of CPT resonances with light tuned to the D1 optical transition of alkali atoms results in a signal with an amplitude larger than those excited with light resonant with the D2 line [6]. A chip-scale atomic frequency reference based on D1 excitation is therefore expected to be more precise. Since single-mode VCSELs are not commercially available at 895 nm (the wavelength of the D1 transition of Cs), but are available at 795 nm (the wavelength of the D1 transition in Rb), a microfabricated frequency reference was constructed based on a cell containing ^{87}Rb. This cell was fabricated in a manner very similar to that used for the previous Cs cells. An isotopically pure compound of ^{87}RbCl was reacted with BaN_6 inside the cell preform before the final window was bonded to the structure. The cell was backfilled with a buffer gas of an Ar/Ne mixture instead of one containing N_2. These buffer gases were used instead of N_2 with the goal of reducing the frequency drift, since Ar and Ne are both inert to reaction with Ba, while N_2 is not.

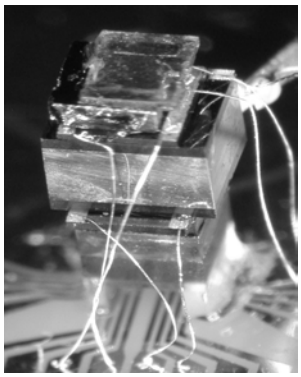


Figure 7. A CSAC physics package based on a cell containing ^{87}Rb and excited with a laser at 795 nm.

A chip-scale atomic clock physics package based on this cell was constructed with a structure identical to that of the Cs device. A photograph of this physics package is shown in Figure 7.

The performance of this physics package was substantially better than that obtained in the Cs/D2 device. The resonance contrast improved to 4.9 %, while the width remained largely unchanged. As a result, the short-term instability was reduced by about a factor of five to 4×10^{-11} at one second of integration. The drift was also reduced to near -5×10^{-9}/day. The frequency as a function of time and corresponding Allan deviation are shown in Figure 4, Trace B and Figure 5, Trace B, respectively. We believe that the residual frequency drift in this physics package is caused by reabsorption of N_2 gas initially created during the chemical reaction to produce the ^{87}Rb and not completely pumped away during the cell filling process.

This device required about 195 mW to operate at an ambient temperature of 22 °C. This larger operating power was a result of the higher operating temperature (120 °C) of the cell, and a lower ambient temperature, as compared to the Cs physics package. The increased cell temperature for the ^{87}Rb device was in part because ^{87}Rb has a lower vapour pressure than Cs at a given temperature. However, it is also likely that optical pumping processes were more important in the ^{87}Rb device than in the Cs device because the former was excited using light resonant with the D1 transition, while the latter was excited on the D2 transition.

6. Microfabricated vapour cells with reduced frequency drift

In order to eliminate the frequency drift due to the presence of Ba inside the vapour cells, a new cell filling technique was developed and is illustrated in Figure 8. In this process, a glass ampoule with a narrow open neck is filled with the mixture of BaN_6 and AlkaliCl instead of the cell itself. The cell preform is first baked at 300 °C to extract water from the glass (Figure 8a). The neck of the ampoule is inserted into the vacuum chamber (Figure 8b), positioned over the cell opening and heated to react the compounds (Figure 8c). Because the vapour pressure of alkali metals is much higher than that of the other materials created in the reaction, a fairly pure beam of alkali atoms emerges from the neck of the ampoule and is deposited into the cell. After the chamber is backfilled with an appropriate buffer gas, a second glass piece is then bonded to the Si, to seal the alkali atoms and buffer gas inside the cell (Figure 8d).

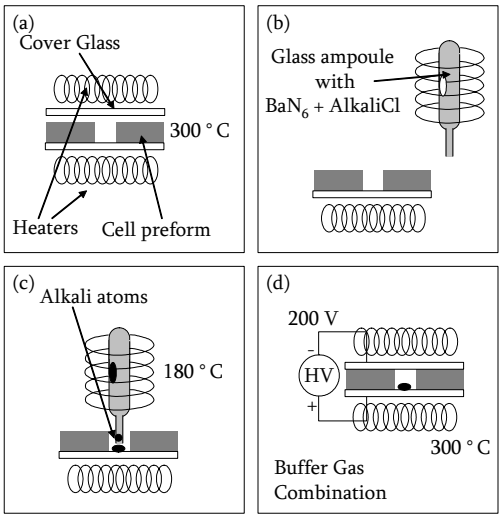


Figure 8. Fabrication process for cells with low frequency drift.

A cell fabricated in this manner showed a significantly reduced frequency drift. The cell was tested in a highly temperature-controlled environment by locking an external oscillator to the CPT resonance excited by the D1 optical transition. The frequency of the locked oscillator is plotted as a function of time in Figure 4, Trace C and the corresponding Allan deviation is shown in Figure 5, Trace C. The frequency drift is reduced to below 5×10^{-11}/day while the short-term frequency stability is maintained.

7. Other applications

While the main goal of our research over the last three years has been to design and build highly miniaturized atomic clocks, the basic technologies developed as part of this research program have significant applications in other areas. One example is highly sensitive and accurate magnetometers [7] based on vapors of alkali atoms. Because atoms have a magnetic moment, different orientations of certain atomic states in an external magnetic field have different internal energies. Since the energy difference is proportional to the magnetic field (at least at low field strengths), measuring this energy difference is equivalent to measuring the strength of the magnetic field. It is therefore possible to determine the magnetic field by comparing the energy difference of magnetically sensitive states to the energy difference of magnetically insensitive states (or some other stable frequency), as shown in Figure 9.

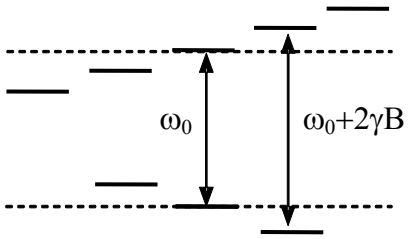


Figure 9. A magnetometer based on coherent population trapping spectroscopy of magnetically sensitive hyperfine transitions.

Operationally, this can be accomplished by simply tuning the frequency of the local oscillator in a chip-scale atomic clock from the magnetically insensitive transition at ω_0 to a magnetically sensitive one at $\omega_0+\gamma B$, where B is the magnitude of the external magnetic field and γ is a constant. This change in LO frequency causes an oscillation between

two atomic states whose energy depends on magnetic field. A magnetometer of this type has been implemented at NIST [8] with a microfabricated structure similar to that in Figure 2 but based on ^{87}Rb atoms, rather than Cs. With this device magnetic fields can be measured with a sensitivity of $40\,pT/\sqrt{Hz}$ in the frequency range of 10 Hz to 100 Hz. The output of the device when subject to an external magnetic field stepped by 2.53 nT every ten seconds is shown in Figure 10.

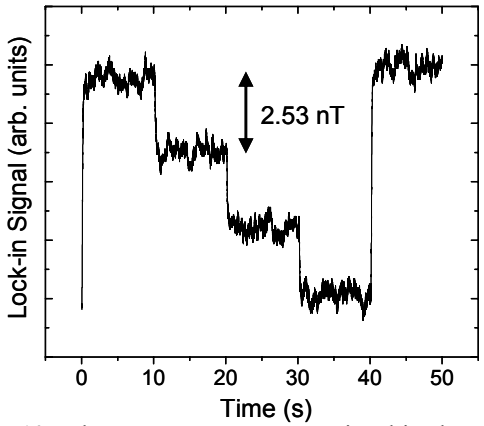


Figure 10. The magnetometer output signal is plotted as a function of time as the magnetic flux is stepped in 10-second intervals. The averaging time is 30 ms.

10. Conclusion

We have described here recent work at NIST to develop highly miniaturized atomic clocks and magnetometers using MEMS microfabrication techniques. The physics packages associated with these devices have volumes in the range of 10 mm^3 and dissipate under 100 mW of power, with significant potential for even lower power dissipation. A critical aspect of the work is the possibility of low-cost, wafer-level fabrication and assembly of devices, which would likely open the door to a wide range of commercial applications for which cost is a considerable factor. However, even a compact atomic clock capable of being operated on batteries for an extended period would be useful in many applications requiring mobility [9,10].

Although the work so far has focussed primarily on the physics package, future work will undoubtedly progress towards integration of this subsystem with a compact local oscillator and control electronics. Development of these other critical subsystems will pave the way toward commercial realization of chip-scale atomic frequency references [11]

Acknowledgements

This work was supported by NIST and the Defence Advanced Research Projects Agency (DARPA). This work is a contribution of NIST, an agency of the US government, and is not subject to copyright. Portions of this manuscript have been published previously in Refs. [3], [4], [5], [8] and [12]

[1] P.J. Chantry, I. Liberman, W.R. Verbanets, C. F. Petronio, R. L. Cather, and W. D. Partlow, "Miniature laser-pumped cesium cell atomic clock oscillator," *Proc. IEEE Int. Frequency Control Symp.,* pp. 1002-1010, 1996.

[2] J. Kitching, S. Knappe, L. Liew, J. Moreland, H. G. Robinson, P. D. D. Schwindt, V. Shah and L. Hollberg, "Microfabricated atomic frequency references," *Proc. 18th Eur. Freq. Time Forum*, April 5-6, Guildford, UK, 2004.

[3] L-A. Liew, S. Knappe, J. Moreland, H. Robinson, L. Hollberg and J. Kitching, "Microfabricated alkali atom vapor cells," *Appl. Phys. Lett*., **84**, 2694-2696, 2004.

[4] S. Knappe, L. Liew, V. Shah, P. Schwindt, J. Moreland, L. Hollberg, and J. Kitching, "A micromachined atomic clock," *Appl. Phys. Lett.*, **85**, 1460-1462, 2004.

[5] J. Kitching, S. Knappe, P. D. D. Schwindt, V. Shah, L. Hollberg, L. Liew and J. Moreland, "Power dissipation in a vertically-integrated chip-scale atomic clock," *Proc. IEEE Int. Freq. Cont. Symp.*, Montreal, Canada, Aug. 24-27, 2004.

[6] M. Stahler, R. Wynands, S. Knappe, J. Kitching, L. Hollberg, A. Taichenachev, and V. Yudin, "Coherent population trapping resonances in thermal Rb-85 vapor: D-1 versus D-2 line excitation," *Opt. Lett.*, **27**, 1472-1474, 2002

[7] H. Dehmelt, "Modulation of a light beam by precessing absorbing atoms," *Phys. Rev.,* vol. 105, pp. 1924-1925, 1957.

[8] P. D. D. Schwindt, S. Knappe, V. Shah, L. Hollberg, J. Kitching, L. Liew and J. Moreland, "A chip-scale atomic magnetometer," *Appl. Phys. Lett.,* **85**, 6409, 2004.

[9] J. Vig, "Military applications of high-accuracy frequency standards and clocks," *IEEE Trans. Ultrason. Ferroelec. Freq. Cont.*, **40**, 522-527, 1993

[10] H. Fruehoff, "Fast "direct-P(Y)" GPS signal acquisition using a special portable clock," *Proc 33rd Ann. Precise Time and Time Interval (PTTI) Meeting*, 359-369, 2001.

[11] R. Lutwak, J. Deng, W. Riley, M. Varghese, J. Leblanc, G. Tepolt, M. Mescher, D. K. Serkland, K. M. Geib and G. M. Peake, "The chip-scale atomic clock – low-power physics package," *Proc. 36th Ann. Precise Time and Time Interval (PTTI) Meeting*, Washington, DC, Dec. 7-9, 2004.

[12] J. Kitching, S. Knappe, L. Liew, P. D. D. Schwindt, V. Shah, J. Moreland and L. Hollberg, "Microfabricated atomic clocks," *Proc. 18th Int. IEEE Conf. Micro Elec. Mech. Sys.*, Jan 30- Feb. 3, 2005, Miami, FL, 2005.

A FEASIBILITY STUDY OF AN OPTICALLY-PUMPED CESIUM BEAM RESONATOR FOR SPACE APPLICATIONS

P. Berthoud, R. Ruffieux, and P. Thomann

Observatoire de Neuchâtel, Rue de l'Observatoire 58, CH-2000 Neuchâtel, Switzerland
Email : Patrick.Berthoud@ne.ch

ABSTRACT

Observatoire de Neuchâtel (ON) has recently started breadboarding activities for an optically-pumped cesium beam frequency standard in the frame of an ESA-ARTES 5 project. The optical setup is based on the simplest optical scheme, which uses a single optical frequency for both the atomic preparation and detection. Preliminary signal-to-noise measurements on the atomic resonator yield an expected frequency stability about 6.5×10^{-12} @1s. The two major noise contributions have been identified as the background light shot-noise and as the residual laser frequency noise converted in amplitude noise by the atomic frequency discriminator.

1. INTRODUCTION

Atomic clocks play a major role in various space applications requiring a stable and/or accurate timing reference source, such as navigation systems, telecommunications, long-term missions, and scientific experiments. To a large extent, the performance of the clock is the major limiting factor of the system capabilities. Even if atomic clock technologies are well mastered for ground applications, they have to be adapted to the space environment in order to exhibit similar performances, but with a strong reduction in terms of mass, volume and power consumption.

The two atomic clocks foreseen to take place onboard the first generation of Galileo satellites are respectively the Space Passive Hydrogen Maser (SPHM) and the Rubidium Atomic Frequency Standard (RAFS). While the RAFS is a very compact clock (2.4 liters, 3.4 kg) [1], the SPHM is bigger (25 liters, 15 kg) [2], but exhibits an 5-fold improvement of the long-term frequency stability ($\sigma_y < 10^{-14}$ for $\tau > 10^4$ s). These two frequency standards being operated in vapor cell conditions, the influence of the environment is dramatic (frequency temperature coefficient). To overcome the long-term frequency instability of these standards, the atomic-beam frequency standard is an elegant alternative presently in wide use for GPS. Compared to a magnetically-deflected atomic-beam device, a laser-pumped device has a better short-term stability, owing to the fact that it makes use of the full atomic velocity distribution. It can even compete with the SPHM in terms of frequency stability. Moreover, due to its inherently simple design, the manufacturing and the reliability of this frequency standard can be strongly improved.

Optical pumping of atomic beams has been extensively studied in the past. Two optical configurations have been tested on similar compact atomic resonators, but unfortunately never extensively compared (in particular line quality factor and atomic flux). These two optical configurations are namely the "1-frequency" scheme and the "2-frequency" scheme:

- The 1-frequency scheme uses the same optical frequency for both the atomic beam preparation and detection, respectively before and after the microwave interaction. Its low photon fluorescence rate (≤ 4 photons/atom) requires a very efficient optical collection efficiency (>40%), but relaxes the laser linewidth requirement. In addition, only one laser diode is required (no additional active optical part, such as a second laser diode or an acousto-optic modulator (AOM)). This simplest optical scheme has demonstrated a short term frequency stability of $\sigma_y = 4 \times 10^{-12}\, \tau^{-1/2}$ using a compact laboratory atomic resonator and a single solitary laser diode (852 nm, 25 MHz linewidth) [3].

- The 2-frequency scheme takes advantage of the high fluorescence rate of a cycling optical transition (≈ 70 photons/atom) for the detection process, while a pumping transition is used for the preparation. Such a scheme relaxes the flurorescence light collection efficiency requirement (>2%), but calls for a narrower laser source, e.g. an extended cavity laser diode (ECLD). In addition, two active optical parts are required, either two laser diodes, or one laser diode and an AOM. With such a 2-frequency scheme, the frequency stability of compact atomic resonators has been improved up to $\sigma_y = 3.8 \times 10^{-12}\, \tau^{-1/2}$ using a commercial ECDL and an AOM [4], or even $\sigma_y = 1.5 \times 10^{-12}\, \tau^{-1/2}$ using a laboratory ECDL and an AOM [5]. An industrial prototype using this optical scheme has also demonstrated a frequency stability of $\sigma_y = 3 \times 10^{-12}\, \tau^{-1/2}$ [6].

Observatoire de Neuchâtel (ON) has recently started breadboarding activities for such an optically-pumped cesium beam frequency standard in the frame of an ESA-ARTES 5 project. Our goal is to demonstrate a frequency stability approaching $\sigma_y = 1 \times 10^{-12}\, \tau^{-1/2}$ with the simplest optical scheme ("1-frequency scheme"). The design of the atomic resonator should be as simple and rugged for best reliability. This development constitutes a fundamental step in the general effort to reduce the mass of the on-board clocks, while keeping or even improving their performances. It is taking advantage of previous activities at ON in the late '80s [7] and of the latest progress in the field of tunable and narrow-band lasers.

We report in these proceedings on preliminary developments on the atomic resonator. First we briefly describe the experimental set-up (§2), then we discuss preliminary results concerning the atomic beam preparation and detection performances, and Ramsey fringe signal-to-noise ratio (§3), and finally we conclude with the next steps to be taken to improve its performance (§4).

2. EXPERIMENTAL SETUP

The architecture of the frequency standard is shown on Fig. 1: the Physics Package (PP) is composed of the Atomic Resonator (AR), while the Opto-Electronics Package (OEP) contains the Optics and Laser modules (OL), the Optics and Laser Control module (OLC) and the Atomic Resonator Control module (ARC).

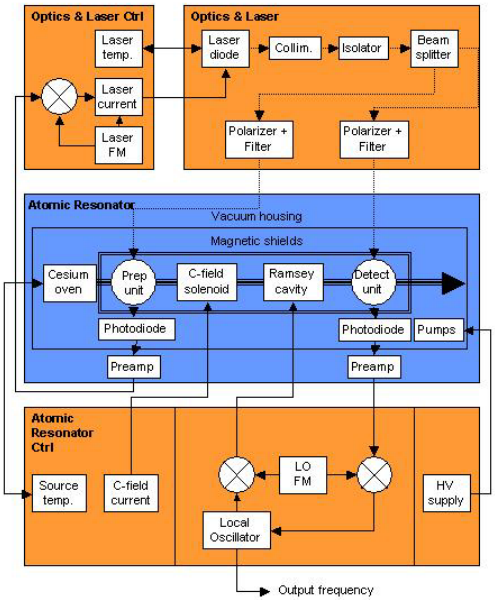


Fig. 1:Block diagram of the optically-pumped cesium beam frequency standard. The blue module is the Physics Package and the orange modules are part of the Opto-Electronics Package.

The AR is fairly standard for an optically-pumped beam resonator. The atomic beam is produced by an oven containing 2g of Cesium, fitted with a multichannel collimator. The outcoming atomic beam crosses a first laser beam in the Preparation Unit (PU), which achieves the required ground state population inversion. Then, the atomic beam crosses a short Ramsey cavity (12 cm), where the microwave resonance takes place. Finally, the atomic beam crosses the second laser beam (from the same laser) in the Detection Unit (DU), in which the final atomic state is optically probed. The two optical units and the microwave cavity are surrounded by a solenoid coil which produces a uniform magnetic C-field (52 mG). The AR is maintained under ultra high vacuum by Cesium getters (graphite) and by a commercial ion pump. In each optical unit, one Si photodiode (25 mm^2) placed under vacuum converts the fluorescence light into photo-electrons. It is followed by an external, low-noise current-to-voltage converter.

The OL has been assembled to operate the frequency standard with the 1-frequency scheme. However, it could be easily adapted for the 2-frequency scheme, by keeping the optical setup for the preparation beam and using a a second laser for the detection beam. The optical beam propagation uses laboratory and commercial free space optics. The single 852nm laser source is a Distributed Feed-Back diode (DFB), with an output power >15mW and a spectral linewidth <2MHz. The diverging laser beam is first collimated, then isolated (40dB), and finally split into two beams. Neutral density filters and polarisers properly prepare their respective power and polarisation before the AR windows.

The ARC is composed of the necessary power supplies and driving electronics for the Cs oven temperature regulation, for the C-field solenoid current provision and for the ion pump high voltage supply. The microwave frequency chain (local oscillator, frequency multiplier and phase modulator and demodulator) are commercial devices. Presently this RF module is however operated in open loop mode.

The OLC uses the Cesium beam inside the AR as the frequency discriminator (DU photo-detector). The laser frequency is modulated by the laser injection current at 1kHz and peak-locked on the pumping transition F=4 $\leftrightarrow$ F'=4 of the Cesium D_2 line by synchronous detection. The laser temperature is regulated with a thermo-electric heater/cooler.

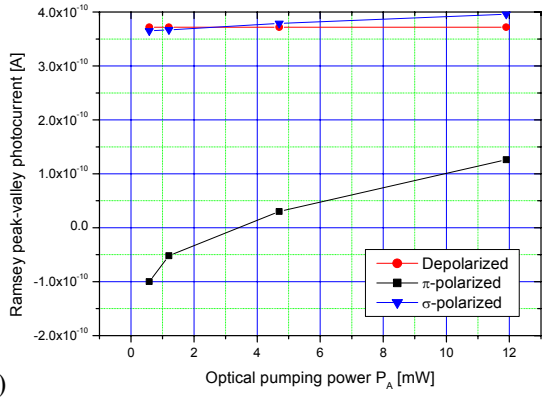
3. PRELIMINARY RESULTS

Before investigating the atomic beam with laser beams, we have calibrated its intensity as a function of the Cs oven temperature. For this purpose we have used an

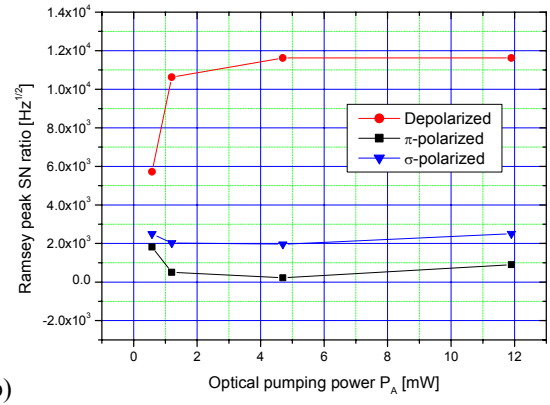
independent atomic ionisation detector with an open surface of 5mm^2 at a distance of 275 mm from the source (solid angle = 6.6x10^{-5} sr). The oven temperature is adjusted for a calibrated flux density of 2 10^9 atoms s^{-1} mm^{-2} (all atomic states). We have verified that DC ion current is characterised by pure shot-noise.

The first optical process we have investigated concerns the atomic beam preparation. As was pointed out by Dimarcq et al., the efficiency of the atomic preparation may critically govern the clock signal-to-noise ratio (SNR) [8]. As we are using an optical pumping transition that creates dark states (Cs D_2: 44'), a significant fraction of the incoming atomic beam can remain trapped in these states. The consequences are firstly a reduction of the useful atomic signal, but secondly a significant noise increase. Various technical solutions have been proposed to overcome this trapping process, either by using an intense magnetic field in each optical unit [9], or by creating an optical polarisation gradient transverse to the atomic beam direction [10]. Although both methods strongly reduce the fraction of trapped atoms, they add a significant complexity to the AR design.

The new optical scheme we are proposing to reduce this dark state trapping is technically simpler than the previous ones and as efficient. It relies on an optical beam polarisation gradient (as in [10]), which is however parallel to the atomic beam direction, and does not require a retro-reflected laser beam (patent pending). Practically, the depolarisation of the polarised laser beam is achieved by simply inserting a birefringent wedged plate before the AR window and by properly orienting its principal axes with respect to the laser beam polarisation axis. We have measured the atomic beam pumping efficiency as a function of the optical pumping power in the PU for three different optical polarisation states, namely π-polarised, σ-polarised and depolarised. Fig. 2a shows that the Ramsey central fringe peak-valley photocurrent is equal for the depolarised and the σ-polarised beams. On the contrary, a π-polarised beam strongly reduces the Ramsey signal and may even suppress it when the hyperfine pumping process perfectly balances the Zeeman pumping process. Although the photocurrent is not increased for a depolarised beam, its current noise density is strongly reduced with respect to linearly polarised beams of either polarisation, as demonstrated by the SNR (Fig. 2b). The SNR saturation behaviour is an indirect proof of an efficient population inversion. By using such a depolarised beam with a typical optical power of 5 mW, the Ramsey central fringe SNR reaches saturation at SNR=11'000 Hz$^{1/2}$. For these measurements the magnetic field amplitude in the PU was set to 52 mG (C-field amplitude), and a similar behaviour is observed even at zero magnetic field.



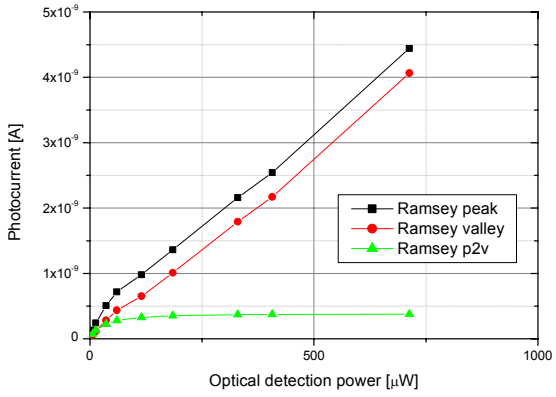
a)



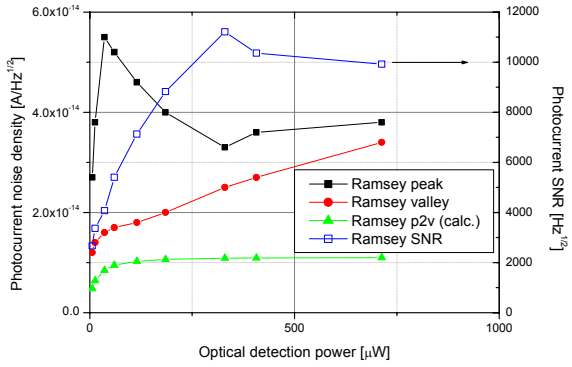
b)

Fig. 2: Atomic beam pumping process: a) Ramsey central fringe peak-valley photocurrent; b) Ramsey central fringe SNR.

The second optical process we have investigated concerns the optical detection. The output signal from the DU photo-detector is used for frequency locking of both the laser and the RF local oscillator. The highest SNR value is aimed at in order to increase the frequency stability of the standard. Fig 3a shows the dependence of the DU photocurrent as a function of the detection beam power after the Ramsey interaction. As we also probe the atomic beam with a pumping transition (same as for the atomic preparation), the Ramsey central fringe peak-valley photocurrent already saturates at around 50 μW. By increasing the detection power, only the background photocurrent rises linearly, which eventually reduces the overall SNR.



a)



b)

Fig. 3: Atomic beam detection process: a) Ramsey central fringe photocurrents; b) Ramsey central fringe noise contributions and SNR.

However, Fig 3b shows that for this low and sufficient detection power of 50 μW, the total current noise density significantly exceeds the RMS sum of the background noise (Ramsey valley) and the effective atomic signal (Ramsey peak-to-valley). At higher detection laser power, the overall noise density even decreases, although the atomic shot-noise is unchanged and the background noise increases. We interpret this behaviour as a reduction of the residual laser frequency noise converted into fluorescent intensity noise by the atomic beam frequency discriminator. Indeed, we observe that increasing the optical power also broadens the atomic frequency response and reduces the frequency discriminator slope. For the present value of laser spectral linewidth, and frequency lock loop parameters, the optimum detection power is about 350 μW, which yields a Ramsey SNR=11'000 $Hz^{1/2}$.

Finally, Fig. 4 shows Ramsey fringes recorded at the optimum functional parameters: the preparation beam is depolarised and has an optical power of 350 μW; the magnetic C-field amplitude is 52 mG; RF power level is adjusted for maximum signal. The AR resonator performances are: an atomic linewidth of 940 Hz yielding an atomic quality factor of 10^7, a Ramsey central fringe peak-to-valley photo-current of 400 pA, a background photo-current of 2 nA, a noise current density at resonance of 37 $fA/Hz^{1/2}$ yielding a SNR=10'800 $Hz^{1/2}$ at resonance. For the foreseen slow RF square frequency modulation and given the above performances, we estimate a short-term frequency stability of $6.5x10^{-12}$ @1s. According to our theoretical estimates, this performance level falls short of the best attainable frequency stability at the present flux density (factor 3 for this "1-frequency" scheme).

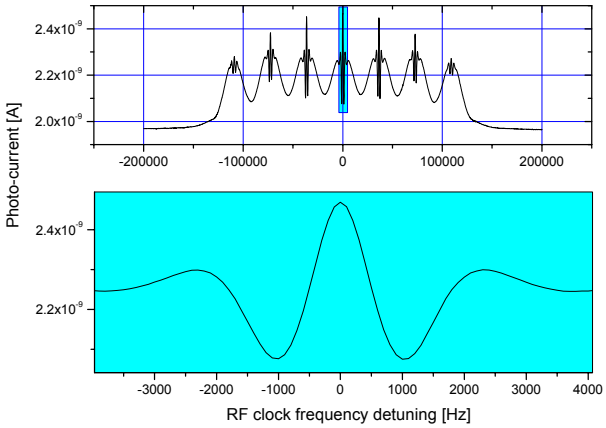


Fig. 4: Ramsey fringes recorded for optical preparation and detection beams using the same pumping hyperfine transition Cs D_2:44'.

4. NEXT STEPS

Presently, the two main limiting factors for the clock performance have been identified as the residual laser frequency noise converted to amplitude noise by the atomic beam frequency discriminator, and by the background noise at high detection laser power level. The residual frequency noise could be reduced in different or complementary ways: by using a narrower linewidth laser or by improving the laser frequency lock loop. We recall that the present laser modulation frequency is only 1 kHz, which typically provides a servo loop bandwidth of about 100 Hz. By increasing the modulation frequency, we should increase accordingly the loop bandwidth and therefore reduce the 1/f contribution to laser frequency noise. In addition, a modulation period smaller than the atomic transit time may improve the pumping efficiency and reduce the contribution of unpumped atoms to the measured noise. By having a less noisy detection laser beam, we should be able to avoid a highly saturated optical detection process. This will allow a strong reduction of the detection power, which will directly reduce the background light level and its associated shot-noise.

5. CONCLUSION

We have reported on the experimental setup of a breadboard optically-pumped cesium beam frequency standard. Its concept relies on the simplest optical scheme, in which the two optical beams for both the preparation and the detection processes are issued from the same laser diode which is frequency locked on a pumping hyperfine transition (Cs D_2: 44'). A distinctive feature of this breadboard resonator is the provision of an independent and absolute calibration of the atomic flux density.

Preliminary results have demonstrated the efficiency of a new original optical pumping setup using depolarised laser beams, which prevent dark state accumulation even in a weak magnetic field and without the need for a retroreflected beam. Based on measured signal-to-noise ratio of the Ramsey fringe, we expect a frequency stability for the atomic resonator of about $\sigma_y = 6.5 \times 10^{-12}\ \tau^{-1/2}$. This is within a factor of three of the theoretical stability achievable with the present atomic flux density, assuming a noiseless laser and negligible background light.

Presently, the two major limiting factors are the residual laser frequency noise and/or the background light shot-noise. An improvement by a factor of three of either noise contribution is sufficient to reach the atomic shot noise limit at the present atomic flux density, and our first frequency stability goal of $\sigma_y = 3 \times 10^{-12}\ \tau^{-1/2}$.

6. ACKNOWLEDGEMENTS

The ON team would like to thank ESA for supporting the development of this "Ultra Stable Atomic Beam Clock for Telecom Space Applications and Long-Term Space Missions" (ESA contract 18040/04/NL/DS).

7. REFERENCES

1. Droz, F., et al. *On-board Galileo RAFS current status and performance*. in *2003 Frequency Control Symposium jointly with the 17th European Time and Frequency Forum*. 2003. Tampa, FL, USA.
2. Berthoud, P., et al. *The engineering model of the space passive hydrogen maser for the european global navigation satellite system Galileo*. in *2003 Frequency Control Symposium jointly with the 17th European Time and Frequency Forum*. 2003. Tampa, FL, USA.
3. Petit, P., et al. *Performance of a 2dm^3 optically pumped cesium beam tube: a progress report*. in *8th European and Time Frequency Forum*. 1994. Munich, Germany.
4. Lutwak, R., et al. *Optically-pumped cesium beam frequency standard for GPS-III*. in *33rd Annual Precise Time and Time Interval (PTTI) Meeting*. 2001. Long Beach, CA, USA.
5. Guérandel, S., et al. *Operation of a compact cesium beam frequency standard: how to improve the stability towards $10^{-12}\ \tau^{-1/2}$ level*. in *2002 IEEE International Frequency Control Symposium and PDA Exhibition*. 2002.
6. Sallot, C., et al. *$3x10^{-12}\ \tau^{-1/2}$ on industrial prototype optically pumped cesium beam frequency standard*. in *2003 Frequency Control Symposium jointly with the 17th European Time and Frequency Forum*. 2003. Tampa, FL, USA.
7. Thomann, P. and F. Hadorn. *Short-term stability of a commercial optically pumped Cs tube*. in *1989 European Time and Frequency Forum*. 1989. Besançon, France.
8. Dimarcq, N., et al., *Analysis of the noise sources in an optically pumped cesium beam resonator.* IEEE Transactions on Instrumentation and Measurement, 1993. **42**(2): p. 115-120.
9. Giordano, V., et al., *New design for a high performance optically pumped cesium beam tube.* IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 1991. **38**(4): p. 350-357.
10. Shirley, J.H. and R.E. Drullinger. *Zeeman coherences and dark states in optically pumped cesium frequency standards*. in *1994 IEEE Prec. Electromagn. Meas.* 1994. Boulder, Co, USA.

Time sequence optimization and optical detuning effects in a Raman-Ramsey frequency standard

T. Zanon, S. Guérandel, E. de Clercq, D. Holleville, N. Dimarcq and A. Clairon
SYRTE – CNRS UMR8630 – Observatoire de Paris
61 avenue de l'Observatoire, 75014 Paris, France

***Abstract*— Efficient dark resonances are obtained with orthogonal linear laser polarizations. Combined with a pulse sequence interrogation, clock signals can be obtained with high contrast and narrow linewidth. Experimental parameters optimizing the pulse sequence are studied. Frequency shifts inherent to the interrogation method are calculated and the conditions for shift cancelation are discussed.**

I. Introduction

Coherent Population Trapping (CPT) or resonant Raman interaction is a well known effect already used to produce narrow atomic resonances below 100 Hz in a three-level atomic system with buffer gas. The interrogation method usually involves continuous excitation of two ground states defining the Raman transition coupled to a common excited state by two coherent laser fields (Fig. 1). Interference between

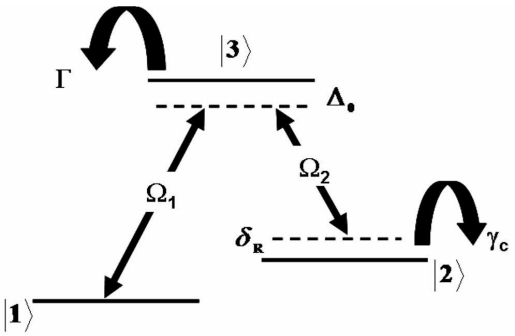


Fig. 1. Three level system under two coherent laser fields excitation for CPT with Ω_1 and Ω_2 Rabi angular frequencies. Δ_0 is the common optical detuning and δ_R is the Raman detuning. Γ and γ_C are respectively spontaneous emission and hyperfine coherence relaxation terms.

amplitudes of transition in the two optical pumping channels cancels both fluorescence and absorption when the Raman resonance condition $\delta_R = 0$ is fulfilled. In the first part of this paper, important experimental details and results are recalled from previous work, in particular the CPT pulsed interrogation method. Then, an optimized interrogation sequence is presented under typical working conditions. Finally, the frequency shifts due to laser detuning are evaluated by changing most of the experimental parameters of interest.

II. Experiment

Our experimental setup has been presented previously [1]. Only important characteristics follow. Two monochromatic phase-coherent laser beams tuned to Cs-D1 pumping transitions provide independent adjustment of frequency, intensity and polarization for each beam. Both beams are superposed on a polarizing beam splitter cube. One output provides the reference beat note at 9.2 GHz, and the other is sent to the atomic cell through an acousto-optical modulator. The cell at room temperature is filled with cesium vapor with 23 torr N_2 buffer gas. A 20 μT longitudinal static magnetic field is created by a surrounding solenoid and the whole physics package is magnetically shielded. With orthogonal linear polarizations, the laser beams parallel to the magnetic field make a double-lambda or lin-per-lin optical configuration. The advantages are (1) reduce the optical pumping into the end Zeeman states and (2) increase the contrast of the central fringe. An experimental CPT continuous signal is plotted in Fig. 2, fitted with a Lorentz absorption shape as the microwave frequency is scanned across resonance.

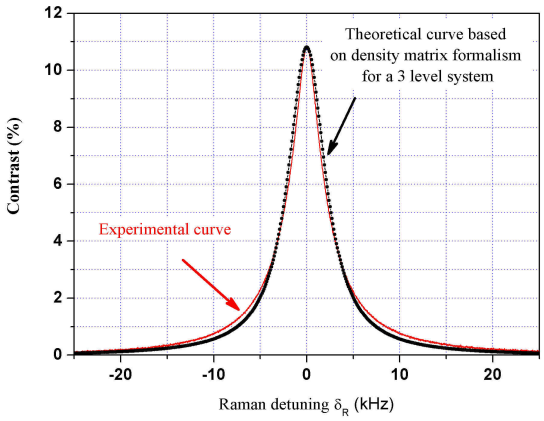


Fig. 2. Continuous CPT absorption across resonance. The experimental curve is fitted by a Lorentz profile derived from a density matrix calculation in a three-level atomic system.

To get clock signals with narrow linewidths, two approaches are possible. On one hand, continuous CPT interaction can be used at low laser intensity, since the width of this resonance is linearly dependent with the intensity, but eventually gives poor S/N. On the other hand, pulsed CPT interaction at high laser intensity has been demonstrated and give good S/N,

inspired from earlier work [2], where linewidths no longer depend on laser intensity but scale as $1/(2T)$. The time pulse sequence is achieved with an optical preparation pulse during τ followed by a free evolution time T in the dark. The intensity measurement is performed at a time τ_m after application of another light pulse (Fig. 3). This sequence is temporally

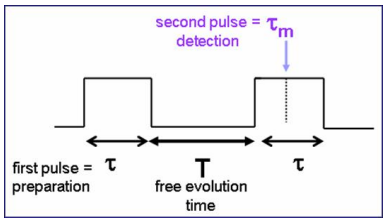


Fig. 3. The pulse time sequence (τ, τ_m, T) used to prepare and probe the CPT signals avoids the limitations due to optical saturation.

equivalent to the usual Ramsey sequence and leads to fringe widths independent of the laser intensity. By combining the lin-per-lin polarizations with the pulsed CPT method, high contrast and narrow width signals can be recorded as in Fig. 4. In particular, the 500 Hz wide central fringe focus in Fig. 5

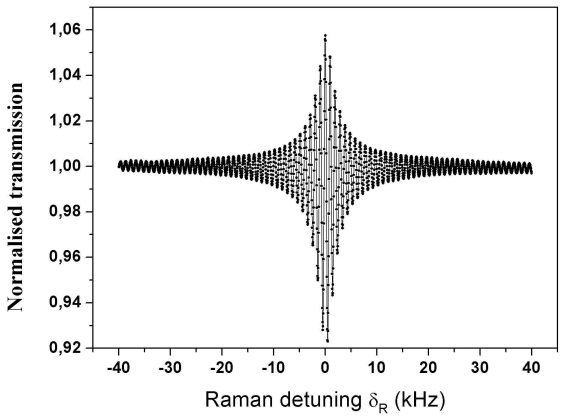


Fig. 4. Experimental saturated Raman-Ramsey fringes obtained with 5 ms preparation time and 5 μs measurement time. The total laser intensity 500 μW/cm^2.

corresponds to $T = 1$ ms free evolution time. It is important to note that two regimes govern the measured transmitted intensity, and the choice of τ_m imposes a choice of a regime. Let us assume that the lasers are optically resonant in the lambda configuration. For short τ_m, the transient atom-laser state is being pumped by spontaneous emission in the so-called dark state. In this transient regime, the oscillating Raman-Ramsey fringes are modulated by the Rabi envelope. Note that τ_m does not affect the fringe width but only the fringe envelope. The fringe width is determined only by the free evolution time T. For long τ_m, the atom-laser state evolves to the asymptotic situation where all the atoms are trapped in the dark state. In this stationary regime, the Lorentz signal width is determined by optical saturation.

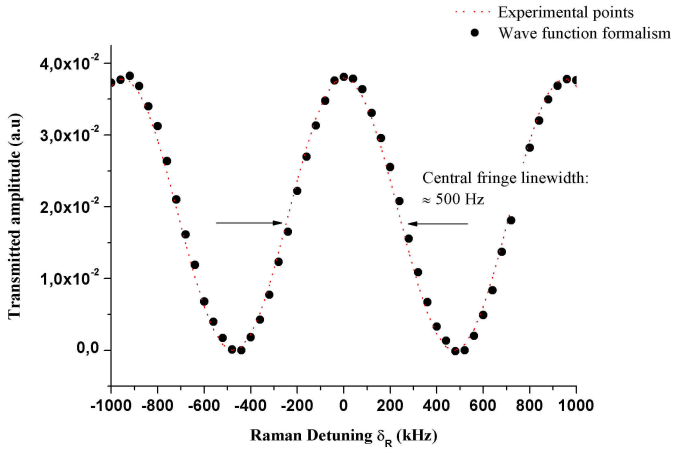


Fig. 5. The 500 Hz central fringe width corresponds to $T = 1$ ms free evolution time.

III. Time Sequence Optimization

To optimize the pulse sequence (τ, τ_m, T), the central fringe contrast is measured separately as a function of the three durations in the sequence. The Raman-Ramsey signal can be written as the product of a function only depending on the interaction times $F(\tau, \tau_m)$ by the oscillating fringe factor which is exponentially damped:

$$F(\tau, \tau_m)\left[1 - e^{(-\gamma_c T)}\cos\left(\delta_R T + \Delta\phi\right)\right]$$

In Fig. 6 the experimental central fringe contrast is plotted versus the preparation time τ. As the preparation time is

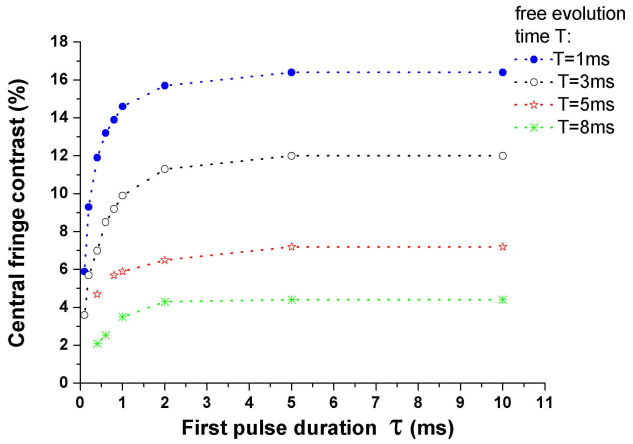


Fig. 6. Central fringe contrast versus the preparation time τ for different values of T. The total laser intensity is 500 μW/cm^2. The optimized sequence needs preparation times longer than 1 ms.

increased the contrast reaches a maximum. Thus after a typical 3 ms preparation time, atoms are pumped into the dark state. If the free evolution time is increased, the maximum level decreases, as described by the exponential decay factor, but the maximum is still reached after a few milliseconds.

In Fig. 7, the central fringe contrast is plotted versus the measurement time τ_m. As the measurement time is increased,

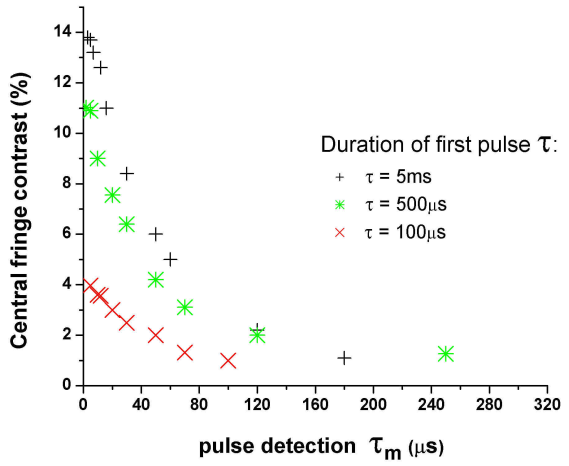


Fig. 7. Central fringe contrast versus the measurement time τ_m for different values of τ. The free evolution time is $T = 1$ ms and the total laser intensity is 500 μW/cm^2. The optimized sequence needs measurement times shorter than 10 μs.

the contrast rapidly decays to zero. Due to the exponential loss of contrast during the interaction time, the faster the measurement (about 1 μs), the higher the contrast (about 10%). As stated above, the contrast is higher for longer preparation times.

Finally, the central fringe contrast is plotted versus the free evolution time in Fig. 8. As the atoms freely evolves

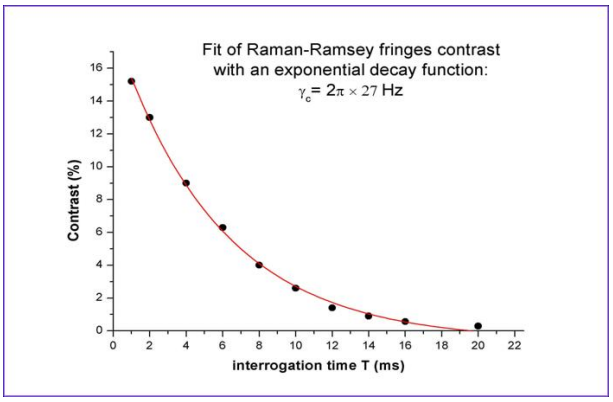


Fig. 8. Central fringe contrast versus the free evolution time T. The hyperfine coherence relaxation in the dark is responsible for the exponential decay. The optimized sequence needs free evolution times shorter than 20 ms.

during the dark time, the hyperfine coherence is exponentially damped due to relaxation effects in the cell like spin exchange, collisions in the buffer gas and diffusion to the walls. By increasing the free evolution time T, the loss of contrast can be fitted with an exponential decay function. The decay time constant 5.9 ms is a measurement of the coherence lifetime. In our experiment T can take values up to 20 ms.

The experimental optimization of the sequence has been validated by numerical calculations with density matrix formalism [4].

IV. Theoretical frequency shifts

The central fringe frequency shifts have already been studied in a Raman-Ramsey frequency standard [**?**], [2]. They are due to coherence phase accumulation during the atom-laser interaction and written as an additional phase shift term $\Delta\phi$ in Eq. 1. The associated frequency shifts mainly depend on the common optical detuning Δ_0, the interaction times and Rabi frequencies. As they are based on the simplest wave function formalism, the theoretical predictions cannot include the possible equipartition of the atoms between the clock levels.

First, the theoretical dependence is plotted versus the optical detuning. Calculations are made for $T = 1$ ms. Generally speaking, the central fringe frequency shift strongly depends on the optical detuning, up to 1.5 kHz for 3 GHz detuning (Fig. 8-9-10). However, there is a range of about 500 MHz around zero optical detuning for which the frequency shift can be canceled. The following discussion deals with this cancelation region. In Fig. 9, the frequency shift is plotted for different preparation times τ. The cancelation occurs for

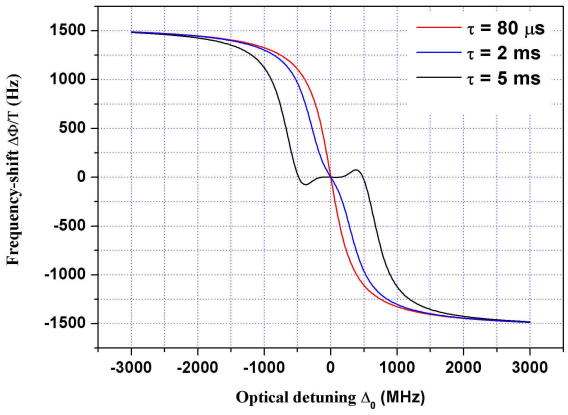


Fig. 9. Frequency shift versus the optical detuning for different preparation times τ and 2.5 μW/cm^2 laser intensity. The cancelation region appears for τ longer than 1 ms.

long τ, about 5 ms. In Fig. 10 the frequency shift is plotted for different laser intensities. Surprisingly, the cancelation occurs for high laser intensity (about 250 μW/cm^2). In both Figs. 9-10, the condition for shift cancelation is reached when the maximum number of atoms are trapped in the dark state for long enough preparation pulse with strong optical saturation. For shorter preparation times and weaker laser intensities, the shift dependence increases and can be linearly dependent with the optical detuning. In Fig. 11 the frequency shift is plotted for different relaxation rates. There again, frequency shift cancelation is possible and occurs for low collision rate, about 30 Hz, meaning weak hyperfine coherence relaxation.

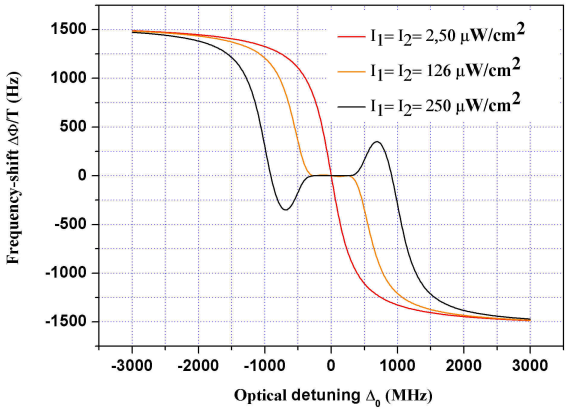


Fig. 10. Frequency shift versus the optical detuning for 80 μs preparation time and different laser intensities. The cancelation region appears for intensities higher than about 100 μW/cm^2.

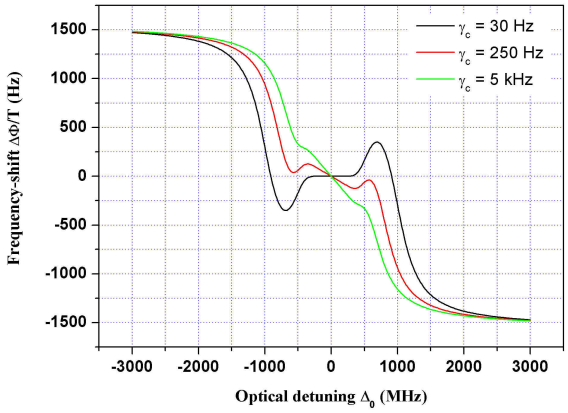


Fig. 11. Frequency shift versus the optical detuning for 80 μs preparation time and 250 μW/cm^2 for different collision rates. The cancelation region appears for rates lower than about 30 Hz.

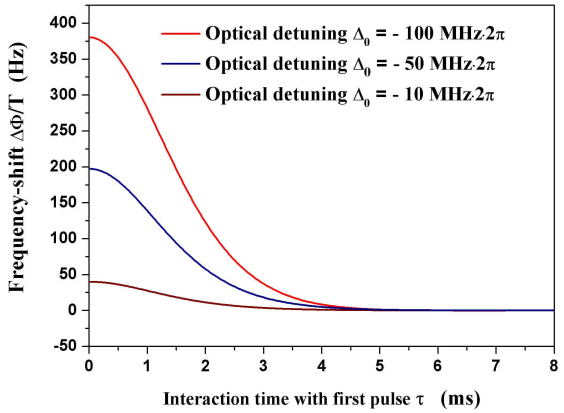


Fig. 12. Frequency shift versus the preparation time for 2.5 μW/cm^2 laser intensity and different optical detunings. The shift cancels for preparation time longer than 5 ms.

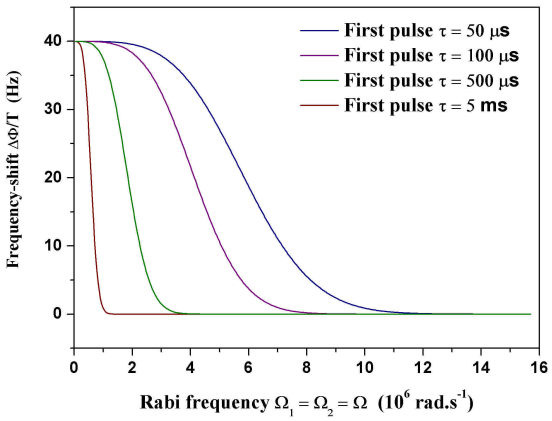


Fig. 13. Frequency shift versus Rabi frequencies for 10 MHz optical detuning and different preparation times. The shift cancels at high laser intensities.

Next, the theoretical dependence is plotted versus the preparation pulse parameters, namely the interaction time τ and the laser intensity (Figs. 12-13). In Fig. 12 the frequency shift is plotted versus τ for different optical detunings. This figure is actually complementary to Fig. 9. As previously mentioned, the frequency shift depends on the optical detuning but can be canceled by increasing the preparation time to about 5 ms. In Fig. 13 the relative frequency shift is plotted versus Rabi frequencies (proportional to the square root of laser intensities) for different preparation times. The relative frequency shift decays to zero by increasing the laser intensities, and once again the cancelation is faster for higher intensities. As the intensity increases, the shift cancelation level goes from 5.10^{-9} with 20 nW/cm^2 and is lower than 10^{-13} with 4 μW/cm^2 laser intensity.

From previous graphs and discussions, the Raman-Ramsey pulse sequence is optimized for long preparation pulses and high laser intensities. Resulting experimental fringes are plotted in Fig. 14. With 5 ms preparation time, the central fringe contrast is 14%. To compare both continuous and pulsed CPT interactions, a figure of merit can be defined as the signal contrast-to-linewidth ratio. This ratio is 65 times higher for pulsed CPT than for continuous CPT.

V. Conclusion

CPT pulses with crossed linear laser polarizations provide high contrast Raman-Ramsey fringes (up to 14%) with small linewidths (down to 60 Hz). The pulse sequence is optimized for long preparation time (5 ms), high laser intensities (a few μW/cm^2) and very short measurement time in order to saturate the Raman process. These values also cancel the relative frequency shift of the central fringe to a level below 10^{-13} for a pessimistic optical detuning set as 10 MHz.

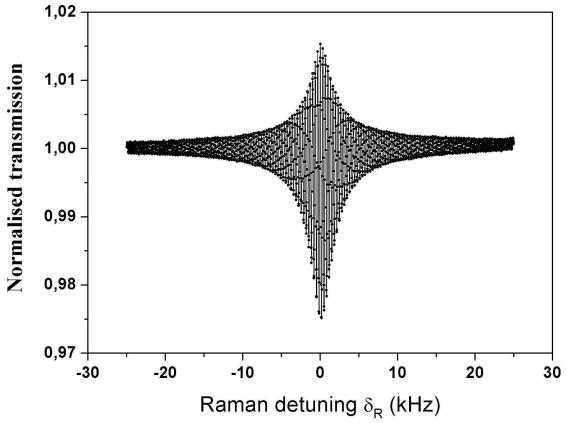


Fig. 14. Experimental Raman-Ramsey fringes obtained with a pulse time sequence optimizing the central fringe contrast-to-linewidth ratio. This time sequence ($\tau = 5$ ms, $\tau_m = 10$ μs, $T = 3$ ms) generates a 166 Hz wide central fringe with 14% contrast.

REFERENCES

[1] T. Zanon, S. Trémine, S. Guérandel, E. de Clercq, D. Holleville, N. Dimarcq, A. Clairon, *Observation of Raman-Ramsey fringes with optical CPT pulses*, IEEE Trans. Instr. Meas., vol 54, No 2 p 776 (2005).

[2] M.S. Shariar, P.R. Hemmer, D.P. Katz, A. Lee, M.G. Prentiss, *Dark-state based three-element vector model for the stimulated Raman interaction*, Phys. Rev. A, vol 55, No 3, p 2272 (1997).

[3] P.R. Hemmer, M.S. Shariar, V.D. Natoli, S. Ezekiel, *AC Stark shifts in a two-zone Raman interaction*, J. Opt. Soc. Am. B, vol 6, No 8, p 1519 (1989).

[4] G. Orriols, *Nonabsorption resonances by nonlinear coherent effects in a three level system*, Il Nuovo Cimento, B 53 vol 1, p 1 (1979).

Short term stability limitations in a compact cold atom clock

S. Trémine[(1)], S. Guérandel[(1)], D. Holleville[(1)], N. Dimarcq[(1)], J. Delporte[(2)] and A. Clairon[(1)]

[(1)]LNE-SYRTE – CNRS UMR8630 – Observatoire de Paris – 61 avenue de l'Observatoire 750014 Paris – France

[(2)]CNES – 18 avenue Edouard Belin – 31055 Toulouse Cedex – France

email: stephane.tremine@obspm.fr

***Abstract*— HORACE is a compact cold atom clock where the atoms are cooled inside the microwave interrogation cavity, and then detected by linear absorption. We report preliminary fringes obtained with this method. The optimization of the short term stability with the cooling duration is presented, and we show that an Allan deviation of $\sigma_y(\tau) = 2.5\ 10^{-13}\ \tau^{-1/2}$ should be obtained on Earth. For the calculations, we take into account the quantum projection noise, the Dick effect, the detection laser noises, and the cesium vapor absorption fluctuations.**

I. Introduction

Compared to traditional clock configurations like atomic fountains or thermal beams, the basic idea is to realize all the different phases of the clock operation in the same region (Fig. 1). In our case, this region is the copper cavity presented in Fig. 2, where we successively cool the atoms from a cesium vapor, then switch off the cooling light and perform a Ramsey interrogation by generating two microwave pulses, and finally measure the transition probability using a linear absorption method through the cloud of atoms.

The internal diameter of the cavity is about $4\ cm$, set by the resonance at 9.2 GHz for the interrogation phase. It is optically polished for the cooling phase. Indeed, we don't use collimated beams as it is usually done, but the cooling light is injected in a multimode optical fiber, then separated in six to feed the cavity. Consequently, the light diverges at the end of the fibers and reflects on the cavity walls. With such a configuration, and using laser parameters suitable for sub-Doppler cooling, we managed to cool about 10^8 atoms at a kinetic temperature of $2\ \mu K$ ($v = 1\ cm.s^{-1}$) [1].

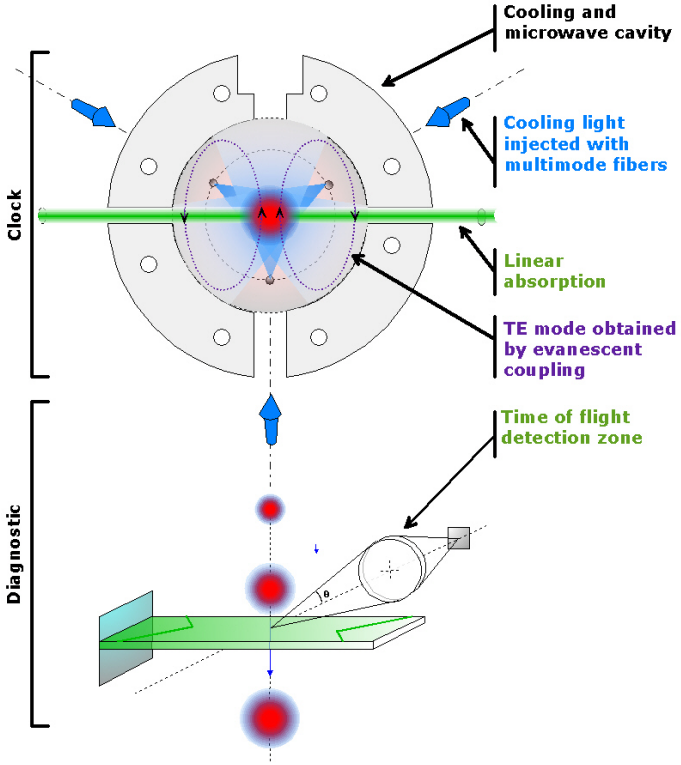


Fig. 1. Physics package. The clock part is composed of a quartz bulb (see Fig. 2) surrounded by a spherical copper cavity resonant at $9.2\ GHz$. The cooling light is injected into the cavity via $\Phi = 1\ mm$ apertures following the (1,1,1) configuration. The microwave field is coupled to the cavity via a $1 \times 6\ mm^2$ vertical slit. On the present device, an horizontal detection beam goes through the cavity via $\Phi = 4\ mm$ apertures. The diagnostic part is composed of a vacuum chamber and a detection zone used for time of flight measurements.

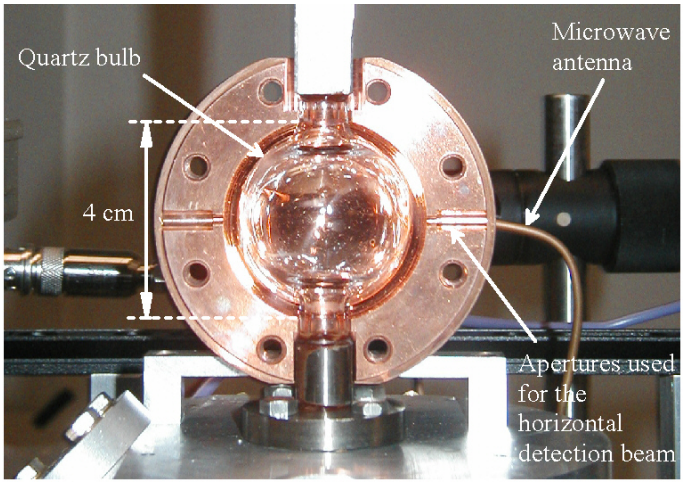


Fig. 2. Inside the microwave cavity, one can see the quartz bulb containing the Cs vapor, in which the vacuum is about 2 10^{-9} $Torr$. A part of the cesium reservoir appears above the cavity. The cavity is optically polished at the level of $\lambda/14$, and the reflection coefficient is about 96%. The quartz bulb has thickness inhomogeneities of about 0.5 mm and is far from a perfect spherical shape.

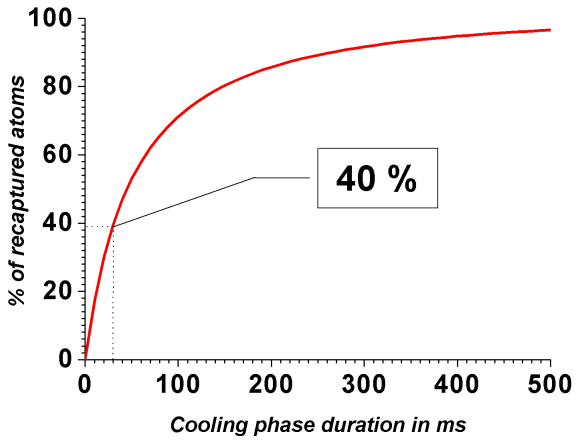


Fig. 3. Stationary number of recaptured cold atoms reached after about ten cycles for a falling time set to 20 ms. This means that the composition of each cycle is [T_{cool} = (given by the abscise on this graph)] + [$T_{fall} = 20$ ms]. On the vertical axis, 100% of recaptured atoms is equivalent to 3 10^8 atoms obtained with capture parameters for the cooling.

As in other pulsed cold atom clocks, the local oscillator frequency is not compared to the atomic resonance during the cooling and the detection times. The associated Dick effect limitation requires to improve the duty cycle given by Eq. (1). T_i is the interrogation time, T_{cool} is the cooling time, and T_{fall} is the falling time including the interrogation and the detection times between two cooling phases. T_c is the cycle duration.

$$\text{Duty cycle} = \frac{T_i}{T_{cool} + T_{fall}} \quad \text{with} \quad \begin{cases} T_{fall} = T_i + T_{detection} \\ T_c = T_{cool} + T_{fall} \end{cases} \tag{1}$$

To reduce the cooling time, we need to recapture a part of the atoms cooled in previous cycles. This is done with Doppler parameters for the cooling, which leads us to a maximum number of 3 10^8 atoms with a kinetic temperature of 35 μK ($v = 5$ $cm.s^{-1}$). This temperature is below the Doppler limit ($125\mu K$) due to Sisyphus processes that occur during the capture phase. The recapture has been presented in previous papers [2]. We have shown that a stationary number of recaptured atoms can be reached after about 10 cycles, with a falling time T_{fall} set to 20 ms. This stationary number $N_{(T_{fall},T_{cool})}$, depends on the two parameters T_{fall} and T_{cool}. This function is presented in figure 3 for $T_{fall} = 20$ ms, and will be used in section III-A to optimize the short term stability. Obviously, the longer the cooling phase, the higher the stationary number, but the worse the duty cycle. This is the basis of the trade-off discussed in section III.

Before dealing with the expected stability, we are first going to present the main results obtained with our device, concerning the interrogation and detection phases. Those results will lead us to describe the detection steps, and the clock sequence that we will follow to reach our stability goal of 5 10^{-13} at 1 s.

II. First results - Normalization sequence

A. First results

First, laser cooling and microwave Ramsey interrogation have been validated in the same region, using a time-of-flight detection located a few centimeters below the cavity (diagnostic part in Fig. 1). By detecting only one population (no normalization), we have obtained the fringes presented in Fig. 4 (span of 2 kHz) and Fig. 5 (zoom of the central fringe). After a Ramsey sequence composed of two 2.5 ms microwave pulses separated by a

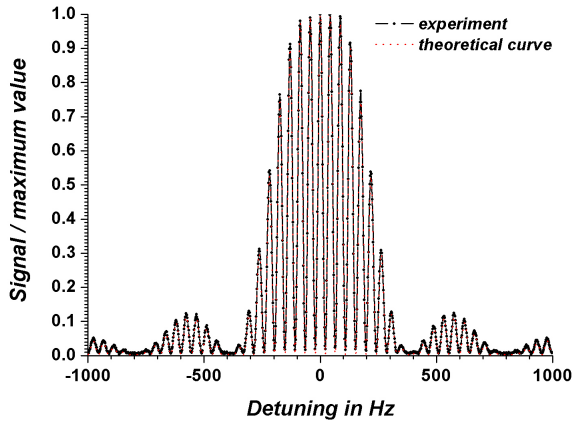


Fig. 4. Ramsey pattern obtained with a time of flight technique after two microwave pulses of 2.5 ms each, separated by 20 ms dead time. The Rabi envelope is characterized by a FHWM of 400 Hz and the Ramsey FHWM is 22 Hz. In this measurement, only the population that has done the transition has been detected. There is no normalization by the total population of the two concerned levels.

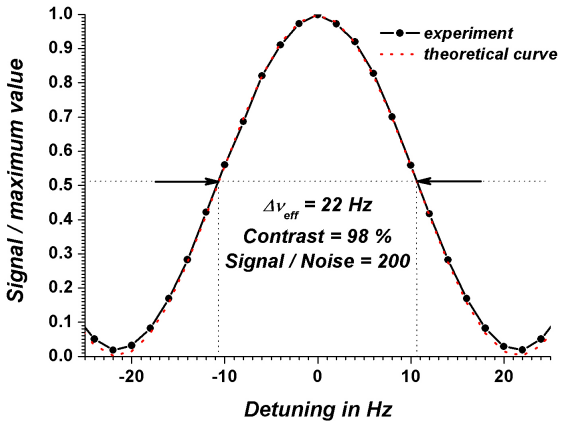


Fig. 5. Zoom on the central fringe of the figure 4. The residual 2% loss of contrast is due to stray light present in the cavity during the interrogation phase. This is responsible of the destruction of the coherence between the two microwave pulses. So, the longer the interrogation phase, the higher the loss of contrast. In our experiment, due to the presence of the optically polished cavity around the interrogation zone and the proximity of the cooling fibers, we are very sensitive to stray light.

20 ms dead time, the full width at half maximum (FWHM) is 22 Hz, the contrast is 98% and the signal to noise ratio is 200 limited by the captured atoms number fluctuations. With such a detection, the narrowest signal we have obtained is 9 Hz wide, corresponding to a Ramsey interrogation time of 50 ms, with a contrast of 85%.

To finally perform all the clock steps at the same place, we then tried to detect the clock signal by linear absorption of an horizontal probe beam, using the apertures available on our device (clock part in Fig. 1). With this configuration, we obtained the fringes presented in Fig. 6 with a poor signal to noise ratio and a FWHM of 86 Hz. Unfortunately, we have poor conditions for this detection, since we cannot realize a standing wave due to the quartz bulb (see Fig. 2), which besides generates stray reflections destroying the cold atom cloud. Moreover, this horizontal configuration is responsible of a fast decrease of the absorption level (atoms are falling below the detection beam), which limits the interrogation time to a few milliseconds.

Therefore, the solution is to have a vertical detection beam that only passes through good quality view-ports. As a matter of fact, the atoms will stay in the detection beam all along their fall, and the maximum interrogation time will only be determined by the atoms leaving the cavity (about

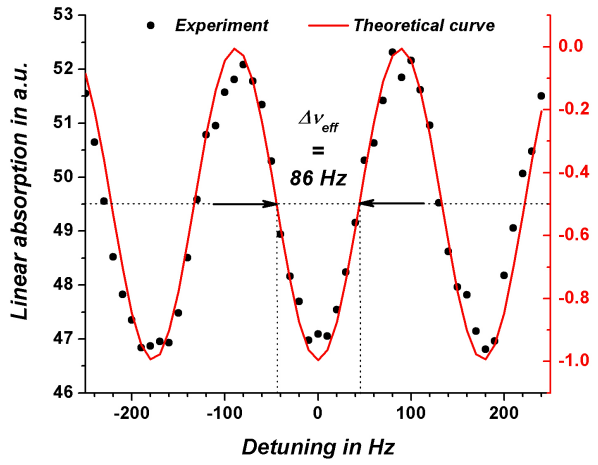


Fig. 6. Fringes obtained with a linear absorption technique through the cloud (clock part in Fig. 1), averaged on three samples. In this measurement, we detect the population that has not made the transition (this choice is due to technical problems), that is why we have an inverted fringe. Moreover, on the contrary to the fringes of the figure 4, the background includes the contributions of the other Zeeman sublevels $|F = 4, m_F \neq 0\rangle$. Therefore, those atoms significatively contribute to the noise. Nevertheless, it is not sufficient to explain the excess of noise observed on the fringes.

60 ms). With this new detection scheme, we expect to retrieve fringes similar to the "time of flight" ones, with a comparable signal to noise ratio. Nevertheless, a S/N value of 200 is not sufficient to reach our short term stability goal below $5\ 10^{-13}$ at $1s$. Indeed, considering a realistic sequence (taking into account the recapture properties, the necessary detection time ... see part II-B $\Rightarrow$ $T_{cool} = 30\ ms$, $T_i = 10\ ms$, $T_{detection} = 10\ ms$, and $S/N = 200$), the expected stability limitation would be about $2\ 10^{-12}\ \tau^{-1/2}$.

B. Normalization and clock sequence

The detection of the two clock level populations allows to reject, in a first time, the atom number fluctuations, and finally reach a better signal to noise ratio. Besides, we absolutely need to minimize the loss of atoms between two successive cooling phases. That is why we propose the normalization sequence shown in Fig. 7 to minimize those losses, and reject the technical noises.

At the end of the cooling phase, all the atoms are in the level $F = 4$, all m_F. The first step consists in optical pumping the atoms from the level 4 to 3. Then the microwave interrogation is applied, which means that half of the $|F = 3, m_F = 0\rangle$ Zeeman sublevel population is transferred to $|F = 4, m_F = 0\rangle$. Finally the population (N_4) of this level is measured with the vertical detection beam. In the next step we clear out the level 4 with a blue-shifted laser light, transfer the atoms left in the level 3-0 to the level 4 using a π microwave pulse, and finally detect this population (N_3).

We know consequently the populations N_4 and N_3, which allows us to calculate the transition probability using Eq. (2).

$$P = \frac{N_4}{N_3 + N_4} \tag{2}$$

One can easily deduce that the probability fluctuations are given (in the case of a $\pi/2$ interrogation, which means $P = 1/2$) by the Eq. (3) considering the noise on the measurements of N_4 and N_3, respectively given by δN_4 and δN_3.

$$\delta P = \frac{1}{2} \times \frac{\delta N_4 - \delta N_3}{N_3 + N_4} \tag{3}$$

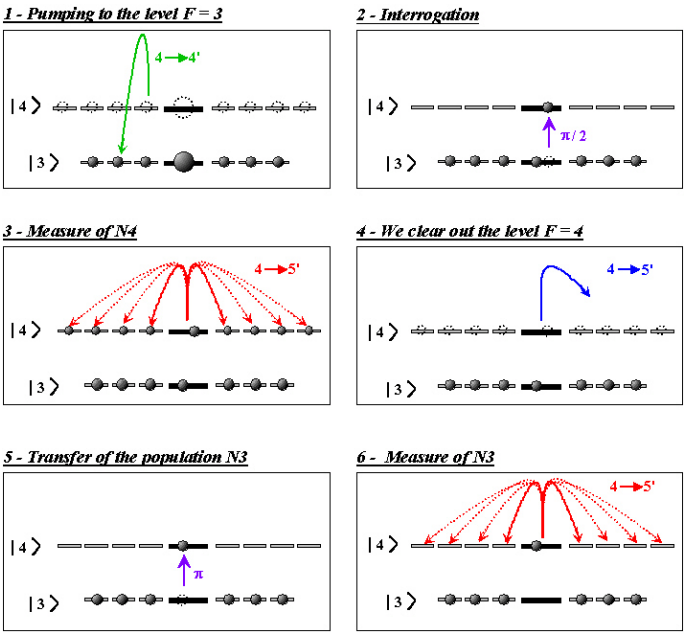


Fig. 7. Normalization sequence that will be used in our next device. After the cooling phase, the atoms are transferred to the level $F = 3$. On the one hand, it allows us to keep all the atoms contained in the Zeeman sublevels $m_F \neq 0$ for the next recapture cycle. On the other hand, we detect the populations on a black background. Moreover, we have more atoms that contribute to the signal, since they are distributed between seven sublevels compared to nine in the case of the $F = 4$ hyperfine level.

We can see that if the noises on those measurements are totaly correlated, like the captured atom number fluctuations, they are completely rejected thanks to the normalization.

In the next part, we will study different noises that can limit the short term stability, using those expressions.

So, now that we know how a cycle is organized, we can present a realistic sequence by setting the durations of the different phases, which will be useful for our calculations (see Fig. 8).

The optimization of the short term stability will be done in this paper by adjusting the cooling time. Since we have studied the recapture for a dead time set to 20 ms (see part I and [2]), we keep this value. This dead time is composed of 10 ms dedicated to the detection time (the previous normalization sequence), and then we have 10 ms for the interrogation. From this sequence we can estimate the number of atoms useful for the clock signal.

We know that the full molasses contain $3\ 10^8$

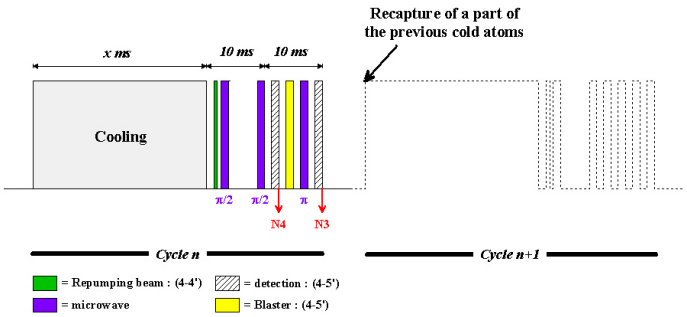


Fig. 8. Realistic clock sequence where the interrogation and the detection times are set to 10 ms each. For the optimization of the short term stability (part III), the cooling phase duration is the adjustable parameter. All the steps of Fig. 7 are presented on this scheme.

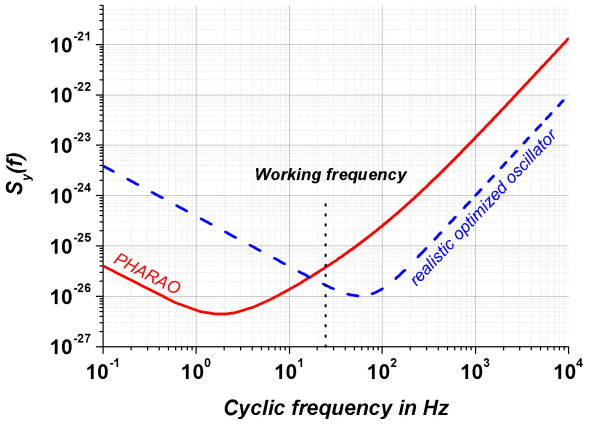


Fig. 9. Spectral density of frequency noise for two different oscillators versus the cyclic frequency. *Continuous line* – hight performances local oscillator used in the PHARAO project. *Dashed line* – realistic optimized oscillator for our clock. Our working frequency region is indicated.

atoms, but we only use 16% of them (measured on a microwave spectrum of the hyperfine structure), the ones contained in the 3-0 Zeeman sublevel. Moreover, due to the size of our detection beam and to the thermal expansion of the cloud, we will only detect 38% of those atoms after $20\ ms$. And finally, if we consider a cooling time of $x\ ms$, the recapture tell us (Fig. 3) that we will have a given fraction of this number. For example, setting the cooling time to $30\ ms$ gives $x = 40\%$, corresponding to $7\ 10^6$ useful atoms.

III. Short term stability limitations

A. Quantum projection noise

With this information, we can determine the ultimate limitation for the short term stability, which is given by the quantum projection noise (Eq. 4) [3].

$$\sigma_y(\tau) = \frac{1}{\pi}\frac{1}{\nu_0.2T_i}\frac{1}{\sqrt{N_{(T_{fall},T_{cool})}}}\sqrt{\frac{T_c}{\tau}} \qquad (4)$$

In this formula, $N_{(T_{fall},T_{cool})}$ is the number of atoms given by the recapture (the function presented in Fig. 3) and only depends of the cooling duration when the falling time is set. The corresponding variance is presented in Fig. 14. From this, the optimization of the stability only limited by the quantum projection noise, is obtained for $T_{cool} = 30\ ms$ and is $1.5\ 10^{-13}\ \tau^{-1/2}$.

B. Local oscillator phase noise

Moreover, we have performed Dick effect [4] [5] calculations considering first an excellent local oscillator used in the PHARAO project [6] (see the continuous line in Fig. 9), and a realistic and more convenient one for our clock (dashed line). The first one is optimized for a fountain configuration, in the range of $1\ Hz$ for the cycle frequency. In our case, we would look for an oscillator optimized in the range of $20 - 30\ Hz$. The corresponding Dick limitations on the stability for those two oscillators are presented in Fig. 14.

We are going to see that for this type of clock, the local oscillator phase noise is one of the main limitations (Dick effect), but since we will use linear absorption detection compared to the fluorescence usually used in most other clocks, the detection laser noise may significatively contribute to the short term stability limitation. This is what we are going to discuss now.

C. Detection laser noise

The vertical detection beam will first pass through a cesium vapor column of $4.5\ cm$ long (Fig. 10), before crossing the cold atoms cloud. The full cesium vapor column is $6.6\ cm$ long. The graph in Fig. 11 presents the different absorption levels versus the time in ms after the cooling light has been switched off. The dashed line corresponds to no absorption, only the laser intensity. The dotted line is the vapor absorption, and the continuous line is the cold atoms absorption. So, one can guess now that the main noise sources seen by the detector

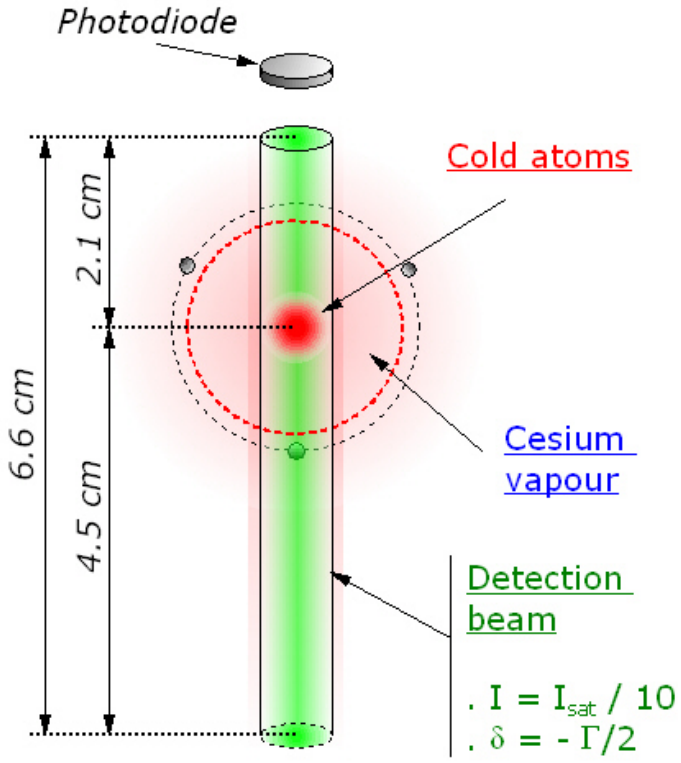


Fig. 10. Detection using a vertical probe beam. Due to technical constraints, we have to take into account a cesium vapor column with the characteristics indicated on the scheme.

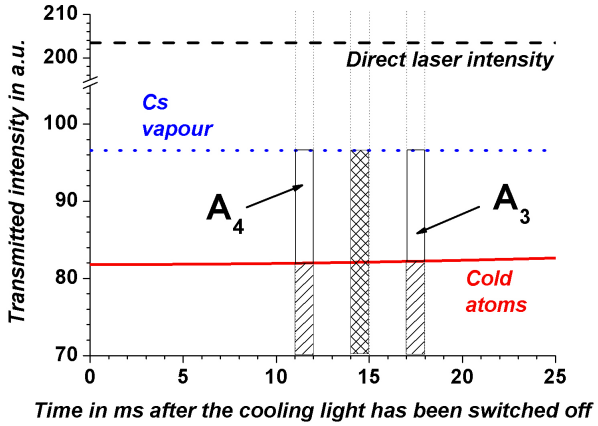


Fig. 11. Considering a vertical detection beam for our next device, we have calculated the different absorption levels due to the cesium vapor (*dotted line*) and to the cold atom cloud (*continuous line*). Without any absorption, the reference level is given by the *dashed line*. About 50% of the light is absorbed by the cesium vapor, and about 10% by the cold atoms. This graph shows how the useful areas A_4 and A_3 will be calculated owing to the complementary areas (*dashed areas*), and to the measurement of the cesium vapor absorption level (*crossed area*).

will be the intensity fluctuations, and the vapor absorption fluctuations. In order to calculate the transition probability, we will realize two different operations that will filter those noise sources [7] [8] [9].

Considering Fig. 11, the populations N_4 and N_3 will be proportional to the areas named A_4 and A_3, integrated between the two absorption levels. During the clock sequence, the available signals will be the two complementary dashed areas. To calculate A_4 and A_3, we will need to measure the cesium absorption level on the same duration in order to make the substractions. This operation is equivalent to a base-line substraction. Then, with A_4 and A_3 we can calculate the transition probability (Eq. 2). Those two operations, "base-line substraction" and "calculation of the transition probability" induce transfer functions that filter intensity and frequency fluctuations of the laser. The important parameters used for the calculations are the detection laser intensity set to $I_{sat}/10$ (where I_{sat} is the saturated intensity), and the detuning set to $\delta = -\Gamma/2$. Those parameters will maximize the sensitivity to intensity and frequency noises, which will lead us to maximize the limitations.

For the base-line substraction, we consider the transfer function presented in Fig. 12.a, which represents the difference between the crossed and one of the dashed areas. After Fourier transform, this leads us to the function presented in Fig. 12.b, that we will use as a filter for the spectral density of tension fluctuations of the detector.

Different sources can be responsible of those tension fluctuations. The first one is the laser intensity noise, and the second one is the fluctuation of the cesium vapor absorption due to laser intensity noise, or laser frequency noise.

For our detection laser, we will consider $S_{\delta I/I}(f) = 3\ 10^{-9}/f$ for the spectral density fluctuations of intensity, and $S_y(f) = 2\ 10^{-26}\ Hz^{-1}$ for the spectral density fluctuations of frequency [8]. The previous characteristics have been measured on an extended cavity laser frequency locked on a saturated absorption. Besides, the power of the laser is stabilized by means of a servo using an accousto-optic-modulator. Those laser fluctuations seen by the detector can be interpreted in term of detected

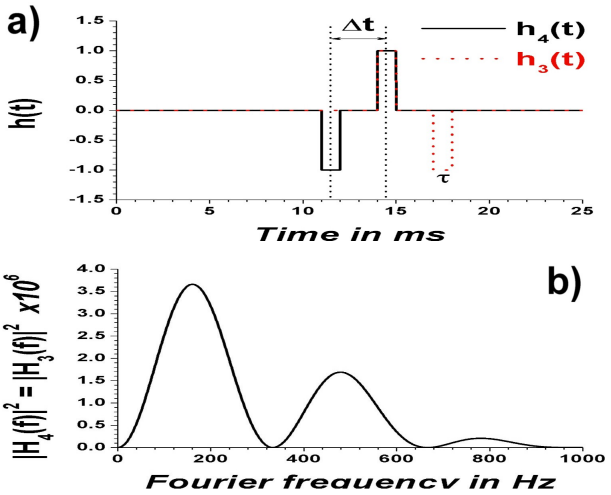


Fig. 12. Base line substraction. (a) Transfer functions induced by the calculation of the area A_4 (Fig. 11) (*continuous line*) and A_3 (*dashed line*). (b) Fourier transform of the transfer functions presented in (a).

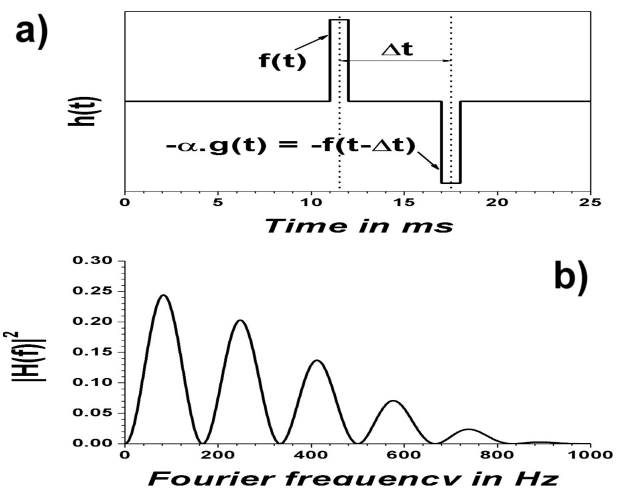


Fig. 13. (a) Transfer function induced by the transition probability calculation (Eq. 2). Since the cold atoms absorption decreases with time (see Fig. 11), the areas A_4 and A_3 that should be equal after a $\pi/2$ interrogation, are a little bit different. These areas are described respectively by the functions $f(t)$ and $g(t)$. In order to have $P = 1/2$ at half maximum of the central fringe, we multiply $g(t)$ by a calibrated constant α such as $\alpha.g(t) = f(t - \Delta t)$. Δt is the time between the two areas measurement. (b) Fourier transform of the transfer functions presented in (a).

atom fluctuations $\sigma_{N_{det}}$, and finally translated in term of stability limitation owing to Eq. (5) where $N_{det} = 7\ 10^6$ from section II-B.

$$\sigma_y(\tau) = \frac{2}{\pi}\frac{\sigma_{\delta P}}{Q_{at}}\sqrt{\frac{T_c}{\tau}} \text{ avec } \sigma_{\delta P} = \frac{1}{\sqrt{2}}\frac{\sigma_{N_{det}}}{N_{det}} \quad (5)$$

After calculations, we find $\sigma_{N_{det}} = 373$ atoms due to the laser intensity fluctuations leading to $\sigma_y(\tau) = 3\ 10^{-13}\ \tau^{-1/2}$, and $\sigma_{N_{det}} = 183$ atoms for the variation of the cesium vapor absorption, leading to $\sigma_y(\tau) = 1.4\ 10^{-13}\ \tau^{-1/2}$, mainly dominated by the intensity fluctuations. Those limitations have been calculated for $T_{cool} = 30\ ms$, leading to $T_c = 50\ ms$ (to continue the example given in section II-B).

Following a similar approach for the transition probability calculation, we find the transfer function presented in Fig. 13, and consequently, a limitation at $1.2\ 10^{-13}\ \tau^{-1/2}$ for the short term frequency stability.

All those stability limitations are summarized in Fig. 14, that presents the short term stability against the cooling phase duration. The quadratic sum of those different effects is given by the continuous lines.

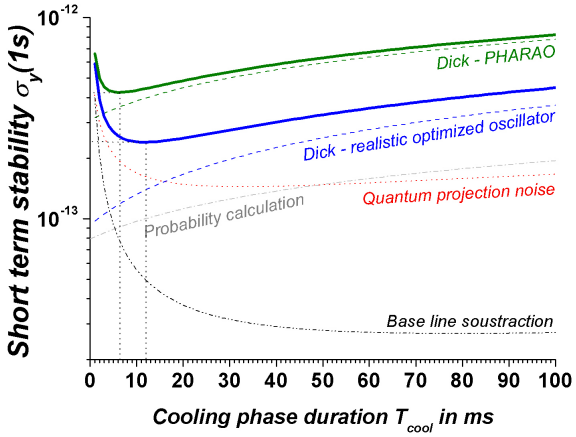


Fig. 14. Optimization of the short term frequency stability with the cooling phase duration. Taking into account the quantum projection noise, the Dick effect, and the detection laser noises, a non optimized local oscillator like the PHARAO one should lead to $\sigma_y(\tau) \sim 5\ 10^{-13}\ \tau^{-1/2}$. With a realistic optimized oscillator, $\sigma_y(\tau) \sim 2\ 10^{-13}\ \tau^{-1/2}$ may be obtained on Earth.

IV. Conclusion

To conclude, we will be mainly limited by the Dick effect at a value of $5\ 10^{-13}\ \tau^{-1/2}$ with a non optimized local oscillator, and we could reach on Earth a short term stability of about $2\ 10^{-13}\ \tau^{-1/2}$ with an optimized one.

The implementation of a new design using the in-situ vertical detection we have presented, is in progress. Then, the recapture will be studied further

to optimize the short term stability with the interrogation time, since the optimization presented in this paper was only based on the cooling duration. We will finally extrapolate those results to a clock operating in space, since we expect to improve those performances in a micro-gravity environment [10] [11].

V. Acknowledgment

This work is supported in part by CNES. S. Trémine is grateful to CNES and SODERN for providing research studentship.

References

[1] P.-E. Pottie, S. Trémine, S. Guérandel, N. Dimarck, A. Clairon, *3D speckle cooling in a microwave clock*, In Proc. of the 17th EFTF, Tampa, Florida (2003).

[2] S. Trémine, S. Guérandel, D. Holleville, N. Dimarck, A. Clairon, *Microwave interrogation in a compact atomic clock*, In Proc. of the 18th EFTF, Guilford, England (2004).

[3] G. Santarelli, Ph. Laurent, P. Lemonde, A. Clairon *Quantum Projection Noise in an Atomic Fountain : A High stability Cesium Frequency Standard*, Phys. Rev. Lett., vol 82, p 4619, (1999).

[4] G. Santarelli, C. Audoin, A. Makdissi, Ph. Laurent, G. Dick, A. Clairon *Frequency Stability Degradation of an Oscillator Slaved to a Periodically Interrogated Atomic Resonator*, IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol 45, p 887, (1998).

[5] C. Audoin, G. Santarelli, A. Makdissi, A. Clairon *Properties of an Oscillator Slaved to a Periodically Interrogated Atomic Resonator*, IEEE Trans. Ultrason., Ferroelect., Freq. Contr., vol 45, p 877, (1998).

[6] P. Lemonde, Ph. Laurent, G. Santarelli, M. Abgrall, Y. Sortais, S. Bize, C. Nicolas, S. Zhang, A. Clairon, N. Dimarcq, P. Petit, A. Mann, A. Luiten, S. Chang, C. Salomon *Cold Atom Clocks on Earth and in Space*, Frequency Measurement and Control, Topics in Applied Physics, A. Luiten ed. 79, p 131, (2000).

[7] E. Simon, *Vers une stabilité et une exactitude de* 10^{-16} *pour les horloges atomiques : - le rayonnement du corps noir - la détection optique*, Thesis from the Université Paris-XI, Orsay (1997).

[8] S. Bize, *Tests fondamentaux à l'aide d'horloges à atomes froids de rubidium et de césium*, Thesis from the Université Paris-VI, Paris (2001).

[9] Y. Sortais, *Construction d'une fontaine double à atomes froids de* ^{87}Rb *et* ^{133}Cs *; étude des effets dépendant du nombre d'atomes dans une fontaine*, Thesis from the Université Paris-VI, Paris (2001).

[10] G. Santarelli et al., *Laser cooling of atoms in micro-gravity*, C. R. Acad. Sci. Paris, t. 316, Série II, p 739-744 (1993).

[11] B. Lounis, J. Reichel, C. Salomon, *Laser cooling of atoms in micro-gravity*, C. R. Acad. Sci. Paris, t. 316, Série II, p 739-744 (1993).

Session 2A / Piezoelectric sensors

F_2 - LASER ABLATION AND SURFACE – PATTERNING OF $GAPO_4$ FOR SENSOR APPLICATIONS

M. Peruzzi, J.D. Pedarnig *, S. Roither, D. Bäuerle
Institute of Applied Physics, Johannes-Kepler-University Linz, A-4040 Linz, Austria

H. Salhofer, R. Schwödiauer
Department of Soft Matter Physics, Johannes-Kepler-University Linz, A-4040 Linz, Austria

W. Reichl, J. Runck, H. Mitter
E+E Elektronik GmbH, A-4209 Engerwitzdorf, Austria

ABSTRACT

Pulsed-laser ablation of $GaPO_4$ single crystals and ceramics and sensor applications of laser–patterned high fundamental frequency (HFF) $GaPO_4$ resonators were demonstrated. Well defined micron-sized holes with sharp edges between ablated and non-ablated areas were obtained at the $GaPO_4$ surface employing projection-mask techniques. F_2 - laser direct writing was employed to reduce the thickness of the piezo-active area and to increase the resonance frequency of $GaPO_4$ thickness shear-mode resonators from ~ 6 up to ~ 12 MHz. The HFF $GaPO_4$ resonators were employed for humidity sensing using a polyimide receptor layer deposited onto the non-ablated side of the $GaPO_4$ crystal. The resonance frequency of the coated 12 MHz HFF $GaPO_4$ revealed a linear decrease with increasing level of relative humidity (RH), δf = -100 Hz / %RH.

1. INTRODUCTION

Ultraviolet lasers are important tools for the processing of advanced materials [1]. The short wavelength / high photon energy of such lasers enables applications where high spatial resolution and / or the processing of wide-band-gap materials is required [2]. F_2 - lasers (wavelength 157 nm, photon energy 7.9 eV) are among the shortest wavelength sources that are commercially available. Fluorine lasers have been employed (among others) for surface patterning and micro-machining of chemically inert and wide-band-gap materials such as polytetrafluoroethylene (PTFE), quartz and gallium orthophosphate [2, 3]. The new material gallium orthophosphate ($GaPO_4$) is especially interesting for piezoelectric applications due to its large electro-mechanical coupling coefficient and very high phase transition temperature as compared to quartz [4]. Piezoelectric $GaPO_4$ thickness-shear- mode (TSM) resonators have quality factors comparable to AT-cut quartz crystals [4] and can be employed, for instance, in micro-balance devices for sensing in gaseous and liquid ambient. The sensitivity of quartz micro-balances strongly increases with frequency [5] and so-called high fundamental frequency (HFF) devices are produced by wet-chemical etching [6] and ion etching [7] mainly. The processing of $GaPO_4$ is more difficult and wet-chemical etching of $GaPO_4$ HFF resonators has been reported only recently [8].

In this paper, we report on pulsed-laser ablation and micro-patterning of $GaPO_4$ single crystals by means of F_2 - laser radiation and the fabrication of HFF $GaPO_4$ TSM resonators for gas sensing. The laser-ablated crystals were investigated by scanning electron microscopy (SEM), optical microscopy, UV-VIS spectrometry, surface profilometry and atomic force microscopy (AFM). For humidity sensing the crystals were coated with a thin polyimide layer. The impedance spectra of $GaPO_4$ and quartz TSM resonators were recorded at room temperature at various levels of relative humidity (RH) employing a commercial humidity generator (HUMOR 10) and an impedance analyser (HP4291B).

2. EXPERIMENTAL

The processing of $GaPO_4$ was performed by pulsed-laser ablation techniques employing the deep UV radiation of a Lambda Physik LPF 202 system (wavelength 157 nm, pulse duration 15 ns, pulse energy up to 60 mJ). The laser beam is guided in a sealed aluminum box which is directly connected to the laser head and continuously flushed with dry N_2 gas. This avoids absorption of the radiation by O_2 and H_2O. The box contains the optical components, translation stages and energy meter. All experiments were performed in a vacuum chamber that was attached to the flushed box and separated via a CaF_2 window. A low-pressure argon atmosphere (p(Ar) = 0.03 mbar) was employed to reduced the formation of debris on the crystal surface. The imaging system did consist of two plan-convex CaF_2 lenses with an effective focal length of ~ 120 mm. The $GaPO_4$ single crystals used (diameter 5 mm, thickness from 0.5 mm to 1 mm) were polished before ablation.

For F_2 - laser direct writing, a rectangular aperture (2 mm × 24 mm) was used as projection mask. The laser fluence on the sample was $\Phi \approx 2 \pm 0.2$ J/cm^2 and the pulse repetition rate was 20 Hz. The $GaPO_4$ samples were mounted on a computer controlled 2D

translation stage which was synchronised to the laser system. The samples were translated perpendicular to the optical axis (step width 0.5 µm) and after each step a sequence of 3 laser pulses was employed for ablation. The $GaPO_4$ crystals used had a cylindrical shape with diameter 7.5 mm and thickness h ≈ 215 µm. The fundamental resonance frequency was in the range of 5.9 to 6 MHz and the quality factor in air was Q ≈ 95000. After laser-patterning the crystals were cleaned in H_2SO_4 (concentration 98 %, temperature 50 °C, duration 45 min) to remove ablated material that was back-scattered onto the crystal surface. For piezoelectric measurements Au (100 nm) / Cr (40 nm) electrode pads and connecting lines were thermally evaporated on both sides of the crystals. The pads were placed in a face-to-face geometry with one pad in the centre of the laser- ablated region.

For humidity sensing thin polyimide layers were spin-coated onto the non-ablated surface of various $GaPO_4$ and quartz TSM resonators. The thickness of the cured polymer layers was h_s ≈ 0.63 µm for 6 MHz $GaPO_4$ and 10 MHz quartz crystals and h_s ≈ 5.7 µm for the laser-patterned $GaPO_4$ crystal, as determined from profilometry measurements. Polyimide adsorbs water from a humid atmosphere and for commercial Kapton® films (DuPont) the mass density ρ_s at room temperature increases from ~ 1.42 g/cm^3 at 0 %RH to ~ 1.46 g/cm^3 at 100 %RH [9]. The impedance spectra of coated and non-coated $GaPO_4$ and quartz resonators were recorded at room temperature at levels of relative humidity ranging from 0 to 95 %. The accuracy of the humidity level setting was ± 1 %.

3. RESULTS AND DISCUSSION

3.1 F_2 - laser ablation of $GaPO_4$

Figure 1 shows single crystal and ceramic $GaPO_4$ patterned by F_2 - laser ablation employing mask projection. The laser fluence employed for the single crystalline and ceramic material was Φ = 1630 mJ/cm^2 (number of laser pulses N_ℓ = 200) and Φ = 1830 mJ/cm^2 (N_ℓ = 800), respectively. Well-defined ablation craters with sharp edges between ablated and non-ablated areas and smooth surfaces at the bottom and the side walls of the holes were achieved at these fluence levels. Ablation of $GaPO_4$ at 1000 mbar N_2 background pressure produced very similar patterns. Backscattering of ablated species due to the high pressure lead to the formation of debris near the processed area.
The strong absorption of F_2 - laser radiation in $GaPO_4$ and the sharp boundaries of ablated areas enable the fabrication of surface microstructures. Figure 2 shows the surface of single crystal $GaPO_4$ after micro-patterning using an electron microscopy grid as projection mask (Φ = 1800 mJ/cm^2, N_ℓ = 200).

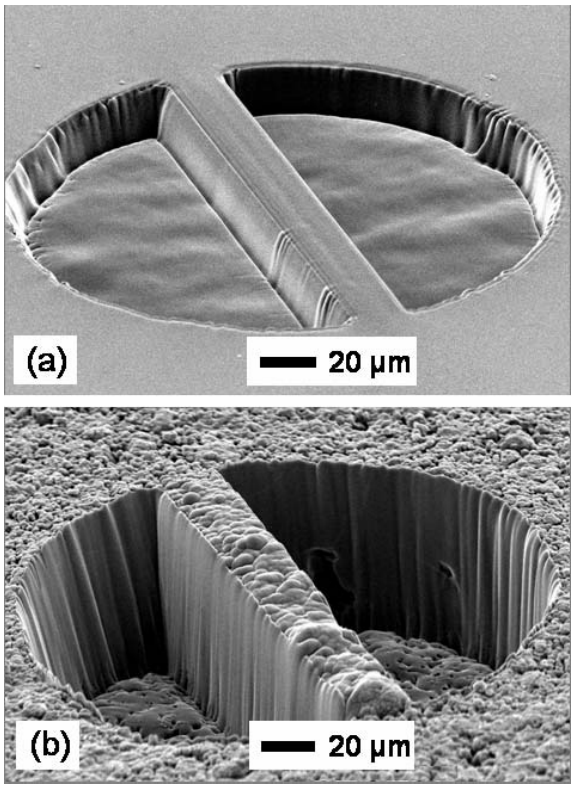


Fig.1. SEM micrographs of single crystalline (*a*) and ceramic (*b*) $GaPO_4$ patterned by F_2 - laser ablation employing a projection-mask technique. [3]

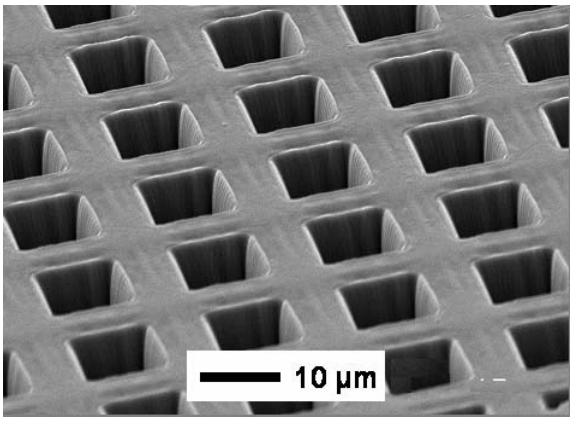


Fig. 2. SEM micrograph of a micro-patterned $GaPO_4$ single crystal surface. A regular array of quadratically shaped holes of size 11 × 11 µm^2 was produced. The depth of holes was 19 ± 1 µm [3].

Figure 3 shows the ablation rate W_A of single crystal and ceramic $GaPO_4$ for different F_2 - laser fluences and number of laser pulses. Experiments were performed in 1000 mbar N_2 (solid symbols) and 0.03 mbar Ar (open symbols) gas background. The ablation rates were determined from the depth of ablated holes measured by a mechanical profilometer and/or an optical microscope and the number of laser

pulses employed. The error bars account for shot-to-shot fluctuations of the pulse energy and uncertainties of spot size and hole depth measurements. Efficient material removal occurred for $\Phi > 200$ mJ/cm^2 and the ablation rates monotonously increased with increasing fluence.

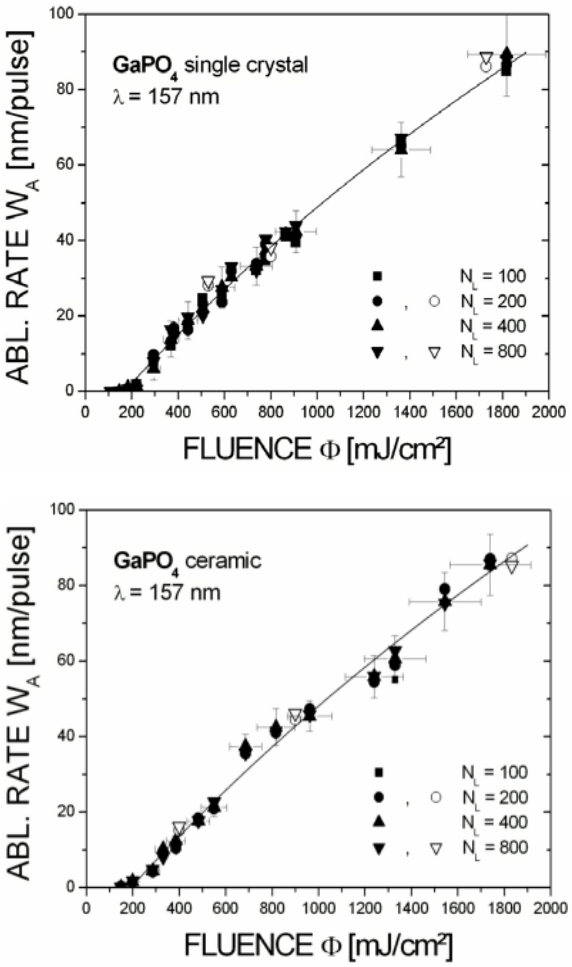


Fig. 3. Ablation rates of single crystal and ceramic $GaPO_4$ for different laser fluence and number of laser pulses. The experiments were performed in 1000 mbar N_2 (solid symbols) and in 0.03 mbar Ar (open symbols) gas background. The solid lines represent fits to the averaged ablation rates [3].

At low fluence, $200 \le \Phi \le 600$ mJ/cm^2, the rates approximately showed a linear dependence on fluence, whereas deviations from a linear behavior were observed at higher fluence values. The ablation rates obtained for single crystalline and ceramic $GaPO_4$ were surprisingly similar and did not essentially depend on the gas background employed. The similarity of rates for the different materials may be due to the small ablated depth per pulse as compared to the typical $GaPO_4$ grain size and the phase purity of the ceramic (x-ray diffraction analysis). At high background pressure, reduced ablation rates and the formation of debris are frequently observed [1]. F_2 - laser ablation of $GaPO_4$ in 1000 mbar N_2 and 0.03 mbar Ar background at a similar laser fluence and the same number of pulses, however, produced ablation craters of comparable depth.

Different processes including thermal, photochemical, photophysical, mechanical and defect-related mechanisms may contribute to nanosecond - laser ablation [1]. For purely photothermal or purely photochemical processes and one-dimensional expansion of the vapor plume, the thickness of the ablated layer per pulse, Δh, can be described by

$$\Delta h = \frac{1}{\alpha_p} \ln\left[1 + \alpha_p B(\Phi - \Phi_{th})\right] \qquad (1)$$

The meaning of α_p and B depends, of course, on the particular ablation mechanism. At very low fluences, effects related to plasma shielding may be ignored ($\alpha_p \Delta h << 1$) and a linear dependence, $\Delta h \approx B\,(\Phi - \Phi_{th})$, is obtained. Equation (1) was used to fit the measured ablation rates averaged over different numbers of laser pulses. The ablation rates could be well fitted in the entire range $200 \le \Phi \le 2000$ mJ/cm^2 (solid lines in Figure 3). The threshold values derived from the fits were $\Phi_{th} = 167 \pm 20$ mJ/cm^2 and $\Phi_{th} = 190 \pm 32$ mJ/cm^2 for single crystalline and ceramic $GaPO_4$, respectively. Linear fits to the data in the low fluence range (200 – 600 mJ/cm^2) gave very similar threshold values. However, linear fits to all data measured in the entire fluence range resulted in threshold values below 100 mJ/cm^2 . The thickness of the ablated layer per pulse at the highest fluences employed, $\Delta h \approx 90$ nm, is roughly comparable to the thermal penetration depth, $l_{Th} = (D \times \tau_\ell)^{1/2} \approx 201$ nm [10] and significantly larger than the optical penetration depth, $l_\alpha \approx 26$ nm, as estimated from reflectivity measurements. Thus, F_2 - laser ablation of $GaPO_4$ is likely to be dominated by photothermal mechanisms.

3.2 Fabrication of HFF $GaPO_4$ resonators

Figure 4 shows micrographs of laser-patterned TSM $GaPO_4$ resonators. In the optical image the patterned region is marked by dashed lines (Fig. 4a). The size of this region was A ≈ 4.2 mm × 2.4 mm and the thickness of the laser-ablated layer was $\Delta h \approx 100$ µm. The average depth etching rate and the volume removal rate were ~ 315 µm/h and ~ 3.2 mm^3/h, respectively.

The crystal was optically transparent after laser processing. The evaporated electrode pad and line

are shown in black in the optical image. The electron microscopy images (Figs. 1b and 1c) revealed a homogeneous depth within the laser - thinned region and smooth edges due to the scanning procedure employed. At the bottom surface in the ablated area a line-like pattern oriented parallel to the scan direction was observed. The rms amplitude of the surface undulation was ~ 1 µm (scan range of the profilometer measurement 1.5 mm).

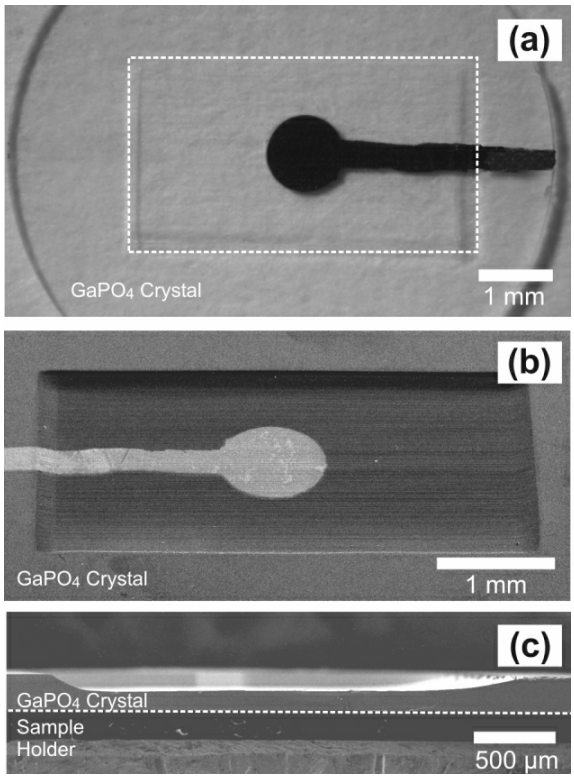


Fig. 4. Optical micrograph of a laser-patterned TSM $GaPO_4$ resonator (a). The thinned region had a rectangular shape and was produced by F_2 - laser ablation (marked by white dashed line). The evaporated electrode pad and line are shown in black. Electron microscopy images of the surface (b, angle of view 45°) and the cross section (c) of the laser-patterned $GaPO_4$ resonator [11].

The microscopic surface roughness as measured by AFM was $\Delta z_{rms} \approx 30$ nm (scan range 40 × 40 µm^2). The formation of the line-like pattern was probably related to non-uniform laser intensity in the spot and could be avoided by using 2D-scan modes or homogenised beam profiles. Wet-chemical etching of $GaPO_4$ was reported to produce smoother surfaces with a roughness of ~ 0.1 µm (etch depth ~ 75 µm, etch rate ~ 5 µm/h, T < 50 °C) [8].

3.3 Piezoelectric properties of laser-processed $GaPO_4$

The laser-processed $GaPO_4$ resonators were piezoelectric and thickness-shear-modes were excited in the thinned areas. Figure 5 shows the impedance spectrum of the patterned crystal. The resonance indicated the excitation of an HFF thickness-shear-mode at a minimum impedance frequency of ~ 11.62 MHz. The quality factor was Q ~ 7000 (in air). The impedance resonances at higher frequencies were due to the excitation of spurious modes (inset in Figure 5).

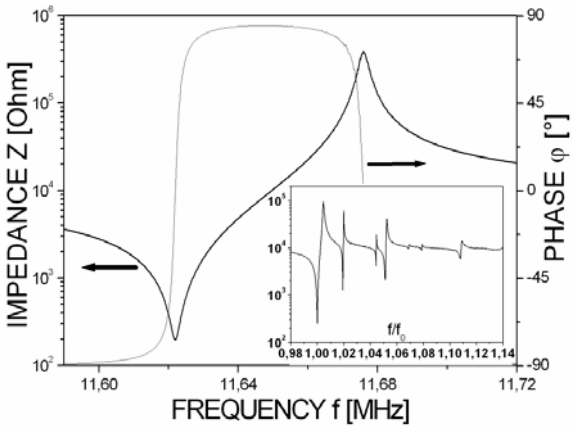


Fig. 5. Impedance spectrum of laser-patterned $GaPO_4$ resonator. The TSM fundamental frequency was increased from ~ 5.93 MHz to ~ 11.62 MHz due to the reduced crystal thickness. Spurious modes were excited at significantly higher frequencies (inset) [11].

The TSM eigen-frequencies f_{nmp} depend on material parameters and the resonator geometry. For thin resonators the fundamental mode frequency is inversely proportional to the crystal thickness [12]. The higher frequency for the laser-patterned crystal (1.96×) correlated reasonably well with the reduced crystal thickness (1.87×). The lower Q factor of the HFF $GaPO_4$ was probably related to the surface roughness and non-uniform thickness of the laser-ablated crystal. An improved laser-patterning procedure could probably reduce the surface roughness and lead to higher Q values. The higher Q value of wet-chemical etched HFF $GaPO_4$ (Q ~ 30000, air-gap measurement) might be due to the smoother surface of such samples [8].

3.4 HFF $GaPO_4$ resonators for humidity sensing

The $GaPO_4$ and quartz crystals revealed a decrease of parallel and serial resonance frequencies and quality factors after coating with the polyimide layer and due to adsorption of water in the polymer. The frequency shift and Q value measured in dry atmosphere were $\Delta f \approx -256$ kHz and Q ~ 5000 and $\Delta f \approx -7.3$ kHz and

Q ~ 90000 for the polymer coated 12 MHz and 6 MHz $GaPO_4$ crystals, respectively. With increasing level of humidity the frequencies and Q factors were decreased further (Figure 6).

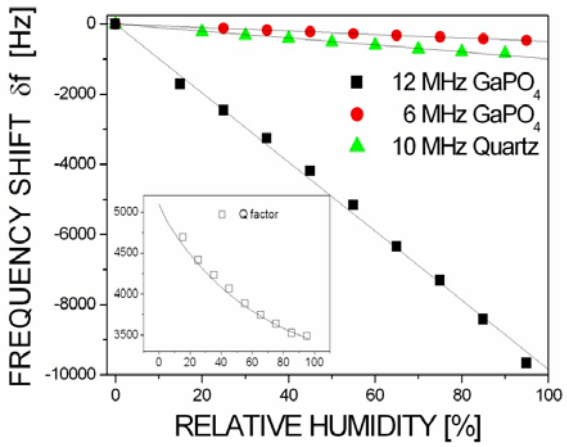


Fig. 6. Shift of serial resonance frequency (δf) of polymer coated $GaPO_4$ and quartz thickness-shear-mode resonators with increasing level of relative humidity at room temperature. The Q factor of the laser-patterned 12 MHz $GaPO_4$ resonator decreased with increasing humidity (inset) [11].

From linear fits to the measured frequency shifts, $\delta f = -S_{RH} \times \%RH$, the sensitivity S_{RH} was determined. The values obtained were $S_{RH} = -98 \pm 1$ Hz / %RH, $S_{RH} = -4.9 \pm 0.01$ Hz / %RH and $S_{RH} = -9.9 \pm 0.2$ Hz / %RH for the 12 MHz $GaPO_4$, the 6 MHz $GaPO_4$ and the 10 MHz quartz resonators, respectively. A smaller frequency shift was observed for non-coated crystals ($|\delta f| \leq 30$ Hz at 80 %RH for 5 MHz quartz resonators). The decrease of the HFF $GaPO_4$ quality factor with increasing humidity could be fitted by an exponential function (solid line in the inset of Figure 6).

The frequency shifts due to the polymer coating (Δf) and the adsorption of water (δf) may be ascribed to mass loading of the TSM transducers. Mass attachment reduces the TSM frequency and for thin rigid coatings a frequency shift proportional to the attached mass was derived by Sauerbrey [13]. The total frequency shift of the polyimide - coated transducers may be approximated by

$$\Delta f + \delta f \approx -c_f h_S \rho_s \left(1 + d_S \frac{\%RH}{100}\right) \qquad (2)$$

and depends on the transducer frequency f and material parameters, $c_f = 2 \times f^2 / \sqrt{(c_{66} \times \rho)}$, and the thickness and density of the polymer layer. A linear increase of the polyimide layer mass density ρ_S with increasing humidity was assumed (parameter d_S). The transducer material properties were $c_{66} = 22.4$ GPa and $\rho = 3570$ kg/m^3 and $c_{66} = 40$ GPa and $\rho = 2650$ kg/m^3 for $GaPO_4$ and quartz, respectively [4, 14]. The calculated frequency shifts due to the polymer attachment, $\Delta f \approx -244$ kHz (12 MHz $GaPO_4$) and $\Delta f \approx -7.0$ kHz (6 MHz $GaPO_4$), were in good agreement to the measured values. The estimated increase in layer density $d_S = 100 \times S_{RH} / \Delta f$ was 3.8 %, 6.7 % and 5.7 % for the 12 MHz $GaPO_4$, 6 MHz $GaPO_4$ and the 10 MHz quartz transducers, respectively. These values were close to the change in density of commercial polyimide foils ($d_S = 2.8$ %, [9]). For a more accurate description of the humidity sensor response including frequency shifts and changes of the Q factor the viscoelastic properties of the polymer layer should be taken into account [15].

4. CONCLUSIONS

The DUV radiation of F_2 - laser has been successfully employed for micro-patterning of $GaPO_4$ (band gap energy 7.1 eV). The high photon energy (7.9 eV) of such lasers allows to remove in a controlled way thin surface layers of $GaPO_4$ material by individual laser pulses.

So-called inverted mesa structures with thickness-shear-mode frequency ~ 12 MHz and quality factor ~ 7000 have been produced by laser-thinning of commercial 6 MHz crystals. $GaPO_4$ and quartz resonators were coated with thin polyimide layers for humidity detection. The crystals revealed a linear decrease of frequency with increasing level of relative humidity. For the same nominal receptor layer thickness the humidity sensitivity of the 12 MHz laser-patterned $GaPO_4$ was more than twice that of the 6 MHz $GaPO_4$ crystal. The laser-processing of ultrathin $GaPO_4$ crystals with resonance frequencies above 100 MHz should be feasible considering the small ablation rates employed (≤ 0.1 µm/laser pulse).

ACKNOWLEDGEMENTS

We would like to thank P.W. Krempl for stimulating discussions and H. Piglmayer-Brezina, A. Mühl-bachler and A. Nimmervoll for technical support. Financial support by the European research project PISARRO (IST-2001-33326) and the Austrian Federal Ministry of Economics and Labour (Micro@Nano – Fabrication Network) is gratefully acknowledged.

REFERENCES

[1] D. Bäuerle: *Laser Processing and Chemistry*, 3rd edition (Springer, Berlin) 2000

[2] R. Delmdahl, Physik Journal **6**, 53 (2003)

[3] M. Peruzzi, J. D. Pedarnig, H. Sturm, N. Huber, D. Bäuerle, Europhys. Lett. **65**(5), 652 (2004)

[4] P.M. Worsch, B. Koppelhuber-Bitschnau, F.A. Mautner, P.W. Krempl, W. Wallnöfer, P. Doppler, J. Gautsch, Mat. Sci.Forum **231-324**, 914 (2000)

[5] A. Janshoff, H.-J. Galla, C. Steinem, Ang. Chemie Int. Edition **39**, 4004 (2000)

[6] B. Zimmermann, R. Lucklum, P. Hauptmann, J. Rabe, S. Büttgenbach, Sensors and Actuators B **76**, 47 (2001)

[7] V.N. Hung, T. Abe, P.N. Minh, M. Esashi, App. Phys. Lett. **81**(26), 5069 (2002)

[8] N. Prudhomme, D. Cachau-Herreillat, P. Papet, O. Cambon, Proceedings of the IEEE International Frequency Control Symposium and PDA Exhibition Jointly with the 17th EFTF, page 688 (2003)

[9] www.dupont.com/kapton/

[10] M. Peruzzi, PhD Thesis, JKU Linz, Austria (2005)

[11] J.D. Pedarnig, M. Peruzzi, H. Salhofer, R. Schwödiauer, W. Reichl, J. Runck, Appl. Phys. A DOI: 10.1007/s00339-004-3186-9 (2005

[12] B. Neubig, W. Briese, *The Crystal Cookbook*, Franzis-Verlag, Feldkirchen (1997)

[13] G. Sauerbrey, Zeitschrift für Physik **155**, 206 (1959)

[14] J.A. Kosinski, Intl. Journal of High Speed Electronics and Systems **10**(4), 1017 (2000)

[15] E.J. Uttenthaler, PhD thesis, Universität der Bundeswehr Munich, Germany (2002)

(*) corresponding author:
Fax +43-732-2468-9242
E-mail: Johannes.Pedarnig@jku.at

THEORETICAL ANALYSIS OF DAMPING EFFECTS OF ACOUSTIC WAVES AT SOLID/FLUID INTERFACES USING A MIXED PERIODIC FEA/BEM APPROACH.

A. Reinhardt, A. Khelif, M. Wilm, V. Laude, W. Daniau, V. Blondeau-Pâtissier, G. Lengaigne, S. Ballandras

Institut FEMTO-ST, UMR CNRS UFC ENSMM UTBM 6174, Département LPMO, 32 Avenue de l'Observatoire, F-25044 Besançon Cedex

Abstract

A theoretical description of ideal and viscous fluid media is exploited to address the problem of modelling damping effects of any guided elastic waves at the interface between viscous fluids and solids. It is based on a boundary element method (BEM) using the Green's function of the considered media mixed with finite element analysis (FEA). This approach is used to compute an harmonic admittance from which propagation characteristics can be extracted. The influence of acoustic radiation and viscosity effects on different kind of waves excited on various substrates is reported and discussed.

1. Introduction

Guided Acoustic Waves can be excited at the surface of any solid material, in stratified media or in plates [1]. These waves may exhibit pure elliptic as well as pure shear polarisation (case of isotropic media), but the practical case of wave propagation at the surface of or in anisotropic materials generally yields any combination of wave polarisation, except along given crystal or symmetry axes [2]. Guided acoustic waves are assumed to propagate without any losses, providing a nice opportunity to manufacture low loss devices such as surface acoustic wave filters and resonators. On the other hand, it is also known that best quality factors and/or smallest insertion losses of such devices are obtained using package closed under vacuum to avoid leakage due to acoustic radiation in air. Also, the use of acoustic wave devices for the development of sensors immersed in fluid media has been widely investigated [3]. Theoretical analysis of acoustic wave excitation and propagation under such working conditions requires the adaptation of existing simulation tools to provide a reliable description of the induced effects by the nature of the surrounding medium on the acoustic wave device response.

In the proposed paper, a Green's function-based description of the acoustic behaviour of fluid media [4] is inserted in an harmonic admittance (HA) computation based on finite element analysis (FEA) [5]. The mathematical developments required in that matter are briefly recalled. Ideal as well as viscous fluids in the limit of Newtonian fluid assumption are simulated. Many cases can be considered, for instance attenuation of Rayleigh and leaky waves due to radiation in fluids but also water damping of Lamb waves. The robustness of the wave propagation is examined when the substrate is in contact with non conductive viscous media. The theoretical results then are discussed. The specific situation of pure shear waves as used in Love wave and in acoustic plate mode (APM) devices also is regarded. These waves are generally assumed poorly affected by water and other weakly viscous fluids. The limit of this hypothesis is examined theoretically.

2. Theoretical fundaments

The theoretical representation of acoustic waves in fluids is usually performed using a pressure formulation. Nevertheless, in order to easily derive the corresponding Green's function, a displacement formulation can be constructed as well. For any fluid, the independent elastic constants required for such a formulation reduce to one, i.e. C_{11} which is also equal to C_{12}, yielding C_{66}=0 consequently. According to [6], a shear effect in a fluid between a moving solid and a reference solid results in a linear stress proportional to the velocity gradient via a coefficient written η called shear viscosity or absolute viscosity of the fluid. The pressure is proportional to the displacement gradient via the compressibility factor. In a very general approach [6], one should also consider a compressive viscosity factor written ζ. For Newtonian fluids, we consider that the stress tensor T_{ij} is given by :

$$T_{ij} = -(P + j\omega\left(\frac{2}{3}\eta - \zeta\right)S_{kk})\delta_{ij} + 2j\omega\eta S_{ij} \qquad (1)$$

where P represents the pressure in the fluid which is proportional to the displacement divergence. The compressive viscosity generally is not evaluated for fluid property characterisation (see for instance [7]) and consequently it has been set to zero in this work. One has to recall that this may be a weakness of the proposed model. According to certain authors [8,9], the compressive viscosity should not be neglected since it significantly contributes to wave attenuation in fluids. However, this parameter must be considered as a correction factor that will not affect the fundaments of the method, but yields less confidence in the results obtained for generally polarized waves than for mostly or pure shear polarized waves. Equation (1) is used to derive the Green's function of the fluid according to the Fahmy-Adler's formulation, as exposed in ref. [4]. We only recall here the main result consisting in the following eigen-mode relation :

$$\frac{\partial}{\partial x_2}\begin{Bmatrix}\sigma_{21}\\ \sigma_{22}\\ \sigma_{23}\\ u_1\\ u_2\\ u_3\end{Bmatrix} = \begin{bmatrix}[\alpha_{11}] & [\alpha_{12}]\\ [A_{22}]^{-1} & [A_{22}]^{-1}(j\omega s_1{}^t[A_{12}] + j\omega s_3[A_{23}])\end{bmatrix}\begin{Bmatrix}\sigma_{21}\\ \sigma_{22}\\ \sigma_{23}\\ u_1\\ u_2\\ u_3\end{Bmatrix} \qquad (2)$$

with

$$[\alpha_{11}] = [(j\omega s_1[A_{12}] + j\omega s_3{}^t[A_{23}])[A_{22}]^{-1}]$$

$$[\alpha_{12}] = \begin{bmatrix} j\omega\rho[I] - j\omega s_1\left([A_{12}][A_{22}]^{-1}\left(j\omega s_1{}^t[A_{12}] + j\omega s_3[A_{23}]\right) + j\omega(s_1[A_{11}] + s_3[A_{13}])\right) \\ -j\omega s_3\left({}^t[A_{23}][A_{22}]^{-1}\left(j\omega s_1{}^t[A_{12}] + j\omega s_3[A_{23}]\right) + j\omega(s_1[A_{13}] + s_3[A_{33}])\right) \end{bmatrix}$$

where the sub-matrices A_{ij} are frequency dependent and combined the acoustic and viscosity properties of the considered fluid [4]. As shown in [10], the Green's function of the medium relating the stress and displacement fields can be derived from (1). This Green function can be used for computing an effective permittivity at the interface between the wave guide and the fluid medium, allowing for the calculation of an HA following Blotekjaer's approach [4]. In this work, it is inserted in a boundary element formulation also derived at the fluid/solid interface, but coupled with a FEA [5] dedicated to periodic wave-guides. This approach is particularly efficient to address problems with undefined bounds (radiation in semi-infinite domains) together with inhomogeneous radiation interfaces (that cannot be described via plane boundaries). According to the usual FEA digitalizing procedure, the problem to solve results in a linear algebraic system as follows :

$$\begin{bmatrix} {}^t C_u^*(\gamma) & 0 \\ 0 & C_\phi^*(\gamma) \end{bmatrix} \left(\begin{bmatrix} K_{uu} - \omega^2 M_{uu} & K_{u\phi} \\ K_{\phi u} & K_{\phi\phi} \end{bmatrix} - \begin{bmatrix} X_{uu}(\omega,\gamma) & 0 \\ 0 & X_{\phi\phi}(\omega,\gamma) \end{bmatrix} \right) \times (3)$$

$$\begin{bmatrix} C_u(\gamma) & 0 \\ 0 & C_\phi(\gamma) \end{bmatrix} \begin{Bmatrix} v \\ \varphi \end{Bmatrix} = {}^t \begin{bmatrix} C_u^*(\gamma) & 0 \\ 0 & C_\phi^*(\gamma) \end{bmatrix} \begin{Bmatrix} F \\ Q \end{Bmatrix}$$

where the matrices K_{uu}, $K_{u\phi}$ and $K_{\phi\phi}$ respectively hold for the purely mechanical, piezoelectric and dielectric stiffness and Muu is the mass matrix. Matrices X_{uu} and $X_{\phi\phi}$ correspond to acoustic and dielectric radiation in the open medium. Matrices C are inserted to account for the periodicity of the problem. Also note that the combination of radiation in the fluid domain as well as in the substrate is accessible to computation. Consequently, only the inhomogeneous part of the problem has to be meshed, yielding a significant increase of precision together with optimised computation delays. It also possible to simulate 3D devices, even if we only investigate her the case of 2D structures.

3. Computation results

The above-described model now is exploited to compute the behaviour of infinitely periodic 2D wave-guides. It allows for the theoretical characterization of guided waves via an HA. The obtained results then cannot prejudge of the effective operation of a real device, but they provide useful indications of practical tendencies and behaviours of the corresponding structures submitted to viscous fluid loads.

3.1. Lamb waves

We first examine the case of Lamb waves on singly-rotated quartz for the sake of simplicity, recalling that anti-symmetrical Lamb waves are poorly excited on such subtrates (bimorph structures are more favourable of course). The plate structure is well-known as one of the most exciting configuration for liquid sensing, since one side is used to excite the wave, the other side being left for the application of the fluid to be tested. We consider the AT-cut (YXl)/36° excited using an inter-digital transducer (IDT) of period 40 µm (arbitrarily fixed), assuming operating frequencies in the range 10-50 MHz. The plate thickness then varies from 5 to 20 µm for a working frequency defined between 5 and 20 MHz. Since the analysis is linear, one can easily transpose the reported results to any dimension, keeping in mind that very thin plates generally are very difficult to handle and process. We then have chosen the plate thickness remembering that the limit operating frequency is provided by the ratio of the wave propagation in the fluid divided by the wavelength (imposed by the device), i.e. 18,75 MHz assuming the phase velocity of pressure wave in water equal to 1500 $m.s^{-1}$.

Figure 1 shows the deformed mesh of the mode under water load, clearly illustrating its anti-symmetrical nature. It exhibits natural directional properties in air, but only the resonance at the beginning of the stop-band still persists once the backside loaded with fluids.

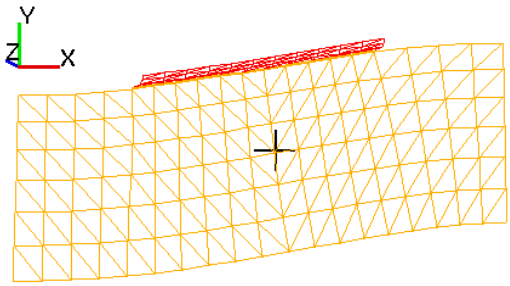


Fig.1 Anti-symmetrical Lamb wave on AT quartz loaded by water on the backside (the top side supports the IDT)

The following curve shows the evolution of the HA corresponding to the mode contribution for different values of viscosity.

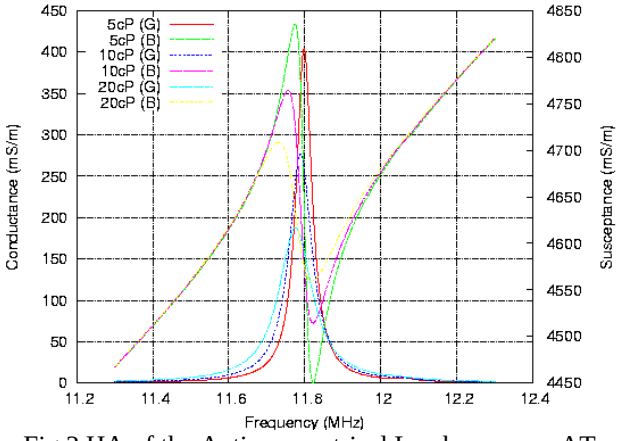


Fig.2 HA of the Anti-symmetrical Lamb wave on AT quartz on a 10 µm thick plate for different values of viscosity (5, 10 et 20 cP.)

Figure 2 indicates that the mode is triply sensitive to viscosity since its quality coefficient, its amplitude as well as its resonance frequency decrease along increasing viscosity values. The strong shift of the resonance frequency compared to the one found in air mainly is due to the mass loading effect due to the application of the

fluid loading condition at the plate bottom. As shown in fig. 3, the dependence of the amplitude and of the resonance frequency versus viscosity is not linear, even if the small number of points does not help in deriving accurate tendencies.

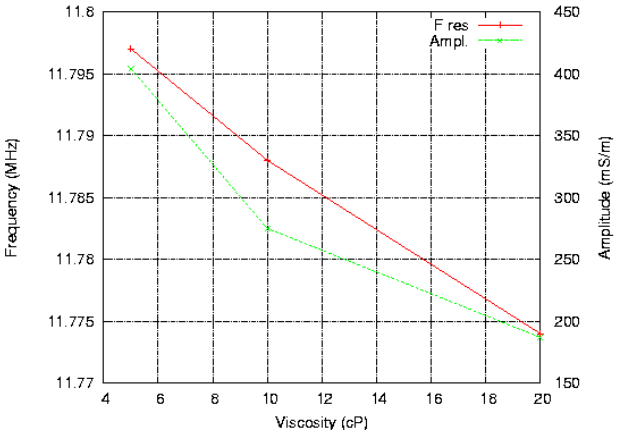


Fig.3 Evolution of the resonance frequency and of the harmonic conductance amplitude versus fluid viscosity for the 10µm thick AT quartz plate

3.2. Micro-machined Ultrasonic Transducers (MUT)

This structure consists in a very thin Si_3N_4 or Si membrane set above a thin electrostatic gap. The excitation is performed by electrostatic forces. It is designed to exhibit the weakest acoustic impedance together with the highest electromechanical coupling to provide an efficient wide band working regime of the device for acoustic imaging applications [11].

Our calculations [15] that there is always a mode guided along the structure which does not radiate energy in the fluid for an alternated +V/-V excitation. This mode is very interesting because contrarily to the usual approach consisting in using a well defined mode in air assumed to persist even in contact with water because of its polarization for instance, it exists only when the cMUT is in contact with water and adequately excited. In this work, we have checked its robustness versus viscosity. We then have simulated the behaviour of the devices for different operating conditions (vacuum, ideal water, viscous water). We assume the simplified plane capacitor assumption for the simulation of the membrane excitation, and we then report the front side velocity versus frequency. Figure 4 shows the implemented mesh for modelling the cMUT. It is assumed to radiate on its back side into silicon, and on the top side in fluids (when loaded).

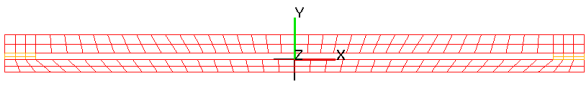


Fig.4 mesh of a cMUT
(backside radiates in Si, topside in fluid)

The following figure shows the influence of the water load on the front side velocity, compared to the device response in vacuum. One has to note that the resonance of the loaded water does not correspond to the one in vacuum but consists in an in-plane wave propagation exhibiting an evanescent behaviour in water.

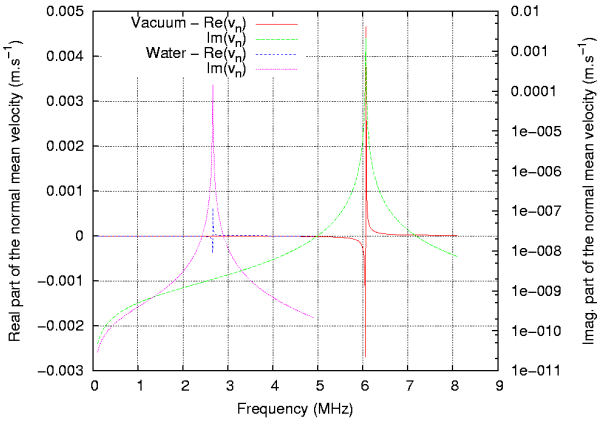


Fig.5 Effect of the water load on a cMUT excited with alternated potential +V/-V

Figure 6 shows the dramatic effect of viscosity that strongly damps the wave propagation. This mode then may be used only with very low viscosity fluids according to our computation.

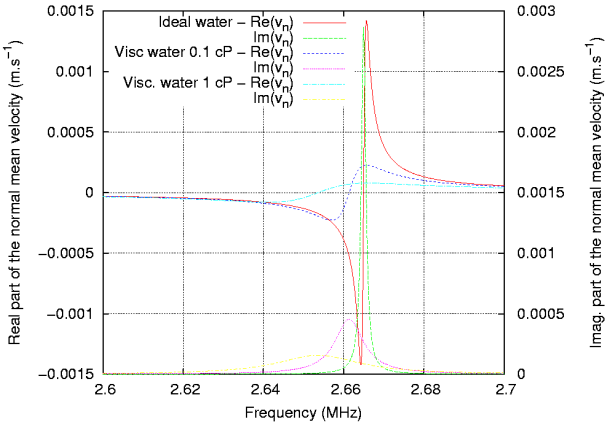


Fig.6 Influence of the viscosity on the cMUT guided mode

3.3. Transverse plate modes

Transverse plate modes have been widely studied because of the possibility to dissociate the excitation zone (IDTs deposited on one side) and the area exposed to the loading medium. Their polarisation a priori is favourable for immersed applications. Finally, one can achieve temperature compensated high quality resonance with those devices (QF in the vicinity of 10^{13} at 100 MHz) [2]. We consider a 50 µm thick plate assuming a 40 µm period (a/p=0,5), yielding resonance close to 60 MHz. Such a configuration also allows for operating near 100 MHz, using higher order modes. However, we focalise here on the first modes for conciseness. Each exhibits 2 resonance respectively corresponding to the beginning and end of the stop band, generally representative of directivity properties [3]. Briot [2] has demonstrated that the corresponding vibrations are of identical nature but the energy distribution is located on one or the other plate surface. One of these modes then should be more robust to the water loading because less damped by the presence of the fluid.

Figure 7 shows the sensitivity of the first mode to viscosity effects. The mass loading effect of water is simulated assuming a very small value for the viscosity. The obtained responses exhibit larger amplitudes than the Lamb wave devices because the modes are more efficiently excited in that latter case. One can note that such modes seem to be more resistant to the viscosity effect, yielding less attenuation than in the other regarded cases.

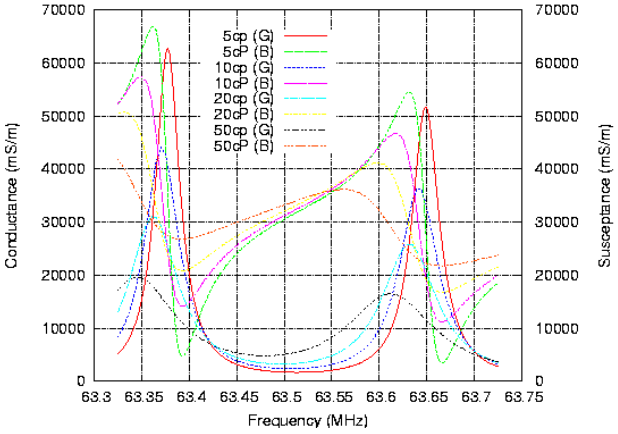


Fig.7 Viscosity effects on the admittance of the 1st pure shear mode on a 50µm thick AT (YXlt)/36/90 cut of quartz

3.4. Love waves

Love waves exhibit a pure shear polarisation and are guided by a thin layer for which the phase velocity of the shear propagation is smaller than the one in the substrate on which it is deposited. Such wave-guide then presents dispersive properties together with a passivated surface. The (YXlt)/36°/90° cut of quartz is well known for its capabilities to support surface transverse waves (STW). It is also very well suited for Love waves if one deposit a fused layer atop. Considering a phase velocity close to 5100 $m.s^{-1}$ in the substrate and 3500 $m.s^{-1}$ for the shear wave in amorphous silica, it is seen that the guiding condition are perfectly fulfilled. We use the mesh of fig.8 to simulate the device using our periodic harmonic computation. The deposition of silica is assumed to be achieved using pressure enhanced chemical vapour deposition (PECVD), which is conformal, explaining the thickening of the mesh above the electrode location. Two quasi-fluid area have been added at each edge of the period to flatten the top interface (to well meet the computation conditions). The HA of the structure is first computed in vacuum (fig.9), showing the arising of two guided modes just above the threshold velocity corresponding to acoustic radiation in the bulk.

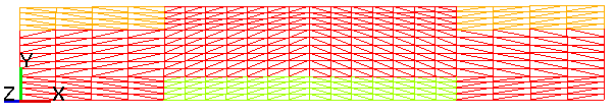


Fig.8 Mesh of the Love wave structure assuming a conformal deposition of the fused silica atop aluminium electrodes

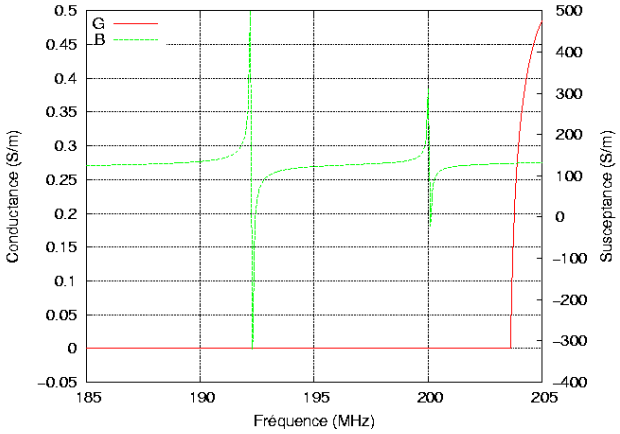


Fig.9 HA of the Love wave guide on (YXlt)/36°/90° quartz convered with a conformal silica layer (h_{silica}=1,5µm, p=12,5µm, a/p=0,5, $h/2p$=2%)

As in the case of transverse plate modes, these two contribution to the HA are of same nature, but they correspond to the 2 ways to satisfy the boundary conditions : a symmetrical mode (the electrode edges vibrate in phase, max vibration at the electrode centre) and an anti-symmetrical mode (electrode edge in phase opposition, no vibration at the electrode centre).

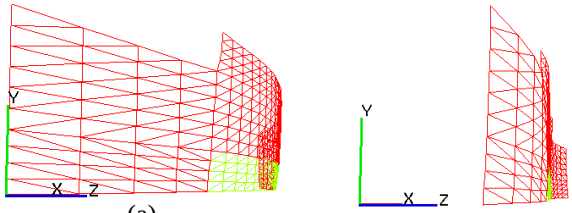


Fig.10 Deformed mesh corresponding to the begining and the end of the stop band of the Love wave resonator (a) Symmetrical mode (b) anti-symmetrical mode

Shear waves do not transfer energy in ideal (i.e. non viscous) fluids. In a 2D representation, no moving part can transmit motion in the fluid since no shear effects are described in the usual pressure formulation of wave propagation in fluids. We now introduce the radiation condition in a viscous fluid. Figure 11 shows the influence of the viscosity on the HA contribution of both modes of the structure for viscosity values respectively equal to 1 et 10 cP.

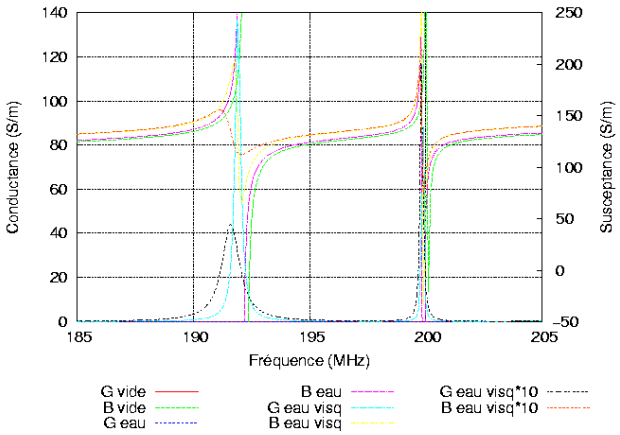


Fig.11 Love wave guide HA for different values of viscosity

One can clearly see that the first resonance is more affected by the viscous fluid than the one at the end of the stop-band. One can imagine that in that case, shear effects in the fluid at the electrode edges may cancel each other because of their phase opposition. We then have investigate the case of a non conformal silica deposition, yielding a flat silica surface (obtained by polishing the conformal deposition for instance). The HA of such a structure shows that both resonance are now closer one to the other (the stop-band is dramatically reduced). Furthermore, we do not preserve the previous behaviour of the end of the stop-band since both exhibit almost the same sensitivity to viscosity. Above 10 cp, it is not possible to distinguish one mode from the other. This result tends to indicate that the shape of the grating can significantly help in the design of robust structure dedicated to immersed applications, able to provide sufficient signals for any analysis processing.

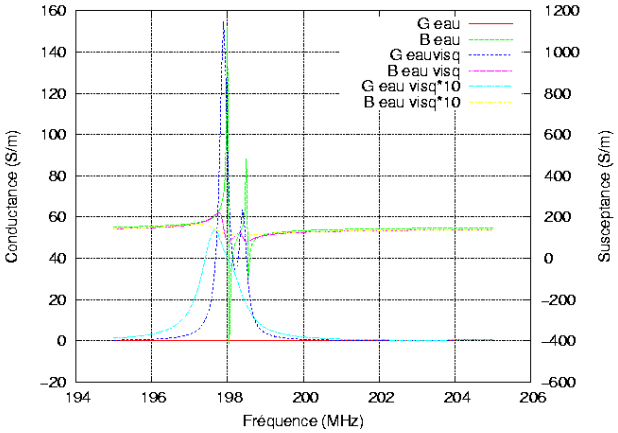


Fig.12 HA of a non conformal (flatten) wave guide grating for different viscosity values

4. Conclusion

We have developed a model allowing for the simulation of acoustic wave guides loaded with fluids exhibiting viscosity properties (in the limit of Newtonian fluid assumption). We have derived a Green's function from these developments, enabling us to compute a boundary element contribution mixed with finite element to predict the influence of the fluid on the harmonic admittance of various wave-guide. Even if bulk viscosity was neglected here, we have shown that flexural waves (Lamb waves, cMUT guided wave) are dramatically affected by the viscosity. Also pure shear polarised wave such as transverse plate modes or Love wave exhibit a relative robustness to viscous effects. These results tend to moderate the confidence in acoustic devices operating under fluid loads. However, they also show that a deep analysis of mode distribution in different wave-guides, particularly taking into account the wave-guide structure (electrode gratings, grooves, passivation layers, etc.), can significantly help to improve the operation of such devices loaded by fluids.

Acknowledgement : This work was partly supported by TEMEX.

References :

[1] S.Ballandras, A. Reinhardt, V. Laude, A. Soufyane, S. Camou, W. Daniau, T. Pastureaud, W. Steichen, R. Lardat, M. Solal, P. Ventura, "Simulations of surface acoustic wave devices built on stratified media using a mixed finite element/boundary integral formulation", J. Appl. Phys., Vol. 96, n°12, pp. 7731-7741, 2004

[2] D. Royer, E. Dieulesaint, *Ondes élastiques dans les solides*, Masson Ed., 1997

[3] In *Sensors, a comprehensive survey*, (Vol.2/3 and 7) W.Göpel, J.Hesse, J.N. Zemel Eds., VCH, 1996

[4] S. Ballandras, A. Reinhardt, A. Khelif, M. Wilm, V. Laude, W. Daniau, V. Blondeau-Pâtissier, W. Boireau, "Theoretical analysis of damping effects of SAW at solid/fluid interfaces", Proc. of the joint EFTF - IEEE Int. Freq. Cont. Symp., Tampa Bay, pp. 907-910, 2003

[5] S. Ballandras, V. Laude, Th. Pastureaud, M. Wilm, W. Daniau, A. Reinhardt, W. Steichen, R. Lardat, "A FEA/BEM Approach to Simulate Complex Electrode Structures Devoted to Guided Elastic Wave Periodic Transducers", Proc. of the IEEE Ultrasonics Symp., pp. 309-312, 2002

[6] R.N Thurston, "Wave propagation in fluids and normal solids", in *Physical Acoustics*, Vol. 1A, W.P. Mason Ed., Academic Press, 1964

[7] R.S. Marvin, "Viscosity of liquids", in *American Institute of Physics Handbook*, 3rd Ed., Mc Graw Hill Ed., pp.187-202, 1972

[8] W.P. Mason, "Piezoelectric crystals and mechanical resonators", in *Physical Acoustics*, Vol. 1A, W.P. Mason Ed., Academic Press, 1964

[9] T.A. Litovitz, C.M. Davis, "Structural and shear relaxation in liquids" in *Physical Acoustics*, Vol. 2A, W.P. Mason Ed., Academic Press, 1965

[10] A. Reinhardt, Th Pastureaud, V. Laude, S. Ballandras, "Scattering matrix method for acoustic waves in piezoelectric, fluid, and metallic multilayers", J. Appl. Phys., Vol. 94, pp. 6923-6931, 2003

[11] B.T. Khuri-Yakub, F.L. Degertekin, X.C. Jin, S. Calmes, I. Ladabaum, S. Hansen, and X.J. Zhang, "Silicon micromachined ultrasonic transducers," Proc. of the IEEE Ultrasonics Symposium, pp. 985-991, 1998

FINITE ELEMENT MODEL OF NANOBALANCE IN $GAPO_4$

L. Delmas*, ** , F. Sthal*, E. Bigler*, R. Bourquin*, M. Minunni**, M. Mascini**
*FEMTO-ST, UMR CNRS 6174, dept LCEP, ENSMM
26 chemin de l'Epitaphe, 25000 Besancon-France
** Università degli Studi di Firenze, Dip. di Chimica- Via della Lastruccia 3,
Polo Scientifico 50019 Sesto Fiorentino (Fi), Italy

Biosensors are more and more used in different applications in food, environmental and clinical analysis. Some recent applications have been reported to genetically modified organism detection in food by the use of a developed quartz crystal microbalance sensors. Most of application are related to the used of quartz crystals. Very recently has been introduced gallium orthophosphate ($GaPO_4$) as interesting new piezoelectric material for application to microbalance -based biosensing.

In this work we present a finite element model of a $GaPO_4$ crystal "nanobalance". In our system the crystal is mounted in a metacrylate cell and one side of its surface is in contact with a liquid. In the model the influence of density, viscosity and temperature of the fluid are reported. The equivalent circuit of the sensor is also obtained and compared to an experimental equivalent circuit.

I. Introduction

There has been an increasing interest in field of biosensors with application in different disciplines: food, environmental and clinical analysis [1]. Some recent applications have been reported to Genetically Modified Organism (GMO) detection by the use of Quartz Crystal Microbalance (QCM) sensors [2-5].

$GaPO_4$ is new quartz-homeotypic piezoelectric crystal [6][7]. It has been introduced for application to microbalance - based biosensing. This new crystal has a better coupling factor than quartz. Thus, we can hope to use it as nanobalance and then to improve sensitivity of biosensors.

This paper reports a Finite Element Model (FEM) of biosensor based on microbalance. Firstly, it's recalled the principle of QCM used as a biosensor. Then FEM is built according to the biosensor used during measurement. A comparison between FEM and measurement is done for quartz crystal. Then, $GaPO_4$ application as nanobalance is demonstrated and an electric equivalent circuit is presented.

II. Principle of QCM

QCM principles were first described by Sauerbrey who determined the frequency to mass relationship [8]. This relationship is outlined by the following equation:

$$\Delta f = -2.3\,10^{-6} F^2 \frac{\Delta m}{A} \qquad (1)$$

Where Δf is the variation in the fundamental frequency of the coated crystal, F is the resonant frequeny of the crystal, A is the area of the coated electrode and Δm is the variation in mass due to material binding to the crystal.

The QCM consists of two electrodes on both sides of a thin disk of AT-cut quartz or $GaPO_4$ resonator. AT-cut refers to the orientation of the plate to the crystal structure [9-11]. Due to the piezoelectric properties and crystal orientation of the quartz, the application of a voltage between these electrodes leads to a shear deformation of the crystal.

A resonant oscillation is achieved by including the crystal in an oscillator, where the electrical and the mechanical oscillation are near of the fundamental frequency of the crystal. The fundamental frequency depends upon thickness, chemical structure, shape, and mass of the resonator.

QCM biosensing devices are based on affinity sensing. A specific probe for the target sequence is immobilized on the sensor surface and targets are free in solution. The hybridization reaction between the immobilized probe and the target sequence present in the solution is observed due to a mass variation[12]. It is analyed through a variation of frequency according to Sauerbrey equation (1).

III. Modelization of nanobalance biosensors

Figure 1 represents the principle of affinity biosensor based on QCM devices. Probe is immobilized onto gold electrode surface of QCM. Targets inside the buffer solution are hydridized with probe. Thus it occurs an adding of mass on QCM and a shift of frequency is observed.

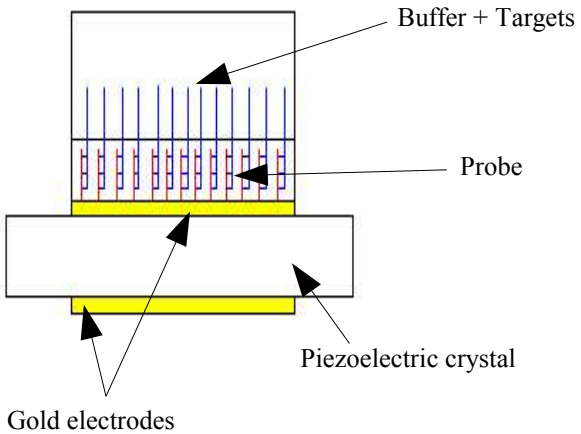


Figure 1: Principle of QCM biosensor.

The FEM consists to modelize each part of the QCM biosensor. The QCM is modelized using piezoelectrical, mechanical and electrical properties [13-16]. Solid or fluid layer can be used to express probe properties. Buffer and targets are modelized with a fluid layer with viscosity, density and concentration of targets inside the buffer.

The modelization is shown in figure 2. Each part of biosensor is represented by different layers with different properties corresponding to element properties.

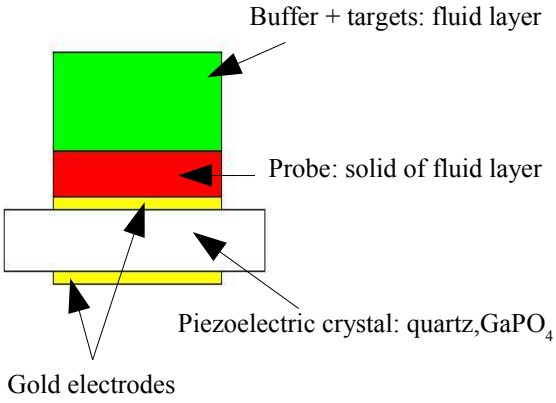


Figure 2: Finite element model of QCM biosensor.

The FEM is built using ANSYS finite element software. The piezoelectric crystal is modelized with Solid5 element. This element takes into account mechanical, piezoelectrical and electrical properties. Electrodes are also modelized with Solid5 but piezoelectric properties are not used. Probes, buffer and targets are modelized using a fluid layer which let taking into account viscosity, density and hybridisation.

IV. Analysis of the biosensor's model

Both modelization, 2D and 3D are made. The first part focus on modal analysis which consists of studying the resonant frequency of the structure. Here the structure is the microbalance with buffer, targets and probes on one side and air on the other one. The microbalance is a resonator at 10 MHz with 10 mm of diameter and 122 µm of thickness for $GaPO_4$ and 180 µm of thickness for quartz.

Figure 3 represents 2D modelization and figure 4 shows 3D model made with Ansys.

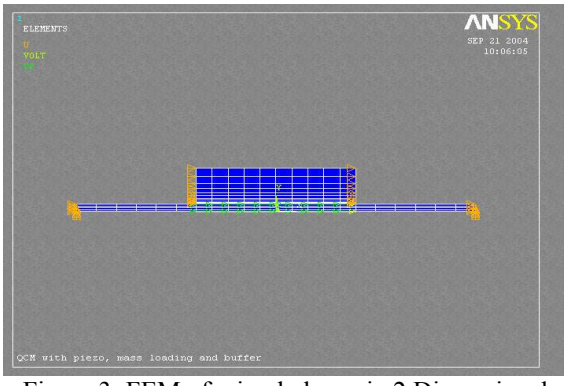


Figure 3: FEM of microbalance in 2 Dimensional.

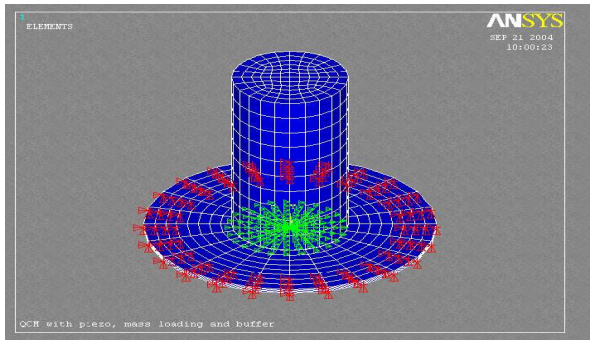


Figure 4: FEM of microbalance in 3 Dimensional.

Using these models and modal analysis, we compare quartz and $GaPO_4$ crystal microbalance. Table 1 shows the evolution of frequency versus mass for both crystals in modal analysis.

Mass	***Frequency (MHz) for quartz***	***Δf (Hz) for quartz***	***Frequency (MHz) for $GaPO_4$***	***Δf (Hz) for $GaPO_4$***
0 in air	9.881732	0	10.134522	0
39.78 ng	9.881724	8	10.134515	7
397.8 ng	9.881656	76	10.134449	73
3.978 µg	9.880908	824	10.133740	782
39.78 µg	9.875964	5768	10.128347	6175
397.8 µg	9.811906	69826	10.058467	76055

Table 1: Comparison between quartz and $GaPO_4$ microbalance for shifting frequency versus mass.

We notice no difference on frequency shift due to mass loading between quartz and $GaPO_4$. However on this assay we analyse only resonant frequency shift of the structure. Then a study of dynamic analysis is required to conclude. The dynamic analysis consists of applying a potential on electrodes and to observe the evolution of the structure. The shift of frequency versus mass is analysed and the equivalent circuit of the resonator is obtained.

Like measurement of experiment, we apply a potential on electrodes and we analyse variation of charges on them. Firstly, we focus on differences between quantity of charges for quartz and $GaPO_4$. At the resonance in air, there is a factor 100 between quartz and $GaPO_4$. This difference changes with mass loading but the charges in $GaPO_4$ are still higher than quartz. Table 2 reports charges' variations for both crystals.

	in air	***M = 0.339 µg***
Quartz	Fres = 9.882 MHz	Fres = 9.881 MHz
	q = 2.25e-7 C	q = 1.95e-7 C
$GaPO_4$	Fres = 10.134 MHz	Fres = 10.1337 MHz
	q = 1.35e-5 C	q = 4.5e-5 C

Table 2: Variation of charges versus mass for quartz and $GaPO_4$.

Secondly, dynamic simulationis achieved. Frequency shift due to mass loading is also analysed. Figure 5a and 5b represent variations of frequency versus mass for quartz and $GaPO_4$, respectively.

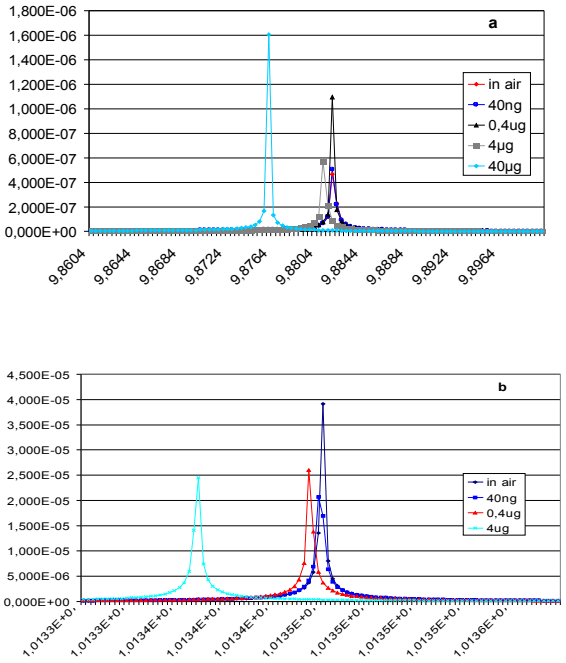


Figure 5a and 5b: Shift of frequency versus mass in dynamic simulation for quartz (a) and for $GaPO_4$ (b).

Instead of $GaPO_4$, we notice that shift frequency for quartz is almost null for small mass loading less than 1 microgramm. So we can see here the using of $GaPO_4$ as a nanobalance. Moreover with this model and dynamic simulation we obtain electric parameters of equivalent circuit of resonator (Table 3).

	quartz	*$GaPO_4$*
In air (M = 0)	C = 24.9468 fF L = 10.3982 mH	C = 119.4186 fF L = 2.0651 mH
M = 39.78 ng	C = 24.9469 fF L = 10.3981 mH	C = 119.4165 fF L = 2.0652 mH
M = 0.3978 µg	C = 24.9481 fF L = 10.3978 mH	C = 119.3948 fF L = 2.0702 mH
M = 3.978 µg	C = 24.9598 fF L = 10.3945 mH	C = 113.5862 fF L = 2.1739 mH
K(%) / Kth(%)	11% / 9.4%	17% / 14.7%
AVL data (in air)		C = 120 fF L = 2 mH K = 13 %

Table 3: Electric parameters of equivalent circuit of resonator for quartz and $GaPO_4$.

Table 3 shows a good agreement between electric parameters obtained with model and measurement made by AVL [17].

V. Comparison between measurement and model results

Some measurements are also made with probe immobilization on quartz crystal microbalance. Figure 6 shows experiment results obtained. At the begining we realize measurement in air (1), then we add water (2), buffer (3), probe solution (4), buffer (5), water (6), buffer(7), water (8) [3]. Buffers are used to compare results obtained with and without immobilized probe. Water is used to flash the surface.

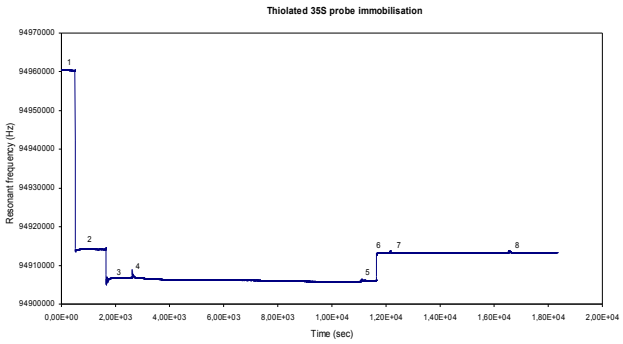


Figure 6: Measurement of Thiolated 35S probe immobilization.

Table 4 summurizes frequency variations of the experiment presented in figure 6. We can see a good correlation on frequency shift order between experiment and finite element model.

	Δf(Hz) measured	*Δf(Hz) simulation*
Δm 3 - 4	951	843
Δm 4 - 5	-279	-85
Δm 3 - 5	672	758

Table 4: Frequency variation versus mass loading for QCM: measured and simulated.

According to this model, we explain temperature effects on measurement. The resonator is compensated in temperature due to the AT-cut orientation. But the fluid and buffer are not compensated in temperature. Thus, variations of temperature influe on measurement due to variation of buffer and water densities.

Temperature (°C)	*Density (kg/m^3)*	*ΔFrequency (Hz)*
10	999.7	
20	998.2	25
30	995.2	75

Table 5: Influences of temperature on density of water.

Table 5 represents influences of temperature on density of water. It's not possible to neglect these temperature variations because frequency shifts are almost the same amplitude as mass loading. In these cases, frequency shifts are: $\Delta f_{10\text{-}20} = 25Hz$ and $\Delta f_{20\text{-}30} = 75Hz$.

VI. Conclusion

The Finite Element Model of Quartz Crystal Microbalance is validated by some measurements. A comparison using FEM between quartz and $GaPO_4$ is also reported.

In this assay, it's shown that $GaPO_4$ has a better sensitivity of mass loading. Nanobalance application with $GaPO_4$ is then demonstrated . However due to this higher sensitivity it's important to take into account some parameters which have less influences on QCM like temperature. Electric device necessary to measure a shift frequency for a nanobalance in $GaPO_4$ is almost the same than QCM. Due to the $GaPO_4$ thermal stability, it's also possible to replace microbalance applications at high temperature by nanobalance using $GaPO_4$ [18-19] .

Acknowledgment

We gratefully acknowledge EU funding of the GAPOGROWTH project under contract n° G5RD-CT-2002-00709 (GROWTH project).

References

[1] E. Luzi, M. Minunni, S. Tombelli and M. Mascini, "New trends in affinity sensing: aptamers for lignad binding", Trends in Analytical Chimistry, vol. 22. No. 11, pp. 810-818, 2003.

[2] S. Tombelli, M. Mascini, A.P.F. Tuner; "Improved procedure for immobilisation of oligonucleotides on gold-coated piezoelectric quartz crystals", biosensors and bioelectronics 17, pp. 929-936, 2002.

[3] I. Manelli, M. Minunni, S. Tombelli, M. Mascini, "Quartz crystal microbalance (QCM) affinity biosensor for genetically modified organisms (GMOs) detection", biosensors and bioelectronics 18, pp. 129-140, 2003.

[4] H. Thanner, P. W. Krempl, W. Wallnöffer and P. M. Worsch, "$GaPO_4$ high temperature crystal microbalance with zero temperature coefficient", Vacuum 67, pp. 687-691, 2002.

[5] M. Minunni, M. Mascini, R. M. Carter, M. B. Jacobs, G. J. Lubrano and G. G. Guilbault, "A quartz crystal microbalance displacement assay for Listeria monocytogenes", Anal. Acta 325, pp. 167-174, 1996.

[6] G. F. Engel and P. W. Krempl, "Stability of alpha-phase of quartz isomorphs", Ferroelectrics, vol. 54, pp. 349-352, 1984.

[7] P. W. Krempl, "Quartz homeotypic gallium orthophosphate: A new high-tech piezoelectric crystal", Ferroelectrics, vol. 202, pp. 65-69, 1997.

[8] G. Z. Sauerbrey, Physik, pp. 155-206, 1959.

[9] F. Krispel, C. Reiter, J. Neubig, F. Lenzenhuber, P. W. Krempl, W. Wallnöffer and P. M. Worsch, "Properties and applications of single rotated $GaPO_4$ resonators", in Proc. IEEE Int. Freq. Cont. Symp. joinly with the 17th EFTF, pp. 668-673, 2003.

[10] J. Detaint, B. Capelle, O. Cambon and E. Philippot, "Properties of AT-cut gallium orthophosphate resonators", Ann. Chim. Sci. Mat., vol. 26, pp. 151-160, 2001.

[11] L. Delmas, F. Sthal, E. Bigler, J. J. Boy, S. Galliou and R. Bourquin, "Experimental study of temperature effects in vibrating beam and thickness-shear resonators of $GaPO_4$ machined by ultrasonic milling", in Proc. IEEE Int. Freq. Cont. Symp., 2004.

[12] Y. Liu, X. Yu, R. Zhao, D. Shangguan, Z. Bo and G. Liu, "Real time kinetic analysis of the interaction between immunoglobulin G and histidine using quartz crystal microbalance biosensor in solution", Biosensors and Bioelectronics 00, pp. 1-9, 2003.

[13] C. Reiter, P.W. Krempl, W. Wallnöffer, P.M. Worsch, "Material properties of $GaPO_4$ and their relevance for applications", Ann. Chim. Sci. Mat., pp. 91-94, 2001.

[14] J. Nosek, "Some second-order effects participating on frequency-temperature behaviour of rotated Y-cut GaPO4 resonators", in Proc. 56th IEEE Int. Freq. Contr. Symp. pp. 337-341, 2002.

[15] C. Reiter, H. Thanner, W. Wallnöffer, P. M. Worsch and P. W. Krempl, "Elastic constants and temperature-compensated orientation of $GaPO_4$", in Proc. 15th EFTF, pp. 50-54, 2001.

[16] R. Bechmann, A. Ballato, T.J. Lakaszek, "High order of temperature coefficient of elastic stiffness and compliances of quartz", in Proc. IRE. Vol.50, pp. 1812-1822, 1962.

[17]P. Krempl, C. Reiter and W. Wallnöffer, Annex 1 of Technical report, pp. 8, Mid-term report, 2002.

[18] P. W. Krempl, G. Schleinzer and W. Wallnöffer, "Gallium orthophosphate $GaPO_4$: a new piezoelectric crystal material for high-temperature sensorics", Sensors and Actuators A 61, pp. 361-363, 1997.

[19] H. Thanner, P. W. Krempl, F. Krispel, C. Reiter and R. Selic, "High temperature microbalance based on $GaPO_4$", Ann. Chim. Sci. Mat. Vol.26, pp. 161-164, 2001.

Quartz tuning fork vibration amplitude as a limitation of spatial resolution of shear force microscopes

J.-M Friedt, É. Carry, Z. Sadani, B. Serio, M. Wilm, S. Ballandras
FEMTO-ST, 32 avenue de l'Observatoire, 25040 Besançon Cedex, FRANCE

Abstract

We compare experimental and finite element modeling of the vibration amplitude of tip-loaded quartz tuning forks. We demonstrate under which experimental conditions the vibration amplitude might become a limiting factor of the lateral resolution of a shear force microscope. Namely, with large excitation amplitudes (> 1 V_{pp}) and under the condition of working at maximum amplitude (as opposed to minimum phase) the displacement amplitude of the tip can be greater than 100 nm, beyond the usual spatial resolution aimed at by scanning probe microscopes (SPM).

1 Introduction

Since Karrai *et al.* [1, 2] showed that quartz tuning forks can be used as sensitive force sensors [3, 4] for scanning probe microscopies [5], many experiments have been developed which monitor various properties of the resonator as a function of the interaction between a tip glued to one prong of the fork and a surface. The spatial resolution of those microscopes are usually attributed to the size of the tip. We wish here to check under which conditions the vibration amplitude of the prongs of the tuning fork might become a limitation of the spatial resolution of scanning probe microscopes.

2 Finite Element Analysis modelling

While a bare, symmetrical quartz tuning fork is accurately modelled as coupled oscillators, finite element analysis [6, 7] is necessary when an additional disturbance such as a tip is asymmetrically attached to one of the prongs. We have first validated our simple quartz tuning fork model with experimental data based on dimensions measured under scanning electronic microscope and tabulated material constants [8, 9]. We have then extended this model by adding tips along one prong with various mechanical properties as used in experiments: either the tip is made of glass (optical fiber) assumed to have mechanical properties similar to that of quartz, of the tip is made of platinum.

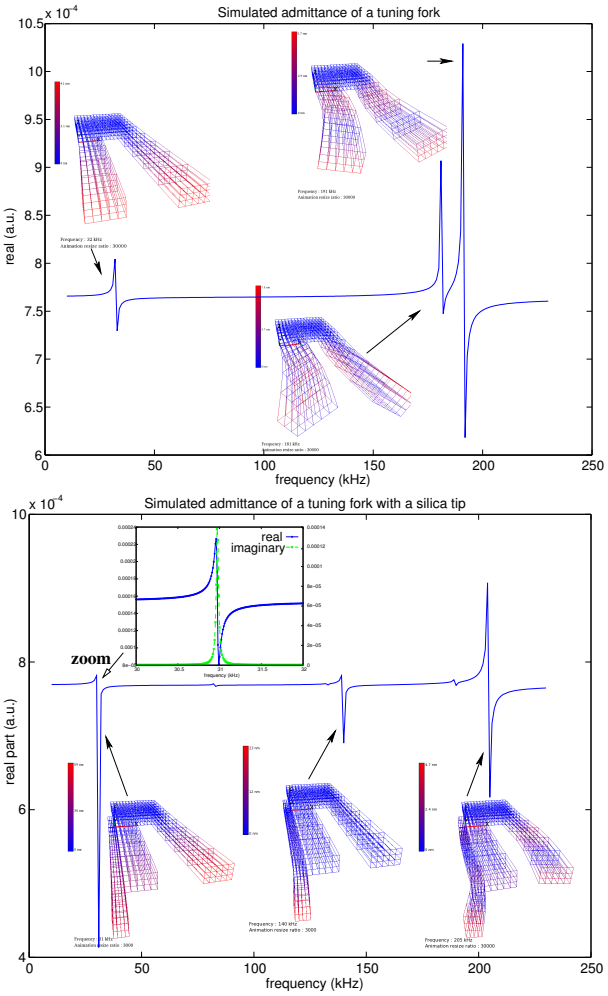


Figure 1: Top: impedance magnitude plot of the bare quartz tuning fork in the 10 to 230 kHz range. The quality factor was taken to be 1000, the driving voltage 0.5 V. The fundamental flexion mode around 33 kHz, the first torsional mode at 181 kHz and the flexion harmonic around 191 kHz are clearly visible. Insets: displacement of the tuning fork. Bottom: same plots for a silica fiber-loaded tuning fork. The asymmetry of the fundamental mode is due to the poor frequency resolution due to limited computation time. Notice the additional mode around 140 kHz which is a resonance of the protruding tip and strongly dependant of the bonding conditions between the fiber and the tuning fork. The 181 kHz torsional mode visible on the bare tuning fork plot has probably been strongly attenuated by the asymmetry due to the glued fiber.

The model is validated by computing the resonance frequencies of the system. An extended version of Modulef (INRIA, France) developed at our institute [10] has

been used for computing the electrical impedance and mechanical motion of the piezoelectric device under consideration for a range on frequencies (Fig. 1). We observe (Fig. 1) as expected the resonances around 32 and 191 kHz [11]. These resonances are associated with maximum displacement amplitudes at the end of the prongs of 110 and 20 nm respectively ($Q = 1000$, 0.5 V driving voltage). An analysis over such a wide range of frequency is incompatible in terms of computation time with the high frequency resolution required by the high quality factor experimentally observed ($Q \simeq 1000$) and included in the model as the division factor between real and imaginary parts of the material constants. Hence zooms were performed around the fundamental resonance frequency in order to provide an accurate result independent of the frequency step of the analysis; as seen in inset of Fig. 1.

The linear dependence of the displacement amplitude with the quality factor of the resonator and the driving amplitude have been verified by modelling the tuning fork with quality factors of 1000 and 10000 (leading to a 10 fold increase of the vibration amplitude) and a driving voltage of 0.1 and 0.5 V.

We observe that a quartz tuning fork with one prong loaded with an optical fiber along its length will display a vibration amplitude of the tip of 55 nm when excited at 0.5 V, assuming a quality factor $Q = 1000$ as experimentally observed. This maximum displacement amplitude is reduced to 17 nm if the optical fiber is replaced by a platinum wire of similar dimensions.

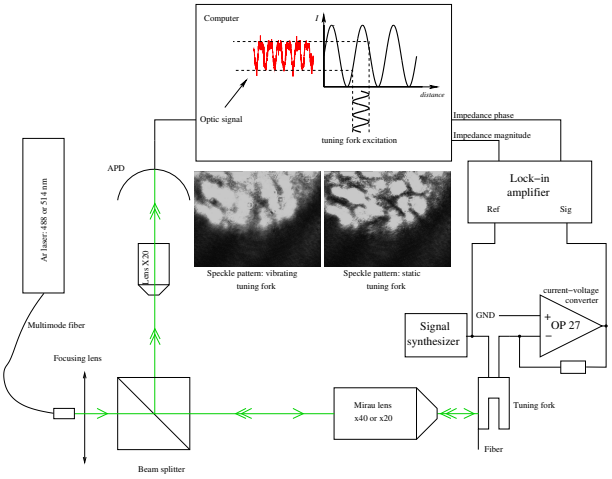


Figure 2: Experimental setup for measuring the displacement of one prong of the quartz tuning fork. Inset: speckle pattern as seen by the photodetector (APD) with a static and a vibrating tuning fork. The blur is due to the integration time of the CCD camera as the tuning fork was vibrating around 32 kHz.

3 Experimental results

3.1 Interferometric method

We have developed a speckle interferometer [12, 13, 14] (Fig. 2) for validating the amplitude informations extracted from finite element modeling. The main issue with the experimental measurement of the vibration amplitude of the quartz tuning fork are :

- the surface roughness of the electrodes which has been measured to be in the 3 to 10 μm range peak to peak. This very large roughness associated with the long coherence length of the Ar laser used implied the use of speckle interferometry

- the large observed amplitudes (under extreme working conditions) – above $\lambda/2$ where $\lambda \simeq 488$ or 514 nm – require an unusual processing step due to folding of the sine-shaped signal over the non-linear interference pattern.

We have varied two parameters: the excitation amplitude (the greater the electrical excitation amplitude, the better the signal to noise ration but the larger the displacement of the tip: Fig. 3) and the excitation frequency in order to be able to link the electrical properties of the tuning fork to mechanical displacements and hence predict the tip displacement without the need for the in-situ interferometric monitoring during actual SPM experiments (Fig. 4). In the latter case, one concludes that whatever the SPM tip-sample feedback mode (working at constant output current magnitude or phase) within the resonance range, the vibration amplitude of the prongs is important over the whole width of the resonance: although our conclusions focus on the maximum vibration amplitude, it can be extended to any feedback method of the tip-sample distance (including feedback of the tip-sample distance for keeping the electrical phase constant).

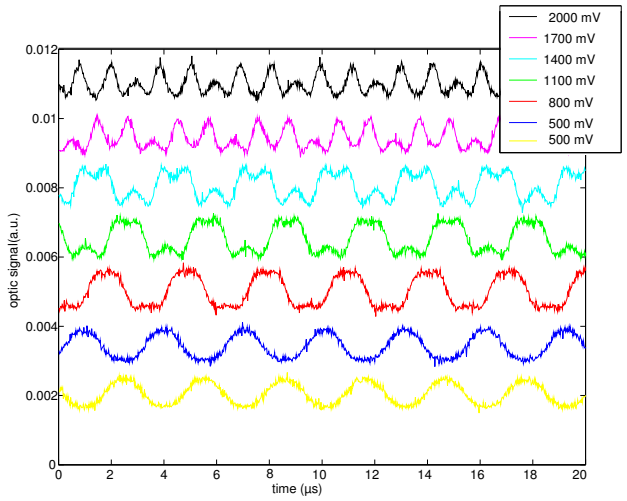


Figure 3: Experimental results of the vibration amplitude as a function of excitation voltage at maximum magnitude frequency

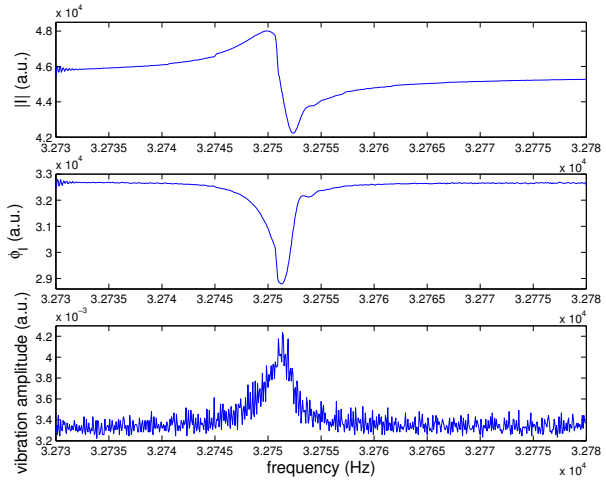


Figure 4: Experimental results of the vibration amplitude as a function of excitation frequency over a range of 50 Hz.

4 Data processing: linking model and experimental data

Data processing of the raw experimental data is performed in the following way :

- for each excitation amplitude a periodically sampled data set has been acquired (sampling rate: 5 Msamples/s, 16 bits/sample)
- each data set is Fourier transformed
- the power of the fundamental frequency components ($f_0 \simeq 32750$ Hz) and of the overtones (due to the non-linear intensity-distance relationship) are computed
- in parallel to processing the experimental data, we simulate the folding of the sine-shaped optical path difference $\delta \propto \omega t$ (with ω the pulsation of the acoustic signal and t the time) over the interference pattern: we obtain an intensity as seen by the photodetector shaped as $\sin(\frac{2\pi}{\lambda} A \sin(\omega t)+\varphi)$, with A the vibration amplitude of the tuning fork and φ a constant offset of the average position of the tuning fork with respect to the linear region of the interference pattern. Here again for each value of A we compute the Fourier transform and extract the Fourier components of the first overtones. We observe that the fundamental and third components are characteristic of the amplitude A while the second component is mostly determined by φ. The only free parameters of this model are the range of A and the phase offset φ.
- from a comparison of the ratio of the powers at the fundamental frequency and overtones (n$\times f_0$, n=[2, 3]) as seen from experimental and simulated data, we extract the quantitative vibration peak-to-peak amplitude in terms of probing laser wavelength: the range of A that best fits our experimental data provides an estimate of the vibration amplitude of the prong under investigation.
- the quantitative peak-to-peak amplitude is obtained from the conversion of the incoming laser wavelength to nanometers (Ar lasers generate two strong signals at either 488 or 514 nm – both have been used in various experiments)

Figs. 5 and 6 display the result of data processing of the experimental data for a bare tuning fork and an optical fiber loaded prong respectively, and a comparison with the modeled power spectrum of a sine wave of varying amplitude folded over the interference pattern. These experimental data display a motion of the cantilever inferior to the one expected from finite element modelling: the latter leads us to predict a peak to peak amplitude for a bare resonator of 174 nm displacement at 0.4 V for a Q=1000 resonator while we experimentally observe a displacement of about 46 nm. Such a result would be consistent with an exceptionally low Q=265 resonator in air which could be attributed to the aging of the electrodes exposed for several weeks to air. The optical fiber loaded tuning fork experimentally displays a vibration amplitude of 11 nm for an excitation voltage of 0.5 V while finite element analysis of a $Q = 1000$ resonator lead us to expect an amplitude of 110 nm for the same driving voltage amplitude. The experimental result would be consistent with a $Q \simeq 100$ while we experimentally observe $Q = 6700$ for this particular setup.

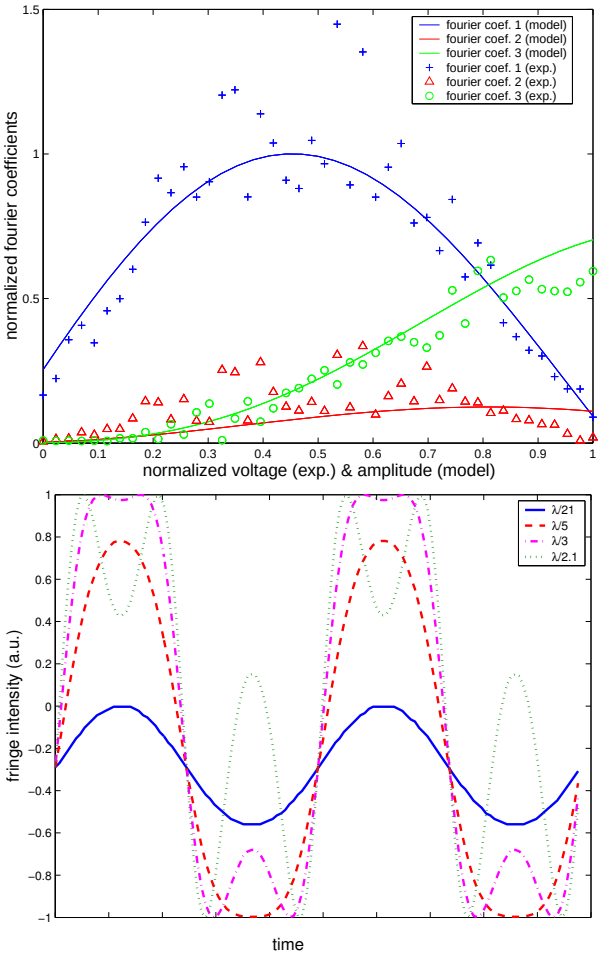


Figure 5: Top: fit of the experimental data to a model of sine wave folding over the interference pattern for extracting quantitative displacement data. The abscissa is graduated in voltage from 400 to 9000 mV (experimental data), which is also equal to (simulated data) a vibration amplitude of $\lambda/21 = 23$ nm to $\lambda/1.7 = 290$ nm ($\lambda = 488$ nm in this experiment). Bottom: result of the model of the folding of the incoming sine wave optical signal over the interference pattern as shown in Fig. 2. Compare with the experimental curves displayed Fig. 3

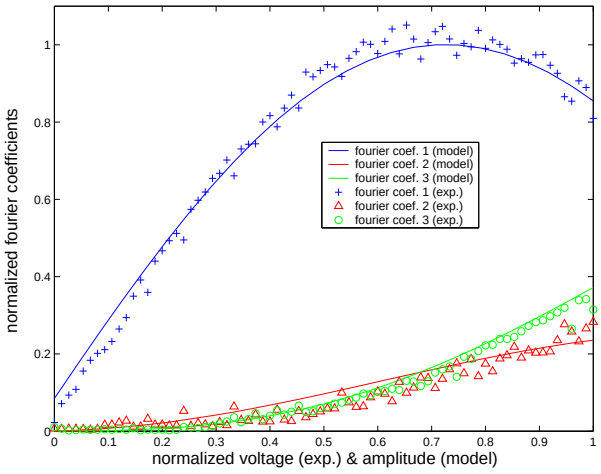


Figure 6: Comparison of the experimental data obtained using an interferometric method on a metallized fiber (the surface of the fiber is smooth enough to obtain here interference fringes rather than speckles) and a model as described in the text. The abscissa spans from 100 mV to 7600 mV amplitude (experiment) which is also equal to $\lambda/126$=4 nm to $\lambda/5$=97 nm (here $\lambda = 488$ nm).

4.1 Stroboscopic method

An alternative to interferometry is to use stroboscopic digital image sampling for measuring the displacement of a surface parallel to the place of the camera sensor [15]. In order to detect minute displacement as required by our application, sub-pixel resolution is obtained by applying an intercorrelation between images of different phase with respect to the oscillation of the tuning fork after interpolation of the images [16].

In this experiment the quartz tuning fork is probed by a square wave signal. Four images phase-synchronized with the oscillation of the tuning fork are taken for each period (Fig. 7).

The result of such a measurement, although sensitive to the roughness of the surface and the induced blurred area of the image due to limited field depth of the microcoscope lens, are as follow : for a peak to peak applied voltage of 0.5 V_{pp} (1.8 V_{pp}, 5 V_{pp} respectively), the displacement amplitude is observed to be 350 nm (850 nm, 3000 nm) as shown Fig. 8. These measurements were performed on a different tuning fork than the one used in the previous experiments, as would be compatible with a resonator with quality factor $Q = 4500$.

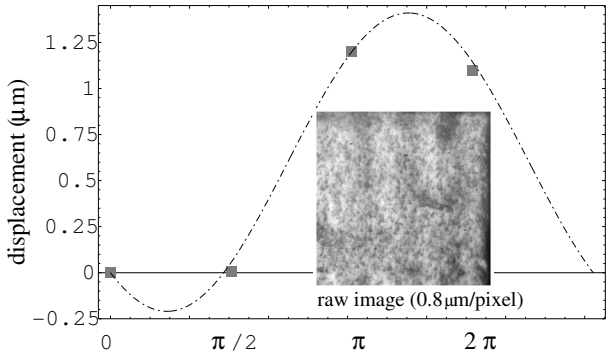


Figure 7: Example of a fit of the sine wave motion of the end of a cantilever powered by a 1.8 V_{pp} square wave. The four squares are the experimental data and the sine wave the fit from which a vibration amplitude of 750 nm is deduced at this pixel position. Inset: one of the four raw images acquired of the surface of the end of a prong: note the obvious roughness of the surface which seems to degrade the resolution of the method.

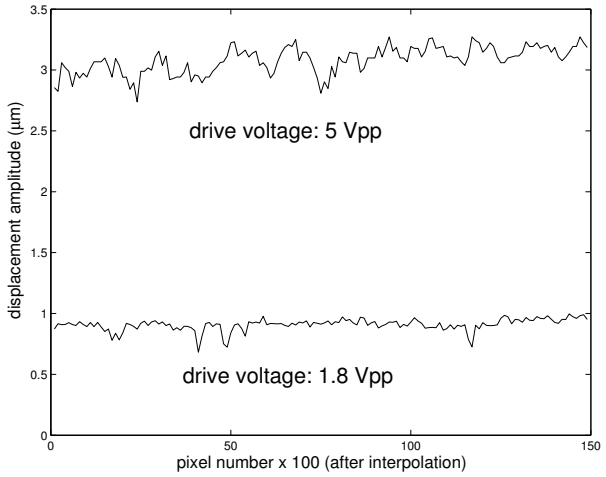


Figure 8: Peak to peak motion of the end of the prong of a tuning fork powered by 1.8 V_{pp} and 5.0 V_{pp} square voltages as computed for all pixels of the image following the method shown Fig. 7. These curves display a standard deviation of 50 nm and 110 nm respectively.

5 Conclusion

We have shown that under many circumstances the vibration amplitude of the quartz tuning fork prong to which the sensing tip of a shear-force microscope is attached might limit the spatial resolution of the scanning probe microscope to a few tens of nanometers, well above the spatial resolution obtained with conventional (contact or tapping) mode atomic force microscopes. These results have been obtained from measurements on bare quartz tuning forks and from preliminary results obtained from tuning forks with one prong loaded with an optical fiber. Experimental results on the bare tuning fork are in agreement with finite element modeling, so that we can confidently extend our finite element analysis of silicate and Pt fiber loaded tuning fork to the same conclusion.

Experimental confirmation of the data computed using finite element analysis software is still needed. A promising extension of this work is the quantitative analysis of the tip-sample interaction by means of electrical impedance recording as well as optical interferometric monitoring of the vibration amplitude. Such interactions can then be compared to classical models of external forces acting on the tip included in the finite element analysis software package.

Acknowledgments

We acknowledge the work of J.-J. Hunsinger (FEMTO-ST, Montbéliard, France) for the development of the electronics of the stroboscopic measurement method.

References

[1] K. Karrai and R.D. Grober, *Piezoelectric tip-sample distance control for near field optical microscopes*, Appl. Phys. Lett. vol 66, num 14, pp.1842-1844, April 1995

[2] Y.-H. Chuang, C.-J. Wang, J.-Y. Huang and C.-L. Pan, *Nonoptical tip-sample distance control for scanning near-field optical microscopy*, Appl. Phys. Lett. vol 69, num 22, pp.3312-3314, November 1996

[3] F.J. Giessibl, *Atomic resolution on Si(111)-(7× 7) by noncontact atomic force microscopy with a force sensor based on a quartz tuning fork*, Appl. Phys. Lett. vol 76, num 11, pp.1470-1472, March 2000

[4] Y. Seo, W. Jhe and C.S. Hwang, *Electrostatic force microscopy using a quartz tuning fork*, Appl. Phys. Lett. vol 80, num 23, pp.4324-4326, June 2002

[5] D. Courjon and C. Bainier, *Le champ proche optique : théorie et applications*, Springer (2001) [in French]

[6] R. Arinero and G. Lévêque, *Vibration of the cantilever in Force Modulation Microscopy analysis by a finite element model*, Rev. Sci. Instrum. vol 74, num 1, pp.104-111, January 2003

[7] H. Hantz, F. Sthal and R. Bourquin, *Model for a quartz-crystal tuning fork with coupling between the arms*, in Proceedings of 18th EFTF, 2004

[8] R. Bechmann, A.D. Ballato and T.J. Lukaszek, *Higher order temperature coefficients of the elastic stiffness and compliances of alpha-quartz*, in Proceedings of IRE, pp. 1812-1822, August 1962

[9] Dubois99 (EPFL)

[10] P.F. Edoa, S. Ballandras and M. Wilm, *A finite-element/boundary-element method for the modeling of piezotransducer radiation in fluids using a polynomial development of the Green's function*, J. Acoust. Am. vol 115, num 6, pp.2974-2953, June 2004

[11] V.A. Bykov, S.A. Saunin, I.V. Duskin and A.V. Mezin, *Second mode (190 kHz) of the vibration of the tuning fork application in shear-force measurements*, in Proceedings of the SPM-2002, 2002, pp.258-260

[12] T.R. Moore, *A simple design for an electronic speckle interferometer*, Am. J. Phys. vol 72, num 11, pp.1380-1384, 2004

[13] C. Shönenberger and S.F. Alvarado, *A differential interferometer for force microscopy*, Rev. Sci. Instrum. vol 70, num 10, October 1989

[14] P.G. Gucciardi, G. Bachelier, A. Mlayah and M. Allegrini, *Interferometric measurement of the tip oscillation amplitude in apertureless near-field optical microscopy*, Rev. Sci. Instrum. vol 76, web reference 036105 (2005)

[15] B. Serio, J.-J. Hunsinger and B. Cretin, *In-plane measurements of microelectromechanical systems vibrations with nanometer resolution using the correlation of synchronous images* , Rev. Sci. Instrum. vol 75, num 10, pp.3335-3341, October 2004

[16] H. Foroosh, J.B. Zerubia and M. Berthod, *Extension of phase correlation to subpixel registration*, IEEE transactions on image processing, vol 11, num 3, pp. 188-200, March 2002

DESIGN OF A NEW MINIATURE RESONANT TEMPERATURE SENSOR IN NLC-CUT OF QUARTZ

♣B. Dulmet, ♦L. Spassov, ♣R. Bourquin, ♣A. Ivan, ♦Ts. Angelov, ♦J. Lazarov

♣FEMTO-ST dpt. LCEP, ENSMM, 26 Chemin de l'Epitaphe 25000 Besançon, France
♦ISSP-BAS, AELab, 72 Tzarigradsko Chaussee, 7148 Sofia, Bulgaria

ABSTRACT

We present the development of new miniature resonant sensors which consist of quartz strip-resonators in NLC-cut [1]. We show how a combination of semi-analytical modeling and FEA was achieved to optimize the miniaturization of sensors while maintaining a performance level quite similar to the one obtained with mass-loaded trapped-energy resonators operating on fundamental mode at 29.3 MHz. A significant effort was devoted to control activity-dips arising from width-resonances of the structures. Dimensions of strips are 0.082 x 2 x 6 *mm*, their Q-factor routinely surpasses 50 000. Temperature sensitivity is near 1100 Hz/K making easy the use of a seamless frequency output on large temperature ranges between 4.5 and 400 K.

INTRODUCTION

Here–presented cooperative work was achieved in the framework of QxSens EU contract from program GROWTH, number GRD1-2001-4816. The main goal of this project is to develop an industry-oriented multi-channel system for monitoring and control which is entirely based on a variety of resonant piezoelectric sensors. Among them, the strip-resonator design was finally retained for temperature sensing applications, for two reasons :

- it fits in much smaller enclosures than classical disk resonators operating in trapped-energy thickness shear [1-2],
- it operates at a much higher frequency than flexural bars, which offer a less stable resonant frequency with respect to most of the perturbing effects.

Then, the smaller volume is expected to significantly improve the response time in comparison with classical trapped energy resonators. Nevertheless, using strip resonators as sensors in large temperature ranges has a counterpart : it requires to control the stability of spectrum with respect to temperature, since interchanges of modes inside the spectrum (activity-dips) perturb the sensor function. In this view, we performed extensive simulations by Finite Element Analysis to analyze the effect of slow temperature changes on the resonant modes spectrum and determine optimal dimensions of the sensors for a given temperature range. Elaborating a semi-analytical model greatly facilitated the optimization of the FEA model to guarantee the convergence of simulations of the main mode and essential spurious modes. This dual approach is required by the large computer resources required by FEA. The modeling effort was accompanied by X-rays topography analysis performed at various temperatures in order to track possible activity-dips when indicated by frequency-temperature measurement. The finalization of design was made throughout a cooperation between LCEP and AELab, whereas pilot series of TSQR (Thermo Sensitive Quartz Sensors) were actually realized and characterized by AELab, which handled a few samples to LCEP for X-rays topography and electrical measurements between liquid nitrogen and ambient or slightly above ambient temperatures.

DESIGN AND MODELLING ISSUES

The NLC-cut [1] is a singly-rotated cut defined as shown on Fig. 1. The actual value of θ angle is -31°30'. It allows a minimal defect of linearity between cryogenic and ambient temperature.

Classical resonators are shaped as disks with circular electrodes designed to trap energy by mass-loading. When operating on fundamental mode at 29.3 MHz, they present a temperature sensitivity close to 1100 Hz/K. Optimal features of series of such TSQRs previously presented in [2-3] are the following :

- disk diameter $\varnothing_q = 5\ mm$
- disk thickness $t_q \approx 82\ \mu m$
- electrode diameter $\varnothing_e = 2\ mm$
- electrode thickness $80\ nm < t_e < 240\ nm$

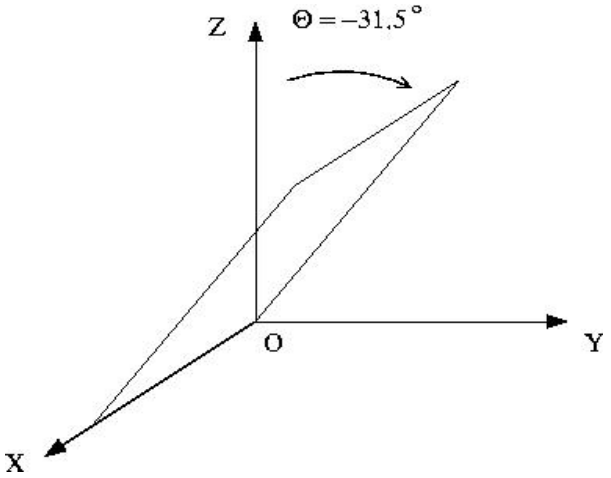


Fig. 1 *Definition of NLC-cut.*

Electrodes were made from silver which provides easier thickness control than the denser gold in the case of thin plates used for high frequency TSQRs. Typical electrical parameters obtained with such features are the following :

$$C1 = 8.3\ fF \quad Q = 75\ 000 \quad R1 = 9\ \Omega$$

which we shall consider as typical reference characteristics to rate the performances of strip resonator design.

From the theoretical point of view, an ideal strip resonator in NLC-cut is just a rectangular plate with deposited electrodes covering the whole width of the strip and only the central region in the direction of length, as shown on Fig. 2. Due to that, mass loading effect induces energy trapping along the length but not along along the width. If the lengths of the plate L and of the electroded region L_e tend to be large enough *w.r.t.* its width w, a useful approximate analysis consists in solving the simpler two-dimensional problem of standing waves in the rectangular cross-section $(x_1 o x_2)$, arising from the condition $\partial/\partial x3 \equiv 0$. Taking into account energy trapping along x_3 turns the problem into a three-dimensional one, treated only by FEA in the present work. Nevertheless, the two-dimensional approach suffices to show the specific features of strip resonator design and in particular the importance of the choice of the aspect ratio w/t of the cross-section of strip as well as the tolerances imposed upon it for reproducibility of results in a series of strip-resonators.

The principle of a two-dimensional semi-analytical modeling of the transverse behavior of narrow plates was presented with some details in Sec. 2 of Ref [2] aimed to the study of intermediate designs between the classical trapped-energy circular design and the actual strip resonator. Although it is a very classical analysis, we take the liberty to recall it below to enlighten some key points of our designing approach.

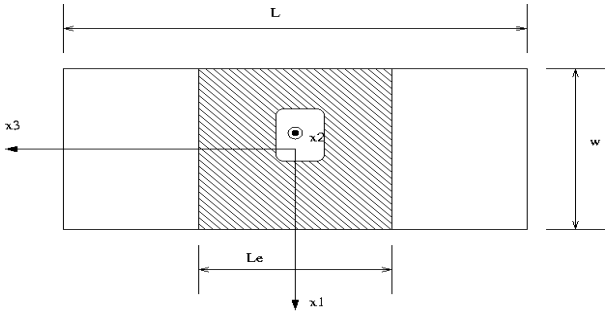


Fig. 2 *An ideal strip resonator*

The initial step consists in the analysis of guided waves propagating along x_1, *i.e.* the direction of the strip width, and having a piezoelectrically-excitable component u_1 of mechanical displacement which is simultaneously antisymmetric along the thickness (x_2 axis) of NLC-cut and symmetric w.r.t. the middle of the strip. In singly-rotated $Y+\theta$ cuts, the most general form of solution of elasticity having this symmetry and obeying free boundary conditions on major surfaces of the plate at $x_2 = \pm$ h is the following :

$$\begin{aligned} u_1 &= \cos(\xi\, x_1) \sum_{n=1}^{3} C_n\, \beta_1^n \sin(\eta_n\, x_2) \\ u_2 &= \sin(\xi\, x_1) \sum_{n=1}^{3} C_n\, \beta_2^n \cos(\eta_n\, x_2) \\ u_3 &= \sin(\xi\, x_1) \sum_{n=1}^{3} C_n\, \beta_3^n \cos(\eta_n\, x_2) \end{aligned} \qquad (1)$$

This leads to the dispersion curves shown on Fig. 3 where the ordinate and abscissa are the frequency and lateral wavenumber respectively normalized by the values of the vertical wavenumber η and frequency of thickness-shear B mode fundamental when $\xi=0$ *i.e.* for pure thickness propagation. Accordingly, the dimensionless frequency Ω and wavenumber γ are defined by

$$\left.\begin{aligned} \Omega &= \frac{\omega}{\omega_B^F} \\ \gamma &= \frac{\xi}{\eta_B^F} \end{aligned}\right\} \qquad (2)$$

where ω_B^F denotes the fundamental resonant frequency of pure shear B-mode and η_B^F is the vertical wavenumber of that mode, which is equal to $\pi/2h$ if $t = 2h$ denotes the thickness of plate (piezoelectricity is neglected here). The left half of Fig. 3 is for purely imaginary γ, while the right region is for purely real γ. Near the Ω axis, we have also represented the second order asymptotic expansion of dispersion curves corresponding to the calculation of Tiersten's dispersion constant *Mn* for *B1*, *A2, C2* quasi-thickness modes.

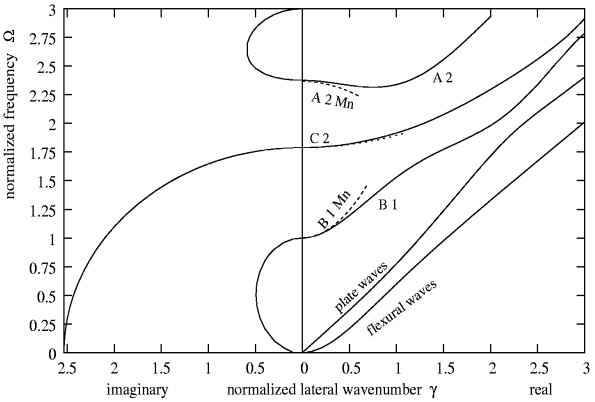


Fig. 3 *Dispersion curves in NLC-cut for modes described by Eq (1)*

One can observe that, in NLC-cut, the actual curves remains closer to the parabolic asymptote for *C2* than for *B1* and *A2*.

The first stage of the two-dimensional model is rigorous and well established, so that the combinations (1) actually describe branches of the dispersion curves as self consistent solutions of elasticity equations and boundary conditions on the major surfaces of plate located at $x_2=\pm h$. But, in the absence of energy trapping along x_1, a single branch of dispersion curve cannot be a self-sufficient solution of the vibration problem for a strip plate with rectangular cross section (x_1ox_2). This comes from the fact that the form of the solution (1) does not allow to identically satisfy the three boundary conditions $T_1=T_5=T_6=0$ for any value of x_2 on the free edges located at $x_1=\pm w/2$. Then the behavior of the vibration is actually more complicate than suggested by (1). Looking for more elaborate initial functions to describe the problem is widely considered as an inefficient approach due to its very strong mathematical difficulty, if not to say its intractability. Instead, considering that the aspect ratio *w/t* of strip is rather large in actual resonant devices, a pragmatic approach consists in building *approximate* solutions by combining branches of dispersion curves in order to obey the edge boundary conditions at $x_1 = \pm w/2$ as precisely as possible.

Any such combination will predict a modulation of the amplitude distribution along the width of plate even if the thickness shear branch *B1* of Fig. 3 is dominant. This essential feature is exemplified on Fig. 4. If the main *B1* branch of such an approximate solution roughly exhibits one half wavelength along *w* (Fig.4-left), then the other branches having a real transverse wavenumber will have a much higher value of γ, as shown on Fig. 4-right, which is an enlarged partial view of the dispersion curves of Fig. 3.

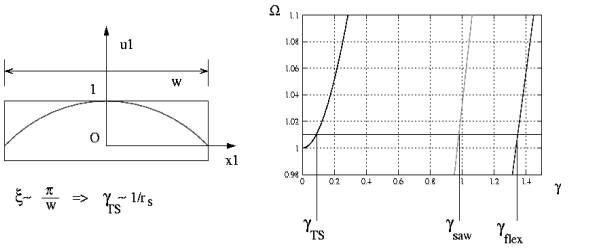


Fig. 4 *A simple way to roughly estimate the transverse behavior of main modes in strip resonators.*

This simple reasoning suffices to obtain the order of magnitude of the wavelength of the spatial modulation of the amplitude distribution which can be expected in a strip resonator with a given aspect ratio. Nevertheless, the amplitude of this modulation depends on the precise balance between the various branches involved in the global combination. There are several ways to actually build the approximation. In Ref. [5], Tiersten clearly showed that a very efficient method for this class of problems consists in using an appropriate form of Hamilton's principle. Here, we can use the same variational formulation which is actually at the kernel of the finite element method :

$$\int_{t_0}^{t} \left\{ \int_V \left[T_{ij}\delta u_{j,i} + \rho \ddot{u}_j \delta u_j \right] dV - \int_S f_i \delta u_i dS \right\} dt = 0 \qquad (3)$$

and is the so-called weak form of Hamilton's principle :

$$\delta \int_{t_0}^{t_1} dt \int_V [\mathcal{T} - \mathcal{U}]\, dV + \delta \int_{t_0}^{t_1} dt \int_S \mathcal{W}\, dS = 0 \qquad (4)$$

In these equations δu_i has the meaning of an arbitrary variation imposed to the function $u_i(x_1,x_2,x_3,t)$ while $\mathcal{T}$, $\mathcal{U}$ and $\mathcal{W}$ respectively denote the kinetic energy, elastic potential energy and the work received by the structure from external actions. Using Green's theorem, it is very easy to demonstrate that Eq. (3) is equivalent to the so-called strong form of Hamilton's principle

$$\int_{t_0}^{t_1} dt \left[\int_V \left(T_{ij,j} - \rho \ddot{u}_i \right) \delta u_i \, dV + \int_S \left(f_i - n_j T_{ij} \right) \delta u_i \, dS \right] = 0 \qquad (5)$$

where f_i is an applied external surface traction and n_i denoes the outer normal unit. Now if one takes as trial function a combination of solutions of the form (1), *i.e.* branches of dispersion curves, a choice which actually fixes their parameters in terms of ω, the variational equation reduces to

$$\int_{-h}^{+h} \left(T_1\delta u_1 + T_6\delta u_2 + T_5\delta u_3 \right)_{x_1=\pm\frac{w}{2}} dx_2 = 0 \qquad (6)$$

because the trial functions make equal to zero the volume integral as well as the surface integral on top and bottom surfaces of the plates. Since the form of the trial functions is fixed, the variation operator δ only

affects the weights of each branch in the combination. Thus, since Eq. (6) must hold for any arbitrary variation δu_i, one readily obtains an homogeneous linear system the dimension of which is the number of branches of the dispersion curves retained in the combination. To admit non trivial solutions, this system must have a null determinant. This condition determines Ω and all other parameters in terms of the aspect ratio *w/t* of the plate. Additional details were given in Sec. 3 of [2]. We only retained the three branches with γ real, which appeared on the our reasoning about Fig.4. This procedure can indeed be reiterated with various value o aspect f the aspect ratio, thereby defining a new kind of dispersion curves, now Ω=f(*w/t*) pertaining to strip plates of rectangular cross-section with free boundaries.

Here-presented model is not state-of-the-art but it is intended only for general understanding and as a first step to speed up the correct settings of 3-D finite element simulations. Various ways to determine the dispersion curves can be found in earlier works by Mindlin *et al.* [6-8]. In particular, we cross-checked the results obtained from here-presented semi-analytical model with the ones provided by a computer program strictly implementing the calculation of dispersion curves after [8]. A local optimum is found when the slope of the dispersion curve Ω=f(*w/t*) is minimal because this situation corresponds to a maximal contribution of thickness shear branch, *B1* in here-studied case (*cf.* Fig. 12 of Ref. [2]). It should be remarked that the amplitude of modulation in the vicinity of successive local optima significantly decreases when the aspect ratios increases, leading to larger value of the motional capacitance C_1. As a whole, this fact pleads for designing plates with larger aspect ratios. From another hand, the precision of simple models may significantly decrease for plates with larger aspect ratios, while 3-D effects involved by the behavior along the second in-plane axis x_3 introduce additional corrections. At this point, FEA analysis is still the best choice to refine the determination of optimal design parameters, but the wider the strip the larger the required computer resources.

Due to the need of miniaturization and good electrical parameters, there is no advantage in designing resonant temperature sensors as wide as classical mass-loaded trapped-energy resonators. After balancing all these considerations with technological feasibility, it was decided to keep the aspect ratios of actual strip sensors around *w/t=25*. Fig. 5 shows the semi-analytically predicted distribution of main mechanical displacement u_1 obtained in that case, while Fig. 6 shows the FEA-computed u_1 amplitude distribution in one half of the strip w.r.t. x_1 and for *w/t=24.8*.

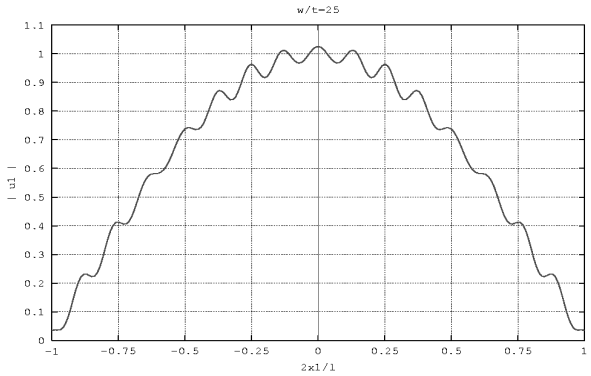


Fig. 5 *Theoretical distribution of u_1 along the width of a strip with w/t=25*

FEA-predicted oscillatory behavior of the amplitude along the width is in good qualitative agreement with the one predicted by semi-analytical modeling. The biggest part of simulation activity consisted in establishing dispersion curves **Ω**=f(w/t) by FEA once the region of actual realizations was decided. Due to their importance, we reproduce on Fig. 7 the FEA-obtained dispersion curves though they were previously presented in [2]. All modes represented on that Fig. have the same symmetry w.r.t. local axes of strip. To obtain the curves, it was found more efficient to keep *w* constant while *t* is varying. The main mode corresponds to the curve labeled (B,1,0,0). A nearly optimal amplitude distribution such as shown on Fig. 6 is only observed near the inflexion point of the curve, *i.e.* in the vicinity of *w/t=25*.

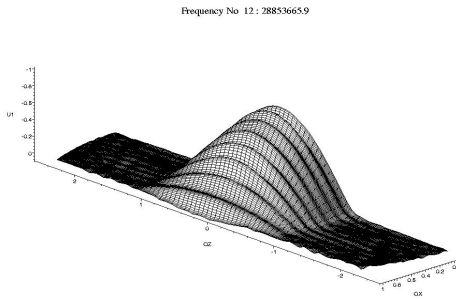


Fig. 6 *A convergent FEA model of strip resonator*

Since the transverse behavior of strip resonators along x_3 is essentially dominated by the energy trapping from electrodes, regular anharmonic modes such as *(B,1,0,1)* and *(B,1,0,2)* with an increasing number of nodal lines fixed along that direction are predicted and this was experimentally confirmed by X-ray topography. It was also confirmed that the transverse behavior of resonant modes along x_1 is not governed by electrodes mass loading, but by above–described coupling mechanism between branches of dispersion curves instead.

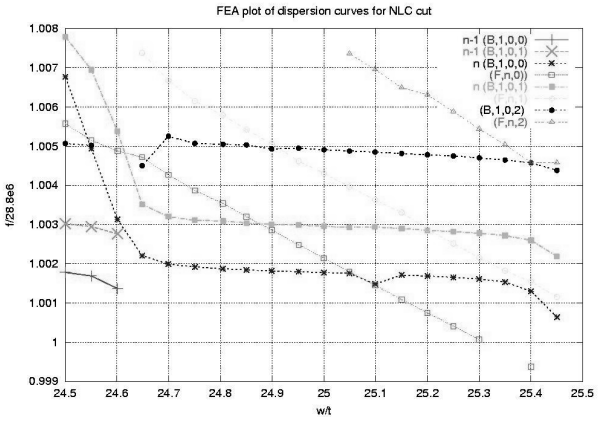


Fig. 7 *FEA-computed dispersion curves* $\Omega=f(w/t)$ *near* $w/t=25$

Due to the reciprocal nature of this mechanism, the solutions with dominant thickness shear are not the only possible ones, so that other modes with dominant quasi-flexure or plate waves components also exist. Modes such as the ones which we have labeled *(F,n,0)* on Fig. 7 belong to this category. Here, we use F for "flexure" in a rather generic fashion since it makes no real difference between the respective contributions of actual flexure and plate waves branches to the combined resonant modes. Although such modes are still governed by the aspect ratio *w/t*, they also exhibit anharmonic resonances along x_3 which correspond for instance to the branches labeled *(F,n,1)* and *(F,n,2)* on Fig. 7. Simulations of Fig. 8 top and bottom respectively present the u1 amplitude distributions of main mode *(B,1,0,0)* and of *(F,n,0)* at ambient temperature and for *w/t=25.*

The modes patterns strongly depend on the position on the dispersion curves, but they can also be somewhat affected by temperature variation. As a matter of fact, FEA simulations showed that modes such as *(B,1,0,0)* and of *(F,n,0)* have different temperature behavior since they are contributed by first-stage branches with different weights and different temperature sensitivities. For instance it was found that *(F,n,0)* had a negative temperature slope whereas *(B,1,0)* has a positive one.

Dispersion curves never exactly cross since the couplings systematically tend to separate degenerated resonant modes. Thus strong activity dips are expected at temperature above ambient when *w/t=24.8*, and at temperatures lower than ambient if *w/t=25.2* for instance, and so on for all *(F,n,x)* branches of dispersion curves with successive values of *n*. It means that strip resonators can be operated as temperature sensors with smooth transfer functions only in dedicated temperature ranges. The tolerances of actual realizations and the amount of required computer resources make an accurate prediction of these ranges rather difficult so that we had to leave it to experimental study.

INVESTIGATION OF ACTIVITY DIPS

Activity dips were searched by X-ray topography with help of a cryostat specially designed with a window transparent to X-rays, previously described in [10]. The cryostat could be operated between liquid nitrogen and slightly above ambient temperature.

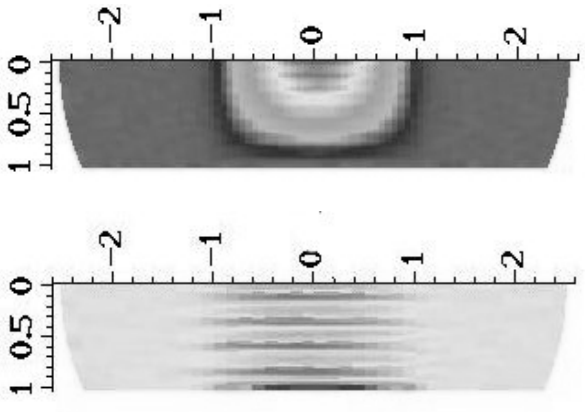


Fig. 8 *Main (top) and spurious (bottom) modes*

For obvious practical reasons, this research had to be restricted to strips sensors presenting clear and repeatable jumps on their transfer function. Due to fabrication tolerances, the field of investigation was actually rather large and it took time before we could find actual jumps in regions allowing a sufficient temperature stabilization of the whole system to make possible the acquisition of neat X-rays topographs. Figs 9-10 show the most convincing results obtained in this field. In the coupling region, the images result from a superposition of mode patterns which have different oscillatory behavior along the width of plate, a fact which is in agreement with above analysis of FEA-computed dispersion curves.

PERFORMANCES OF ACTUAL SENSORS

A pilot series of around 50 strip temperature sensors was made from circular disks with diameter *6 mm*, thereafter lapped to obtain the actual strip shape. Results obtained on the series are summarized on Table 1. Several tens of sensors were investigated and actually did not exhibit any disturbance of the temperature response while keeping good electrical parameters along the specified temperature range. Then only three sensors exhibited repeatable defects in their frequency-temperature response. Consequently, we cannot precisely determine at this moment the actual limits of smooth operation range concluded from above theoretical analysis. A smooth sensor function of a typical strip sensor is shown on Fig. 11.

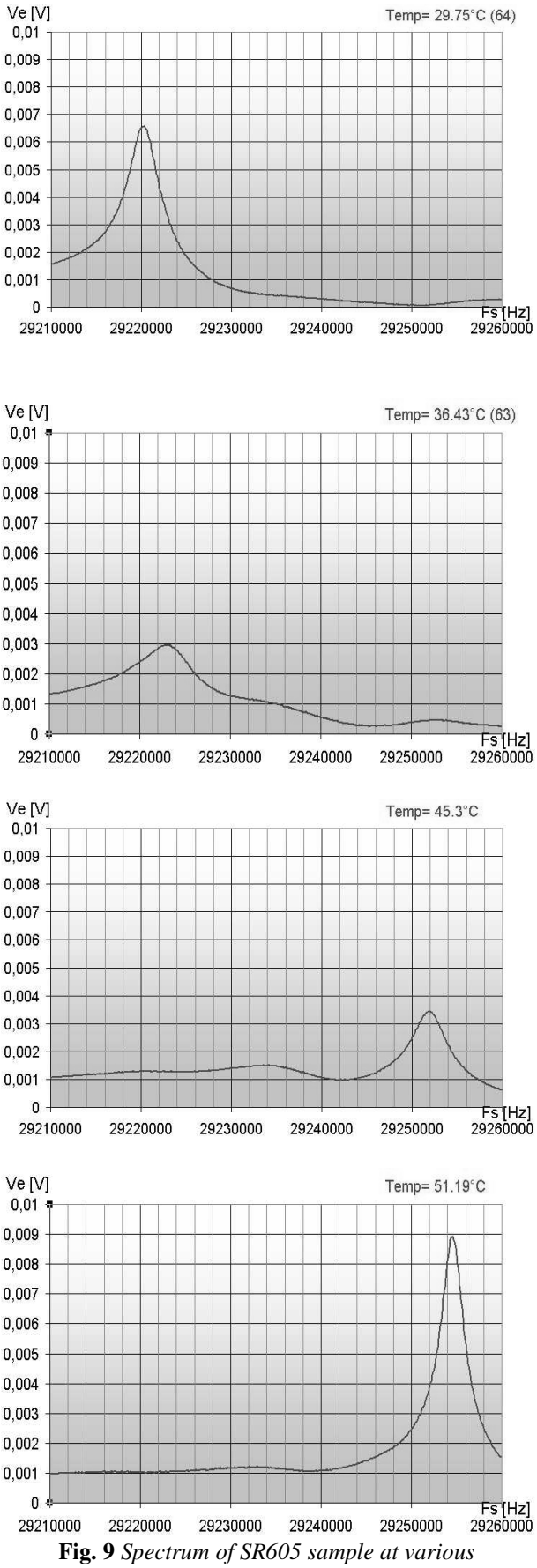


Fig. 9 *Spectrum of SR605 sample at various temperatures in the region of an isolated activity-dip.*

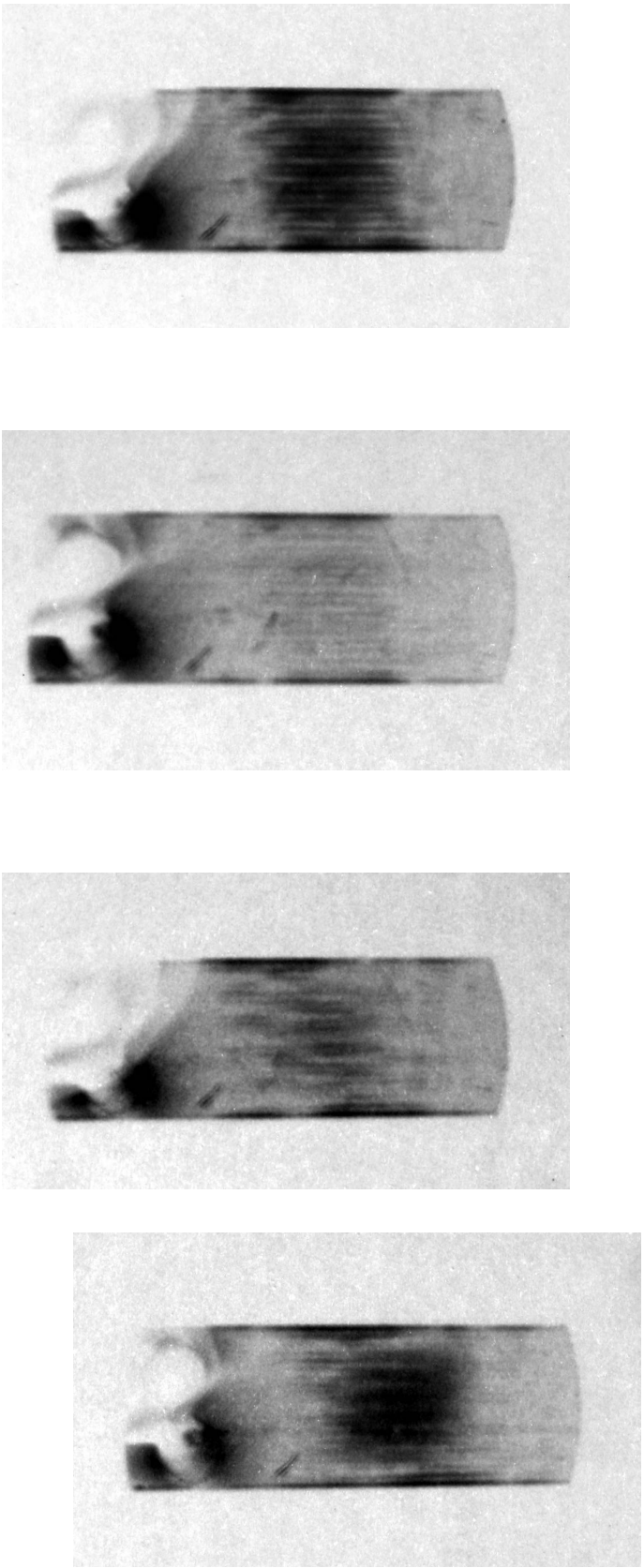


Fig. 10 *X-ray topographs corresponding side-to-side to the spectra presented on Fig. 9.*

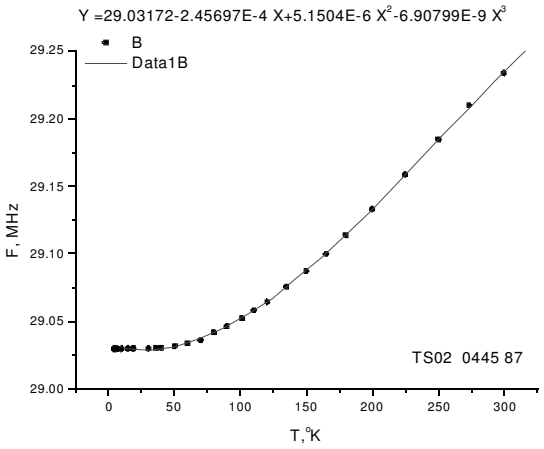


Fig. 11 *A typical frequency-temperature characteristic of strip sensor.*

		typical	***range***
Series frequency	f_S	29 080 kHz	29 045—29 180
Motional resistance	R_1	10 Ω	4.2—20.0
Motional inductance	L_1	3.50 mH	3.10—3.92
Motional capacitance	C1	8.24 fF	7.51—8.86
Static capacitance	C_0	2.32 pF	2.20—2.42
Quality factor	Q	50 000	40 000—80 000

Table 1 *Electrical parameters in range 4.2K—300K*

Fig. 12 shows the results of a simple experiment aimed to determine the dynamic response of strip sensors. It consisted in quickly moving the sensors back and forth between a bath of boiling water and a container kept at the temperature of triple point. Although this measurements may seem rough and have to be followed by more detailed characterizations, they show that the stabilization time is quite short, in the order of 10s.

CONCLUSION

Strip resonators expand temperature-sensing capabilities of piezoelectric resonators because they offer

- the really small dimensions of flexural bars,
- the higher Q and resonant frequency of thickness-shear plates,
- a shorter dynamic response time than thickness-shear plates.

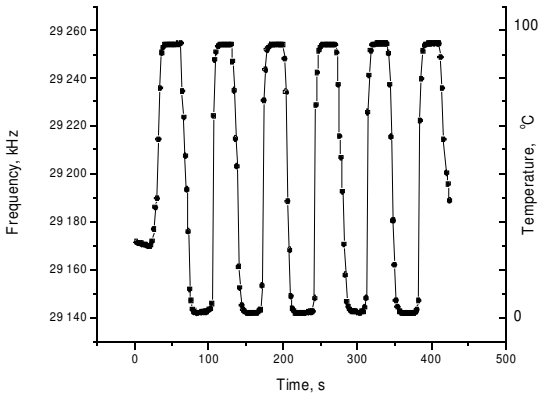


Fig. 12 *Dynamic response to temperature changes between 0 and 100°C*

The development of such miniature sensors requires an effective cooperation between the laboratory and the production unit. Although activity dips arising from spurious transverse resonances along the width seem inherent to their design, they can be controlled by a precise optimization and tolerance-keeping of the width-to-thickness aspect ratio of strip-plates scheduled for a given temperature range. In particular, smooth frequency temperature responses have been obtained between 4.5 K and 300 K on most of the realized samples of a pilot series with an average Q factor of 50 000 slightly above 29 MHz.

REFERENCES

[1] L. Spassov, E. Yossiffov, V. Georgiev, L. Vergov, *A rotated Y-cut quartz resonator with a linear temperature-frequency characteristic*, Sensors & Actuators, A58, pp.185-189(1997).

[2] B. Dulmet, L. Spassov, Ts. Angelov, *Finite element analysis for a new design of NLC-cut quartz crystal resonant temperature sensor*, Proc. EFTF 2004.

[3] H. F. Tiersten, *Analysis of Trapped Energy Resonators Operating in Overtones of Coupled Thickness Shear and Thickness Twist*, J. Acoust. Soc. Am., 59, P. 879 sq., (1976).

[4] D. Stevens, H. F. Tiersten, *An Analysis of Doubly Rotated Quartz Resonators utilizing Essentially Thickness Modes with Transverse Variations*, J. Ac. Soc. Am., 79, n° 9, pp. 1811-1826, (1986).

[5] H. F. Tiersten, *Linear Piezoelectric Plates Vibrations*, Plenum Press Ed., New York, (1969).

[6] R.D. Mindlin, *Thickness-Shear and Flexural Vibrations of Crystal Plates*, J. A.o P., pp. 316-323, 1951.

[7] R.D. Mindlin, *High Frequency Vibrations of Crystal Plates*, Quarterly of Applied Mathematics, April 1961, n°1, pp. 51-61.

[8] R.D. Mindlin, D.C Gazis, *Strong resonances of rectangular AT-cut quartz plates*, Proc. 4-th U.S. Nat. Congress of Appl. Mech., pp. 305-310, 1962.

[9] Plate waves extend the concept of surface waves when both top and bottom surfaces of a thin plate must be simultaneously taken into account.

[10] L. Spassov, R. Velcheva, L. Vergov, R. Bourquin, A. Ivan, B. Dulmet, *Investigation of electrical parameters and mode patterns of NLC-cut quartz resonators at low temperature,* Proc. EFTF 2004.

Links page

Next EFTF's :

2006 : Braunschweig, Germany (http://eftf2006.ptb.de)

2007 : Geneva, Switzerland

EFTF's related links :

Permanent website for EFTF :
http://www.eftf.org

Société Française des Microtechniques et de Chronométrie (S.F.M.C.) :
http://sfmc.obs-besancon.fr

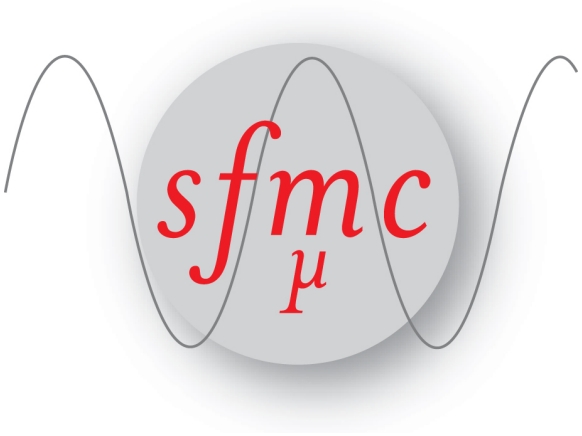
Fondation Suisse pour la Recherche en Microtechniques (F.S.R.M.) :
http://www.fsrm.ch

E-mail addresses :

EFTF 2006 : eftf2006@ptb.de

S.F.M.C. : sfmc@obs-besancon.fr

F.S.R.M. : eftf@fsrm.ch



Opening Session



Monsieur Jean-Louis FOUSSERET, Maire de Besançon and Président de la Communauté d'Agglomération du Grand Besançon.
Monsieur le Professeur Raymond J. BESSON

19th European Frequency and Time Forum, Besançon, France



From left to right :
Monsieur Vincent FUSTER, du Conseil Général du Doubs
Madame Paulette GUINCHARD-KUNSTLER, Députée du Doubs
Monsieur Jean-François ROBERT, Président du Conseil Economique et Social (CESR)
Madame Françoise BRANGET, Conseillère municipale de la Ville de Besançon
Madame Antoinette GILLET, Vice-Président du Conseil régional de Franche-Comté
Monsieur le Professeur Daniel RONDOT, Vice-Président de l'Université de Franche-Comté
Monsieur le Professeur Roger BOURQUIN, Directeur du Département LCEP de l'Institut FEMTO-ST et Technical Chairman de l'EFTF 2005

19th European Frequency and Time Forum, Besançon, France



From left to right :
Monsieur le Professeur Claude AUDOIN et Monsieur le Professeur Raymond J. BESSON awarding the European Award to M. Masao TAKAMOTO for M. Hidetoshi KATORI, for a brilliant break-through in the field of optical frequency standards.
Madame Antoinette GILLET, Vice-Président du Conseil régional de Franche-Comté

19th European Frequency and Time Forum, Besançon, France



From left to right :
Acknowledgements of Dr Peter W. KREMPL for his European Award on his pioneering research on the physical and electro-acoustical properties of Gallium Phosphate and the development of industrial growth processes for large crystals including the epitaxial growth of GaPO4 on quartz, leading to state of the art resonators and sensors.
Madame Antoinette GILLET, Vice-Président du Conseil régional de Franche-Comté, Monsieur le Docteur Bernd NEUBIG and Monsieur le Professeur Raymond J. BESSON.

Prix Jeune Chercheur / Young Scientist Award



Monsieur le Professeur Raymond J. BESSON, Madame Antoinette GILLET, Vice-Président du Conseil régional de Franche-Comté awarding the Young Scientist Award to Monsieur John KITCHING for the challenging development of the first highly miniaturized atomic frequency standard for chip implementation and large scale production.

City Hall of Besançon



Welcoming words by Monsieur Sébastien MAIRE, Adjoint au Maire de la Ville de Besançon
Monsieur le Professeur Raymond J. BESSON



Congressmen listening…

19th European Frequency and Time Forum, Besançon, France



Monsieur François VERNOTTE, Directeur de l'Observatoire de Besançon with some of the Russian congressmen and a few delegates from the CNES (Monsieur Eric LUVISUTTO et Monsieur Michel BRUNET).

Closing Session



Monsieur le Professeur Roger BOURQUIN and Monsieur le Professeur Raymond J. BESSON

19th European Frequency and Time Forum, Besançon, France



Monsieur Pierre THOMANN, Technical Chairman EFTF 2007

Awards Ceremony

Prix Européen / European Award



From left to right :
Madame Antoinette GILLET, Vice-Président du Conseil régional de Franche-Comté
and Monsieur le Professeur Raymond J. BESSON awarding the European Award to Monsieur Valentin N. KULESHOV for Monsieur George MANSFELD on his fundamental contributions in understanding physical and quantum acoustics, acoustoelectronics in piezoelectric semiconductors and acoustomagnetic effects in ferrites. For applications to frequency control and high Q microwave resonators using ultra low loss materials.

19th European Frequency and Time Forum, Besançon, France



Monsieur le Docteur Fritz RIEHLE, Physikalisch-Technische Bundesanstalt (PTB), Technical Chairman EFTF 2006, Monsieur le Professeur Raymond J. BESSON and Monsieur le Professeur Roger BOURQUIN

ACHKAR, J.

Time and frequency comparisons between four European timing institutes and NIST using multiple techniques

Calibration of four European TWSTFT earth stations with a portable station through Intelsat 903

Typical combined uncertainty evaluation on the Ku band TWSTFT link

AIDA, M.

TWSTFT plans in the Pacific-Rim region and Europe link

Calibration system of internal delay Difference for TWSTFT station

BALLANDRAS, S.

A perturbation method for accurate predictions of mass loading and thermal effects on acoustic devices computed using finite element analysis

New results on high frequency surface acoustic waves excited on an oriented LiNbO3, single crystal layer transferred onto Silicon

Longitudinally-coupled bulk acoustic wave resonators filter for 5 GHz applications

Theoretical analysis of damping effects of acoustic waves at solid/fluid interfaces using a mixed periodic FEA/BEM approach

Quartz tuning fork vibration amplitude as a limitation of spatial resolution of shear force microscopes

BAUCH, A.

Time and frequency comparisons between four European timing institutes and NIST using multiple techniques

Galileo system test bed V1 : results on the experimental Galileo System Time

Galileo Time Service Provider

Design and first results on the New PTB caesium fountain CSF2

Time scale prediction using information from external sources

BAZIN, N.

Simple model for the mode splitting effect in whispering gallery mode resonator

Ti3+ doped microwave temperature compen-sated sapphire resonator oscillator at 34K

Observation of a bistability effect in a cryogenic microwave whispering gallery mode resonator

Temperature regulated sapphire resonators for low phase noise oscillators

BERNIER, L.G.

A prediction method applicable to steered timescales

Geodetic frequency transfer. Actual developments in the AIUB-METAS collaboration

Predictability of a hydrogen maser generated timescale

BIZE, S.

Direct comparison between two atomic fountain clocks

Comparison between remote Cs fountain primary frequency standards

New microwave synthesizers based on 1 GHz frequency dissemination for metrological applications

BOUDOT, R.

Phase noise modelling and optimisation of microwave SiGe amplifiers - 1/f and HF noise contributions

Temperature regulated sapphire resonators for low phase noise oscillators

BOURGEOIS, P.Y.

Simple model for the mode splitting effect in whispering gallery mode resonator

Observation of a bistability effect in a cryogenic microwave whispering gallery mode resonator

BOURQUIN, R.

Frequency-temperature characteristic of flexural resonators using the extension of Timoshenko's model

Finite element model of nanobalance in GaPO4

Design of new miniature resonant temperature sensor in NLC-cut of quartz

BOY, J.J.

Ferrobielastic behaviour of 32 class crystals: LGS, LGT and GaPO4

Space radiation effects on ultra stable oscillators

BRUNET, M.

Dual frequency absolute calibration of GPS receiver for time transfer

Space radiation effects on ultra stable oscillators

The French time reference UTC(OP) and the connection of the EGNOS network time

BRUNIAUX, M.

Development of a parametric quartz crystal oscillator for low noise application

A very high speed method to simulate quartz crystal oscillator

BUSCA, G.

Results of field trial campaign for the performance assessment of an innovative time & frequency dissemination system based on ESA patent 407

Effect of hydrogen flux on loop control model in passive hydrogen Maser

CHAMBON, D.

Direct comparison between two atomic fountain clocks

New microwave synthesizers based on 1 GHz frequency dissemination for metrological applications

CIBIEL, G.

Phase noise modelling and optimisation of microwave SiGe amplifiers - 1/f and HF noise contributions

Dual frequency absolute calibration of GPS receiver for time transfer

Space radiation effects on ultra stable oscillators

Development of a parametric quartz crystal oscillator for low noise application

New frequency comparisons using GPS carrier phase at CNES

CLAIRON, A.

Direct comparison between two atomic fountain clocks

Comparison between remote Cs fountain primary frequency standards

New microwave synthesizers based on 1 GHz frequency dissemination for metrological applications

Optimization of the factors determinating the short term stability in a Raman-Ramsey CPT frequency standard

Short term stability limitations in a compact cold atom clock

CORDARA, F.

Galileo system test bed V1 : results on the experimental Galileo System Time

The Galileo time service provider: A progress report

DACH, R.

Time and frequency comparisons between four European timing institutes and NIST using multiple techniques

Geodetic frequency transfer. Actual developments in the AIUB-METAS collaboration

DANIAU, W.

New results on high frequency surface acoustic waves excited on an oriented LiNbO3, single crystal layer transferred onto Silicon

Theoretical analysis of damping effects of acoustic waves at solid/fluid interfaces using a mixed periodic FEA/BEM approach

DELPORTE, J.

Coherent population trapping resonances in thermal 87RB vapour

A very high speed method to simulate quartz crystal oscillator

New frequency comparisons using GPS carrier phase at CNES

Short term stability limitations in a compact cold atom clock

DIMARCQ, N.

Optimization of the factors determinating the short term stability in a Raman-Ramsey CPT frequency standard

Long term stability and accuracy of TAI

DUDLE, G.

Geodetic frequency transfer. Actual developments in the AIUB-METAS collaboration

Two-dimensional collimation of a continuous cesium fountain in an optical lattice

FOKS, A.

Calibration of GPS time receivers

Frequency biases calibration of GLONASS P-Code time receivers

FUJIEDA, M.

Development of time management system for satellite positioning system using quasi-zenith satellites

TWSTFT plans in the Pacific-Rim region and Europe link

Calibration system of internal delay Difference for TWSTFT station

GIORDANO, V.

Simple model for the mode splitting effect in whispering gallery mode resonator

Ti3+ doped microwave temperature compen-sated sapphire resonator oscillator at 34K

Observation of a bistability effect in a cryogenic microwave whispering gallery mode resonator

Phase noise modelling and optimisation of microwave SiGe amplifiers - 1/f and HF noise contributions

Temperature regulated sapphire resonators for low phase noise oscillators

GUERANDEL, S.

Long term stability and accuracy of TAI

Optimization of the factors determinating the short term stability in a Raman-Ramsey CPT frequency standard

HARTNETT, J.G.

Ti3+ doped microwave temperature compen-sated sapphire resonator oscillator at 34K

Rotating Michelson-Morley experiment based on a dual cavity cryogenic sapphire oscillator

HLAVAC, R.

Time and frequency comparisons between four European timing institutes and NIST using multiple techniques

Calibration of four European TWSTFT earth stations with a portable station through Intelsat 903

HOLLEVILLE, D.

Optimization of the factors determinating the short term stability in a Raman-Ramsey CPT frequency standard

Long term stability and accuracy of TAI

ITO, H.

New cesium atomic fountain frequency standard at NICT

Development of time management system for satellite positioning system using quasi-zenith satellites

KALLIOMAKI, K.

Short and long term stability of Kvarz hydrogen Masers CH1-76 and CH1-75A

A new 25 MHz standard frequency and time transmitter for local service

The effect of counter's trigger level on timing below nanosecond

KERSALE, Y.

Simple model for the mode splitting effect in whispering gallery mode resonator

Ti3+ doped microwave temperature compen-sated sapphire resonator oscillator at 34K

Observation of a bistability effect in a cryogenic microwave whispering gallery mode resonator

KOSYKH, A.V.

Independent excitation of two own oscillation modes in resonators with mixed (inclined) type of excitation field

Computer spice simulation of oscillation stability area for dual-mode crystal oscillator

Dual-mode crystal oscillators for DTCXO and OCXO

LARDET-VIEUDRIN, F.

Realization of a 5 and 10 MHz high performance isolation distribution amplifier for short term frequency stability measurements of frequency sources

Time and Frequency metrology accredited laboratories in Besançon

LAUDE, V.

New results on high frequency surface acoustic waves excited on an oriented LiNbO3, single crystal layer transferred onto Silicon

Longitudinally-coupled bulk acoustic wave resonators filter for 5 GHz applications

Theoretical analysis of damping effects of acoustic waves at solid/fluid interfaces using a mixed periodic FEA/BEM approach

LEWANDOWSKI, W.

On the evaluation of uncertainties in UTC-UTC(k)

Calibration of GPS time receivers

Frequency biases calibration of GLONASS P-Code time receivers

LIAO, C.S.

Ionospheric effects and modeling on GPS time transfer

Using DSP technique to average ADC samples for improving the frequency measurement resolution

Design and Implementation of a standard time tracking and keeping system for the real-time clock

LLOPIS, O.

Phase noise modelling and optimisation of microwave SiGe amplifiers - 1/f and HF noise contributions

Temperature regulated sapphire resonators for low phase noise oscillators

LORINI, L.

Comparison between remote Cs fountain primary frequency standards

Time and frequency comparisons between four European timing institutes and NIST using multiple techniques

Galileo system test bed V1 : results on the experimental Galileo System Time

MAENO, H.

TWSTFT plans in the Pacific-Rim region and Europe link

Calibration system of internal delay Difference for TWSTFT station

MANSTEN, T.

Short and long term stability of Kvarz hydrogen Masers CH1-76 and CH1-75A

A new 25 MHz standard frequency and time transmitter for local service

The effect of counter's trigger level on timing below nanosecond

MARION, H.

Direct comparison between two atomic fountain clocks

Comparison between remote Cs fountain primary frequency standards

New microwave synthesizers based on 1 GHz frequency dissemination for metrological applications

MEDVEDEV, S.Y.

Frequency calibration tests of new industrial passive hydrogen maser frequency and time standards VCH-1006

Designing and investigation of digital multi-channel temperature control system with high performance for hydrogen time and frequency standard

MEYER, F.

Time and Frequency metrology accredited laboratories in Besançon

Frequency transfer using low cost GPS receivers, accessing national frequency references. SYREF, example of an integrated solution

NAWROCKI, J.

Calibration of GPS time receivers

Frequency biases calibration of GLONASS P-Code time receivers

ORGIAZZI, D.

Galileo system test bed V1 : results on the experimental Galileo System Time

The Galileo time service provider: A progress report

Determination of GPS/GALILEO time offset: insights from the GSTB-V1 results

PETIT, G.

Long term stability and accuracy of TAI

Time and frequency comparisons between four European timing institutes and NIST using multiple techniques

QUOC TUNG, L.

TWSTFT plans in the Pacific-Rim region and Europe link

Calibration system of internal delay Difference for TWSTFT station

RATIER, N.

Development of a parametric quartz crystal oscillator for low noise application

A very high speed method to simulate quartz crystal oscillator

REINHARDT, A.

Longitudinally-coupled bulk acoustic wave resonators filter for 5 GHz applications

Theoretical analysis of damping effects of acoustic waves at solid/fluid interfaces using a mixed periodic FEA/BEM approach

SALOMON, C.

Direct comparison between two atomic fountain clocks

Comparison between remote Cs fountain primary frequency standards

SALZENSTEIN, P.

Realization of a 5 and 10 MHz high performance isolation distribution amplifier for short term frequency stability measurements of frequency sources

Time and Frequency metrology accredited laboratories in Besançon

SANTARELLI, G.

Direct comparison between two atomic fountain clocks

New microwave synthesizers based on 1 GHz frequency dissemination for metrological applications

SESIA, I.

Galileo system test bed V1 : results on the experimental Galileo System Time

Determination of GPS/GALILEO time offset: insights from the GSTB-V1 results

SHMALIY, Y.

Power dependence of distributed cavity phase induced frequency biases in atomic fountain

Statistically limiting errors estimation in passive remote SAW sensing employing differential phase measurements

SPASSOV, L.

The effect of different electrode films on the resonator parameters

Design of new miniature resonant temperature sensor in NLC-cut of quartz

STEICHEN, W.

New results on high frequency surface acoustic waves excited on an oriented LiNbO3, single crystal layer transferred onto Silicon

Longitudinally-coupled bulk acoustic wave resonators filter for 5 GHz applications

STHAL, F.

Frequency-temperature characteristic of flexural resonators using the extension of Timoshenko's model

Finite element model of nanobalance in GaPO4

TAKAHASHI, Y.

Development of time management system for satellite positioning system using quasi-zenith satellites

TWSTFT plans in the Pacific-Rim region and Europe link

TAVELLA, P.

On the evaluation of uncertainties in UTC-UTC(k)

Galileo system test bed V1 : results on the experimental Galileo System Time

The Galileo time service provider: A progress report

Steering a time scale by the application of the theory of control

Determination of GPS/GALILEO time offset: insights from the GSTB-V1 results

THOMANN, P.

Two-dimensional collimation of a continuous cesium fountain in an optical lattice

A Feasibility study of an optically-pumped cesium atomic resonator for space applications

TOBAR, M.E.

Direct comparison between two atomic fountain clocks

Ti3+ doped microwave temperature compen-sated sapphire resonator oscillator at 34K

Rotating Michelson-Morley experiment based on a dual cavity cryogenic sapphire oscillator

New microwave synthesizers based on 1 GHz frequency dissemination for metrological applications

TSARAPKIN, D.

Comparison between remote Cs fountain primary frequency standards

Resistive coupling in microwave frequency stabilization

UHRICH, P.

Time and frequency comparisons between four European timing institutes and NIST using multiple techniques

The French time reference UTC(OP) and the connection of the EGNOS network time

UNDERHILL, M.

A Spice System Simulation Model for Phase-Lock Loops and anti-Jitter Circuits

Demodulation of noise very close to carrier and estimation of total jitter

WILM, M.

Theoretical analysis of damping effects of acoustic waves at solid/fluid interfaces using a mixed periodic FEA/BEM approach

Quartz tuning fork vibration amplitude as a limitation of spatial resolution of shear force microscopes

WOLCOFF, B.

High frequencies monolithic quartz filters

Temporal stability measurement on OCXO running under severe environmental conditions

NCXO : New developments for avionic applications

New high stability 10 MHz OCXO family

WOLF, P.

An optical clock with atoms confined in an accelerated shallow trap

Rotating Michelson-Morley experiment based on a dual cavity cryogenic sapphire oscillator

ZHGOON, S.

Small RF filters made with thin-film high temperature superconductors

Controllable trimming of SAW filters and resonators on Langasite and Lithium Tantalate

Low-Loss NSPUDT SAW filters using the inclined reflectors on Lithium Tantalate

19th European Frequency and Time Forum, Besançon, France

Opening session

Synthetic aperture radars – Simon WATTS, UK (Invited paper)

Session 4C / Optical Frequency Standards
Chairman : Andreas BAUCH

Optical lattice clock – H. KATORI, M. TAKAMOTO, R. HIGASHI, F.L. HONG, Japan (Invited paper)

Stability and accuracy of 171Yb+ single-ion optical frequency standards – T. SCHNEIDER, C. TAMM, E. PEIK, Germany.

An optical clock with atoms confined in an accelerated shallow trap – P. LEMONDE, P. WOLF, France.

Session 4A / Cold atom clocks
Chairman : Pierre THOMANN

Direct comparison between two atomic fountain clocks – C. VIAN, P. ROSENBUSCH, H. MARION, S. BIZE, L. CACCIAPUOTI, S. ZHANG, M. ABGRALL, D. CHAMBON, I. MAKSIMOVIC, C. MANDACHE, P. LAURENT, G. SANTARELLI, A. CLAIRON, A. LUITEN, M. TOBAR, C. SALOMON, France/Australia.

Power dependence of distributed cavity phase induced frequency biases in atomic fountain frequency standards – S.R. JEFFERTS, J.H. SHIRLEY, N. ASHBY, E.A. BURT, G.J. DICK, USA.

Microwave cavities for atomic clocks :phase variations and associated power dependences – R. LI, K. GIBBLE, USA.

Comparison between remote Cs fountain primary frequency standards - S. BIZE, H. MARION, A. CLAIRON, S. WEYERS, A. BAUCH, R.WYNANDS, D. CALONICO, L. LORINI, F. LEVI, K. SZYMANIEC, W. CHALUPCZAK, D. HENDERSON, France/Germany/Italy/UK.

Session 5A / Time scales, algorithms and analysis
Chairman : Elisa Felicitas ARIAS

A prediction method applicable to steered timescales – L.G. BERNIER, Switzerland.

Long term stability and accuracy of TAI – G. PETIT, France.

On the evaluation of uncertainties in UTC-UTC(k) – W. LEWANDOWSKI, D. MATSAKIS, G. PANFILO, P. TAVELLA, France/USA/Italy.

The development of a Kalman filter clock predictor – J.A. DAVIS, C.A. GREENHALL, R. BOUDJENA, UK/USA.

UTC(SU) steering time scale strategy data stimulation and first results – N. KOSHELYAEVSKY, E. ZAGIROVA, Russia.

19th European Frequency and Time Forum, Besançon, France

Session 5C / Time and frequency transfers
Chairman : Patrizia TAVELLA

Time and frequency comparisons between four European timing institutes and NIST using multiple techniques – A.BAUCH, R. DACH, J. DAVIS, L. LORINI, T. PARKER, G. PETIT, J.Y. RICHARD, Gemrany/Switzerland/Italy/UK/USA/France.

Results of field trial campaign for the performance assessment of an innovative time & frequency dissemination system based on ESA patent 407 – D. GIUNTA, A. DELLA TORRE, M. VITTA, V. DE PERINI, R. GRIMOLDI, G. BUSCA, P. ROCHAT, X. STEHLIN, Switzerland/Italy.

Geodetic frequency transfer. Actual developments in the AIUB-METAS collaboration – R. DACH, T. SCHILDKNECHT, U. HUGENTOBLER, L.G. BERNIER, G. DUDLE, Switzerland.

Session 3C / Cryogenic oscillators
Chairman : Mike UNDERHILL

Simple model for the mode splitting effect in whispering gallery mode resonator – V. GIORDANO, P.Y. BOURGEOIS, N. BAZIN, Y. KERSALE, France.

Ti3+ doped microwave temperature compen-sated sapphire resonator oscillator at 34K – Y. KERSALE, N. BOUBEKEUR, J.G. HARTNETT, M.E. TOBAR, N. BAZIN, V. GIORDANO, France/Australia.

Observation of a bistability effect in a cryogenic microwave whispering gallery mode resonator – V. GIORDANO, P.Y. BOURGEOIS, M. OXBORROW, N. BAZIN, Y. KERSALE, France/UK.

Rotating Michelson-Morley experiment based on a dual cavity cryogenic sapphire oscillator – P.L. STANWIX, M.E. TOBAR, J. WINTERFLOOD, P. WOLF, E.N. IVANOV, M. SUSLI, J.G. HARTNETT, F. VAN KANN, Australia/France.

Session 5B / GNSS time scales
Chairman : Michel BRUNET

Galileo system test bed V1 : results on the experimental Galileo System Time – F. CORDARA, R. COSTA, L. LORINI, D. ORGIAZZI, V. PETTITI, I. SESIA, P. TAVELLA, P. ELIA, M. MASCARELLO, M. FALCONE, J. HAHN, Italy/Netherlands.

Galileo Time Service Provider - The focal point for European timing expertise – J. HAMMESFAHR, A. MOUDRAK, J. FURTHNER, A. SCHROTH, J. DALLAT, R. PIRIZ, A. BAUCH, C. BRUYNINX, P. DEFRAIGNE, Germany/Belgium.

The Galileo time service provider: A progress report – J. LAVERTY, A. BAUCH, F. CORDARA, D. ORGIAZZI, P. TAVELLA, UK/Germany/Italy.

Calibration of GPS time receivers – J. NAWROCKI, P. MIELNIK, Poland.

Developpement of the $^{113}Cd^{+}$ ion atomic clock - B.M. JELENKOVIC, J.D. PRESTAGE, L. MALEKI, USA.

19th European Frequency and Time Forum, Besançon, France

Session 1A / Piezoelectric resonators, filters and materials
Chairman : Emmanuel BIGLER

Frequency-temperature characteristic of flexural resonators using the extension of Timoshenko's model – H. HANTZ, F. STHAL, R. BOURQUIN, France.

A perturbation method for accurate predictions of mass loading and thermal effects on acoustic devices computed using finite element analysis – S. BALLANDRAS, France.

The influence of the material constants on Langasite resonators – I. MATEESCU, C. BRAN, S. GEORGESCU, Romania.

New results on high frequency surface acoustic waves excited on an oriented LiNbO3, single crystal layer transferred onto Silicon – T. PASTUREAUD, B. BIASSE, B. ASPAR, W. DANIAU, W. STEICHEN, V. LAUDE, R. LARDAT, A. LAENS, S. BALLANDRAS, France.

Chemical controlled GaPO4 dissolution High frequency antimesa resonator manufacturing – N. PRUDHOMME, D. CACHAU-HERREILLAT, P. PAPET, O. CAMBON, France.

Posters Session 1

Design and first results on the New PTB caesium fountain CSF2 – S. WEYERS, A. BAUCH, D. GRIEBSCH, R. SCHROEDER, R. WYNANDS, Germany.

Small RF filters made with thin-film high temperature superconductors – S. ZHGOON, A. BARINOV, O. SHTEYNBERG, P. MOZHAEV, Russia.

Controllable trimming of SAW filters and resonators on Langasite and Lithium Tantalate – S. ZHGOON, A.S. SHVETSOV, Russia.

Statistically limiting errors estimation in passive remote SAW sensing employing differential phase measurements – Y. SHMALIY, O. IBARRA-MANZANO, Mexico.

Low-Loss NSPUDT SAW filters using the inclined reflectors on Lithium Tantalate – S. DOBERSTEIN, A. SHVETSOV, S. ZHGOON, Russia.

Coherent population trapping resonances in thermal 87RB vapour – G. KAZAZOV, I. MAZETS, Y. ROZHDESTVENSKY, B. MATISOV, Russia.

Design of GaAs resonant structures for sensors applications – T. LEBLOIS, C. TELLIER, France.

The effect of different electrode films on the resonator parameters – V. GEORGIEVA, L. SPASSOV, T. YORDANOV, Bulgaria.

Effect of hydrogen flux on loop control model in passive hydrogen Maser – Q. WANG, G. BUSCA, Switzerland.

Towards an optical frequency standard at 729 nm – G. HAGEL, C. CHAMPENOIS, T.B. EKOGO, M. HOUSSIN, M. KNOOP, C. LISOWSKI, M. VEDEL, F. VEDEL, France.

19th European Frequency and Time Forum, Besançon, France

The vortices frequency measurement in piezoceramic vortex flow meter and their applications– K. WEISS, S. GAWOR, A. LISOWIEC, M. TUPAJ, Poland.

Shangai astronomical observatory's hydrogen masers – Z.C. ZHAI, C.F. LIN, W.Q. ZHANG, J.X. PENG, Y.P. ZHANG, Y. CAI, China.

Improved theory for higher-order blackbody radiation shift in Cs clocks – Y.S. DOMNIN, V.G. PAL'CHIKOV, Russia.

Short and long term stability of Kvarz hydrogen Masers CH1-76 and CH1-75A – K. KALLIOMAKI, T. MANSTEN, Finland.

Independent excitation of two own oscillation modes in resonators with mixed (inclined) type of excitation field – A.V. KOSYKH, A.N. LEPETAEV, Russia.

Two-dimensional collimation of a continuous cesium fountain in an optical lattice – N. CASTAGNA, G. DI DOMENICO, G. DUDLE, M.D. PLIMMER, A.V. TAICHENACHEV, P. THOMANN, V.I. YUDIN, Switzerland/Russia.

An optical clock with ultracold calcium atoms – T. NAZAROVA, C. GRAIN, C. DEGENHARDT, P.E. POTTIE, H. SCHNATZ, U. STERR, F. RIEHLE, C. LISDAT, H. STOEHR, Germany.

New cesium atomic fountain frequency standard at NICT – M. KUMAGAI, H. ITO, M. KAJITA, M. HOSOKAWA, Japan.

High frequencies monolithic quartz filters – B. D'ALBARET, B. WOLCOFF, S. MICHEL, P. CAPPELLE, France.

Atomic fountain clock and cold atom gravimeter : two atomic interferometers – F. DEYZAC, S. BOUSSEN, F. LIENHART, A. BRESSON, France.

Investigation of spectral devices source for miniaturized rubidium gas cell atomic frequency standards – A.B. BASSEVICH, O.V. BERESOVSKAYA, G.M. SMIRNOVA, Russia.

Information and Reference System for Acoustoelectornic device development – O.L. BALYSHEVA, V.V. KLUDZIN, L.P. KONOVALOVA, S.V. KULAKOV, K.I. KURSANOVA, O.N. NOVIKOVA, Y.G. SMIRNOV, Russia.

Longitudinally-coupled bulk acoustic wave resonators filter for 5 GHz applications – A. REINHARDT, F. MARTIN, R. LANZ, P. MURALT, V. LAUDE, S. BALLANDRAS, W. STEICHEN, France/Switzerland.

Third-Order Elastic Constants for Si from Fisrt Principles – K. VOLIANSKIY, Russia.

Ferrobielastic behaviour of 32 class crystals: LGS, LGT and GaPO4 – J.J. BOY, M. ASSOUD, K. YAMNI, A. ALBIZANE, Marocco/France.

Investigation of frequency stability of active hydrogen maser with autonomous cavity auto tuning using frequency switching method – N. DEMIDOV, V.A. LOGATCHEV, V.S. GOREV, V.G. VORONTSOV, Russia.

Signal processing in microwave autodyne sensor – V.V. BOLOZNEV, E.V. SAFONOVA, M.A. STANCHENKOV, F.I. SULTANOV, Russia.

19th European Frequency and Time Forum, Besançon, France

Frequency calibration tests of new industrial passive hydrogen maser frequency and time standards VCH-1006 – A. BELIAEV, S. MEDVEDEV, I. PAVLENKO, B. SAKHAROV, P. SMIRNOV, M. BLOCH, O. MANCINI, Russia/USA.

The efficiency investigation of 87Rb atomic beam laser pumping for designing a quantum discriminator for high space-borne atomic beam frequency standard – A. BESEDINA, A. GEVORKYAN, V. ZHOLNEROV, Russia.

Session 3B / VCO & synthetisers
Chairman : Jean-Pierre AUBRY

A Spice System Simulation Model for Phase-Lock Loops and anti-Jitter Circuits – M. UNDERHILL, J. BRODERICK, UK.

An integrated differential single-chip VCO for S-Band – H. SALMINEN, P. ESKELINEN, J. HOLMBERG, Finland.

Phase noise modelling and optimisation of microwave SiGe amplifiers - 1/f and HF noise contributions – S. GRIBALDO, R. BOUDOT, G. CIBIEL, V. GIORDANO, O. LLOPIS, France.

New microwave synthesizers based on 1 GHz frequency dissemination for metrological applications – D. CHAMBON, S. BIZE, F. NARBONNEAU, H. MARION, M. LOURS, A. CLAIRON, M. TOBAR, G. SANTARELLI, France/Australia.

Session 5D / Measuring methods and instrumentation
Chairman : Giovanni BUSCA

Calibration of four European TWSTFT earth stations with a portable station through Intelsat 903 – D. PIESTER, R HLAVAC, J. ACHKAR, G. de JONG, B. BLANZANO, H. RESSLER, J. BECKER, P. MERCK, O. KOUDELKA, Germany/UK/France/Netherlands/Austria.

Dual frequency absolute calibration of GPS receiver for time transfer – G. CIBIEL, M. BRUNET, J. DANTEPAL, J.F. DUTREY, A. DE LATOUR, France.

On twofold use of dual-mixer time-difference multiplication – R. BARILLET, J. CERMAK, L. SOJDR, France/Czech Republic.

Time and frequency transfer across Gigabit Ethernet networks – M. KIHARA, K. HISADOME, Japan.

Session 3A / Environmental effects on frequency standards
Chairman : Vincent GIORDANO

A Phase-locked crystal oscillator design to withstand long term vibration loading– H. ESKELINEN, J.M. HEINOLA, E. HAVIA, P. SILVENTOINEN, Finland.

Space radiation effects on ultra stable oscillators – M. BRUNET, G. CIBIEL, R. ECOFFET, G. ROLLAND, J.J. BOY, France.

Temporal stability measurement on OCXO running under severe environmental conditions – O. FRANQUET, O. ROMAND, L. CAMELIN, B. WOLCOFF, France.

Temperature regulated sapphire resonators for low phase noise oscillators – R. BOUDOT, C. ROCHER, N. BAZIN, V. GIORDANO, O. LLOPIS, France.

Session 3D / Oscillators and noise
Chairman : Pekka ESKELINEN

Demodulation of noise very close to carrier and estimation of total jitter – M. UNDERHILL, UK.

Realization of a 5 and 10 MHz high performance isolation distribution amplifier for short term frequency stability measurements of frequency sources – P. SALZENSTEIN, F. LARDET-VIEUDRIN, France.

Development of a parametric quartz crystal oscillator for low noise application – S. GALLIOU, R. BRENDEL, G. CIBIEL, France.

A very high speed method to simulate quartz crystal oscillator – N. RATIER, M. BRUNIAUX, J. DELPORTE, France

Posters Session 2

Noise power conversion in oscillatory systems: the method – Y. SHMALIY, Mexico.

New development of PM and AM noise analysis in crystal oscillators : an influence of Wide band noise sources – V. KULESHOV, Russia.

New development of PM and AM noise analysis in crystal oscillators: an influence of 1/f noise sources – T. I. BOLDYREVA, Russia.

Predictability of a hydrogen maser generated timescale – L.G. BERNIER, Switzerland.

Time scale prediction using information from external sources – A. BAUCH, A. SEN GUPTA, E. STALIUNIENE, Germany/India.

Influence of the operational dead time of atomic fountain frequency standards to its comparison with TAI at NMIJ – S. YANAGIMACHI, T. IKEGAMI, T. KUROSU, Japan.

Development of time management system for satellite positioning system using quasi-zenith satellites – Y. TAKAHASHI, J. AMAGAI, M. FUJIEDA, S. YOKOTA, K. KIMURA, H. ITO, S. AMA, T. MORIKAWA, I. KAWANO, S. KOGURE, Japan.

TWSTFT plans in the Pacific-Rim region and Europe link – H. MAENO, M. FUJIEDA, M. AIDA, L. QUOC TUNG, R. TABUCHI, F. NAKAGAWA, T. SUZUYAMA, Y. TAKAHASHI, Japan.

Steering a time scale by the application of the theory of control – P. TAVELLA, L. GALLEANI, M. FARINA, Italy.

19th European Frequency and Time Forum, Besançon, France

Typical combined uncertainty evaluation On the Ku Band TWSTFT Link – J. ACHKAR, P. MERCK, France.

A new 25 MHz standard frequency and time transmitter for local service – K. KALLIOMAKI, T. MANSTEN, Finland.

The effect of counter's trigger level on timing below nanosecond – K.KALLIOMAKI, T. MANSTEN, J. MANNERMAA, Finland.

Computer spice simulation of oscillation stability area for dual-mode crystal oscillator – A. KOSYKH, A. GUBAREV, S. ZAVJALOV, I. KHOMENKO, Russia.

Dual-mode crystal oscillators for DTCXO and OCXO – A. KOSYKH, Russia.

Sub nanosecond calibration of regional time links using continuous GNSS carrier phase observations – C. RIECK, P. JARLEMARK, K. JALDEHAG, J. JOHANSSON, Sweden.

Time metrology in Romania – A. NICULESCU, Romania.

Ionospheric effects and modeling on GPS time transfer – H.M. PENG, F.D. CHUU, C.S. LIAO, Taiwan.

Timing and optical fiducial systems for the Laser MégaJoule – M. PRAT, P. RAYBAUT, France.

Dynamic parameters of synchronization signals for non-overlapping and overlapping data segments – M. KASZNIA, A. DOBROGOWSKI, Poland.

Calibration system of internal delay Difference for TWSTFT station – M. FUJIEDA, M. AIDA, H. MAENO, L. QUOC TUNG, Japan.

Determination of GPS/GALILEO time offset: insights from the GSTB-V1 results – I. SESIA, D. ORGIAZZI, P. TAVELLA, Italy.

Frequency biases calibration of GLONASS P-Code time receivers – W. LEWANDOWSKI, A. FOKS, J. NAWROCKI, Poland/France.

NCXO : New developments for avionic applications – P. DE JOFFREY, P. VUILLEMIN, B. WOLCOFF, France.

New high stability 10 MHz OCXO family – E. GIRARDET, H. CHARTOGNE, P. RENOULT, B. WOLCOFF, France.

The French time reference UTC(OP) and the connection of the EGNOS network time – P. UHRICH, D. VALAT, M. BRUNET, J. MARECHAL, T. BELTAN, France.

Using DSP technique to average ADC samples for improving the frequency measurement resolution – S.S. CHEN, C.S. LIAO, Taiwan.

Design and Implementation of a standard time tracking and keeping system for the real-time clock – C.C. LIN, C.S. LIAO, Taiwan.

Designing and investigation of digital multi-channel temperature control system with high performance for hydrogen time and frequency standard – A.Y. PAVLENKO, S.Y. MEDVEDEV, P.N. SMIRNOV, Russia.

Time and Frequency metrology accredited laboratories in Besançon – P. SALZENSTEIN, F. MEYER, E. TISSERAND, G. MARTIN, J. PETETIN, N. FRANQUET, F. LARDET-VIEUDRIN, O. LAVEDRINE, M. VINCENT, France.

Frequency transfer using low cost GPS receivers, accessing national frequency references. SYREF, example of an integrated solution – F. MEYER, N. GAUTHEROT, France.

New frequency comparisons using GPS carrier phase at CNES – J. DELPORTE, F. MERCIER, G. CIBIEL, France.

Session 4B / Compact clocks
Chairman : François VERNOTTE

Chip-scale atomic frequency references – J. KITCHING, S. KNAPPE, L.A. LIEW, P.D.D. SCHWINDT, V. SHAH, J. MORELAND, L. HOLLBERG, USA (Invited Paper).

A Feasibility study of an optically-pumped cesium atomic resonator for space applications – P. BERTHOUD, R. RUFFIEUX, P. THOMANN, Switzerland.

Optimization of the factors determinating the short term stability in a Raman-Ramsey CPT frequency standard – T. ZANON, S. GUERANDEL, E. DE CLERCQ, D. HOLLEVILLE, N. DIMARCQ, A. CLAIRON, France.

Short term stability limitations in a compact cold atom clock – S. TREMINE, S. GUERANDEL, D. HOLLEVILLE, A. CLAIRON, J. DELPORTE, N. DIMARCQ, France.

Session 2A / Piezoelectric sensors
Chairman : Peter KREMPL

F2-laser ablation and surface-patterning of GaPO4 for sensor applications – M. PERUZZI, J.D. PEDARNIG, H. SALHOFER, R. SCHWOEDIAUER, W. REICHL, J. RUNCK, H. MITTER, Austria.

Theoretical analysis of damping effects of acoustic waves at solid/fluid interfaces using a mixed periodic FEA/BEM approach – S. BALLANDRAS, A. REINHARDT, A. KHELIF, M. WILM, V. LAUDE, W. DANIAU, V. BLONDEAU-PATISSIER, G. LENGAIGNE, France.

Finite element model of nanobalance in GaPO4 – L. DELMAS, F. STHAL, M. MINUNNI, E. BIGLER, R. BOURQUIN, M. MASCINI, France/Italy.

Quartz tuning fork vibration amplitude as a limitation of spatial resolution of shear force microscopes – J.M. FRIEDT, E. CARRY, B. SERIO, J.J. HUNSINGER, M. WILM, S. BALLANDRAS, France.

Design of new miniature resonant temperature sensor in NLC-cut of quartz – B. DULMET, L. SPASSOV, R. BOURQUIN, A. IVAN, T. ANGELOV, J. LAZAROV, France/Bulgaria.