Accuracy Metrics for Judging Time Scale Algorithms

R.J. Douglas, J.-S. Boulanger and C. Jacques
Time and Frequency Standards Group
Institute for National Measurement Standards
National Research Council of Canada
Ottawa, CANADA K1A 0R6
(613) 993-5186, Fax (613) 952-1394

Abstract

Time scales have been constructed in different ways to meet the many demands placed upon them for time accuracy, frequency accuracy, long-term stability and robustness. Usually, no single time scale is optimum for all purposes. In the context of the impending availability of high-accuracy intermittently-operated cesium fountains, we reconsider the question of evaluating the accuracy of time scales which use an algorithm to span interruptions of the primary standard.

We consider a broad class of calibration algorithms that can be evaluated and compared quantitatively for their accuracy in the presence of frequency drift and a full noise model (a mixture of white PM, flicker PM, white FM, flicker FM and random walk FM noise). We present the analytic techniques for computing the standard uncertainty for the full noise model and this class of calibration algorithms. The simplest algorithm is evaluated to find the average-frequency uncertainty arising from the noise of the cesium fountain's local oscillator and from the noise of a hydrogen maser transfer-standard. This algorithm and known noise sources are shown to permit interlaboratory frequency transfer with a standard uncertainty of less than 10^{-15} ***for periods of 30-100 days.***

Introduction: The Need for Evaluating Algorithm Accuracy

For the near future, new primary (cesium fountain) frequency standards [1] [2] are likely to operate intermittently, rather than to operate continuously as primary clocks. Other new frequency standards of high potential accuracy, such as single-ion optical frequency standards coupled to a divider chain [3], are also likely to operate intermittently, at least initially. More reliable secondary standards of high stability (such as hydrogen masers) will be needed to span the gaps between periods of operation of the primary standard. To evaluate the accuracy of time scales that are to be calibrated with these new standards, one must address the question of how the random noise of the primary standard, of its local oscillator, and of the secondary standards all combine to influence the accuracy of the time scale or its average frequency. We want to examine how these noise sources affect the results of different interpolation and extrapolation algorithms, and to predict the accuracy that could be delivered, in the presence of mixed types of noise, to a local time scale or to TAI. Our main interest is in the frequency accuracy of the secondary time scale after calibration

using some algorithm, but many of the same ideas are also applicable to time accuracy around an interval.

This is not a wholly new question. Intermittent operation of primary cesium frequency standards was universal two decades ago. Continuous operation of these standards as primary clocks, first at NRC [4] and then at PTB, waited until time laboratories had sufficiently evaluated their primary cesium frequency standards and had improved their mean time between failures. Standard techniques for analyzing frequency standards' stability and characterizing mixed types of noise were adopted and refined [5] [6] [7]. A body of very useful guidance was developed [8] for extrapolation in the presence of different (but unmixed) types of noise.

The main new element to be addressed is the quantitative estimation of accuracy for the mixed types of noise which have to be faced for our problem, and which has not been needed for previous standards where the dominant noise has usually been flicker frequency noise for primary and secondary standards. Simulations [9] could give the required guidance, but fully analytic techniques are preferable, and are developed here for a widely used class of algorithms.

Metrics

In choosing and in judging accuracy of the "optimum" algorithm for a purpose, a *metric* should be used for ordering possible algorithms and for guidance of minimal ambiguity. A priori, there are many possible metrics.

The class of metrics of interest to us quantifies the difference between any two functions of time, $A(t)$ and $B(t)$, sampled at a set of discrete times $\{t_i\}, i = 0..N$ during the time interval $[t_0, t_N]$. A metric expresses as a real number the difference between two vectors $\mathbf{A}$ and $\mathbf{B}$ in this N-dimensional vector space: $\|\mathbf{A} - \mathbf{B}\|$, and permits the unambiguous ordering of the quality of a fit from "good" to "bad". Any metric must meet the requirements that (1) $\|\mathbf{A}\| \geq 0$, (2) $\|a\mathbf{A}\| = |a|\|\mathbf{A}\|$, and (3) $\|\mathbf{A} + \mathbf{B}\| \leq \|\mathbf{A}\| + \|\mathbf{B}\|$. A very useful subclass of metrics is the class of "Holder norms", or L_p-norms ($p \geq 1$): $\|\mathbf{A} - \mathbf{B}\| = [\sum_{i=0}^{N} w_i \mid (A_i - B_i) \mid^p]^{(1/p)}$. The weights w_i are positive definite. If the difference between A_i and B_i is a random variable z_i distributed around Z_i, and is described by a probability distribution $e^{-[w_i^p|(z_i - Z_i)|^p]}$, then the minimizing the corresponding L_p-norm will give the maximum likelihood fit of $\mathbf{A}$ to $\mathbf{B}$. When fitting an approximation to a mathematical function, the norm ($\lim p_{\to\infty}$) is usually used, as the min-max norm, to minimize the maximum error between the function and its approximation on the interval. The absolute-value norm ($p = 1$) is occasionally used as an uncritical way of fitting to give minimum fitting weight to erroneous outliers while formally retaining a metric. When fitting experimental data, where normal (gaussian) distributions are common, p=2 is generally appropriate. It is appropriate for describing our expected distributions, and we will concentrate on this type of metric.

When measuring the quality of a fit to the measurements at the N times t_i, the value of the metric, divided by the degrees of freedom (N minus the number of fitting parameters), is often used. For a least-squares fit ($p = 2$) this measure of the quality of fit becomes the square root of the familiar reduced χ^2, and for unweighted least-squares fits ($w_i = 1$) it is the even more familiar root-mean-square residual. The residual is formally a metric in an N-dimensional vector space. As they are

formally defined, weights can in principle be used to change the scale factor on each axis of the vector space, and even to project the metric into a vector subspace. Minimizing such a reduced metric rather than using the full maximum-likelihood weights can be advantageous: for example, the 10-day average frequency of a commercial cesium clock can be better determined from only the end points than by a least-squares fit to many points distributed across the 10-day interval [8].

Thus the quality of the fit at the experimental points is *not* all that is required. Rather, since the fitted function is used for interpolation to other values of t between t_i and t_{i+1}, or for extrapolation to values of t outside $[t_0, t_N]$, it is the accuracy at these points which can be more important. The accuracy of an algorithm is not uniform, but varies with t in a way which depends on the set of fitting points $\{t_i | i = 0..N\}$, the fitting algorithm and the random (and deterministic) noise. Considering this type of problem from the perspective of the residuals seems to require the magic of rotations into a different vector space. Interestingly, exactly this task can be done for the L_2 norm and a rather broad class of fitting functions, although the metric projection picture is unhelpful in determining the fitting accuracy at an arbitrary time.

The accuracy can be determined for any system experimentally by repeated cycles of measurements, doing repeated fitting of one particular pattern of time samples and by statistical analysis of residuals determined at unfitted points. Another approach would be computer simulation of this process - if a sufficiently good description of the noise model is available; or it might be done analytically. We show that a rather broad class of noise models and fitting procedures *can* be treated analytically, to obtain an accuracy estimate for the interpolation or extrapolation of many commonly used algorithms.

Modelling the Difference: Deterministic plus Random Noise

To describe the time-dependence $x(t)$ of the time difference between the primary frequency standard's time scale and the secondary time scale, we model it with $x_m(t)$, and explicitly include a random part $x_0(t)$ as well as a deterministic part. The deterministic part allows for a time offset, a frequency difference, and a drift rate of the frequency of the secondary time scale with respect to the primary standard.

$$x_m(t) = a^k + b^k t + \frac{1}{2} c^k t^2 + x_0(t). \tag{1}$$

The superscript k labels the uninterrupted intervals of operation of the primary standard. For each interval, a new value of a^k is required, and other values of b^k and c^k may (or may not) be used. We will examine the accuracy of a class of fitting functions $x_p(t)$, fit on the interval k, linear in the fitting parameters $\{d_l\}$, and including a broad class of basis functions $g_l(t)$:

$$x_p^k(t) = d_1^k + d_2^k t + d_3^k t^2 + \sum_{l=4}^{n} d_l^k g_l^k(t). \tag{2}$$

In the random noise part, $x_0(t)$, we include the "full" noise model that is usually covered in discus-

sions of frequency standard stability [5]: a sum of five noise processes, each normally distributed about the mean (but with variances which depend on the time sampled in different ways) that have spectral densities of phase noise ($S_x(t)$) that are power laws which range from flat to increasingly divergent at low frequencies. Expressing the five terms in terms of the spectral density of the mean-square of the fluctuations in $\frac{dx_0(t)}{dt}$ (or $y_0(t)$) at a frequency f, $S_y(f)$, each noise term is described by an amplitude h_α which is taken to be independent of any time translations (stationarity and random phase approximations). The sum includes ($\alpha = 2$) white phase noise in x, ($\alpha = 1$) flicker ($1/f$) noise in x, ($\alpha = 0$) white frequency noise and random walk phase noise, ($\alpha = -1$) flicker frequency noise and ($\alpha = -2$) random walk frequency noise. The spectral density of the mean-square fluctuations in $x_0(t)$ is $S_x(t)$, and for this noise model

$$S_y(f) = \sum_{\alpha=-2}^{2} h_\alpha f^\alpha \qquad \text{and} \qquad S_x(f) = \sum_{\alpha=-2}^{2} \frac{h_\alpha f^{(\alpha-2)}}{(2\pi)^2}. \tag{3}$$

A Useful Tool

For many cases we might expect mean-square accuracy estimates to be calculable from the autocorrelation function $\langle x_0(t)x_0(t+\tau)\rangle$. Although it is not easy to use, the autocorrelation function of frequency-standard noise has been rather neglected in our view. It is divergent for four of our five types of noise unless a low-frequency cutoff is applied, and even then can challenge the accuracy and dynamic range capacities of classical computing. Analytic expressions for this autocorrelation function are given in the Appendix for each type of noise, and modern arbitrary-precision computer languages should routinely be able to cope directly with the autocorrelation function. In our analysis of the uncertainty associated with any useful time algorithm we expect no divergences to infinity, and so the combinations of the autocorrelation functions must have their divergent parts cancel, with the algorithm itself acting as low-frequency cutoff. To simplify the numerical evaluation of combinations of the autocorrelation, it can be useful to find an analytic expression for the general two-interval covariance of the random noise model, that is the covariance of the time-scale departure over the time interval $[t_1, t_2]$ with the time-scale departure over the time interval $[t_3, t_4]$:

$$\mathcal{S} \;=\; \langle [x_0(t_2) - x_0(t_1)]\,[x_0(t_4) - x_0(t_3)]\rangle \;=\; \int_{t_1}^{t_2}\int_{t_3}^{t_4} \langle y_0(t')\,y_0(t'')\rangle\, dt''\, dt'. \tag{4}$$

This is just $(t_2 - t_1)(t_4 - t_3)$ times $\left\langle \bar{y}_{[t_1,t_2]}\bar{y}_{[t_3,t_4]} \right\rangle$, the general covariance of the average frequency, which is a generalization of the two-sample variance of the average frequency. The generalization includes the possibility of an overlap of the intervals (as well as the possibility of a "dead time" between the intervals), and incorporates the possibility of considering the frequency average over two time intervals of different duration. Just as for the two-sample variance of y, and for the autocorrelation function of $x(t)$, the covariance separates into the five terms of the noise model.

Analytic forms for the five terms of the autocorrelation function of $x(t)$ and for the five terms of the general cross-correlation of $\bar{y}$ are presented in as Equations 18 to 25 in the Appendix, derived with only the usual assumptions about high and low frequency limits to the noise bandwidth. Some comments are made on practical methods for computing values using these forms.

Algorithms for Accuracy Evaluation

For many, but not all, widely-used interpolation or extrapolation algorithms, it is possible to use the cross-correlation expressions and a knowledge of the noise to calculate the expected root-mean-square (i.e. $p = 2$) standard uncertainty [10] of the time or of the mean frequency extrapolated or interpolated by the algorithm. Based on the noise model (and a very large body of confirming experiment), the distributions expected for deviations from the fit are normal, and so the root-mean-square deviation calculated for the standard uncertainty is rigorously correct, and can be used in the conventional way for predicting confidence intervals from normal distributions [10]. The fitting process yields parameters for a parameterized functional form of $x_p(t)$, which may be of the general form (around the k^{th} interval of continuous operation) given in Equation 2. Nonlinear fitting parameters are not considered.

The algorithm should satisfy two criteria. *Firstly*, the algorithm should be unbiased by the deterministic part of the noise model: any change in a deterministic parameter (a^k, b^k or c^k in Equation 1) should be taken up by the algorithm and not bias the final result. Note that it is the final deviation which is to be unbiased, and some apparently biased forms for $x_p(t)$ may still be used in ways that are unbiased. In addition to being patently desirable, this condition also removes any need for considering cross-correlations between the deterministic and random noise parts of $x_m(t)$. *Secondly*, the coefficients d_l^k must depend linearly on sums over differences in $x(t)$. This includes fitting functions that are constrained to go through one point: least squares fits of polynomials of general order, and other functions with linear coefficients. It includes constrained weighted least-squares fits, as long as the weights do not themselves depend on the values of $x^k(t_i)$ or the variances on the k^{th} fitting interval. With this condition we ensure that we do not have to calculate any higher-order correlations than the general covariances evaluated in the Appendix.

What algorithms does this exclude? The first condition would not seem to exclude any serious contenders for fitting methods and fitting functions: if one is taking care to evaluate accuracy, then presumably one also values an unbiased fitting form. The second criterion admits many algorithms easily. However, to analyze rigorously the accuracy of a Kalman filter, where fitting weights depend on past variances, appears to require a study of higher-order autocorrelations of the random noise, at least to the level of identifying the magnitude of these corrections. We therefore exclude this important class of algorithms from our present considerations.

Many extrapolation and interpolation algorithms that are useful for time-scale purposes are very simple: for example, constrained to go through the last experimental point [8], or constrained to go through both the first and last points on an interval. However, it is instructive to consider first the most general least-squares fitting case for N points on a single calibrating interval.

Weighted Fits of General Functions with Linear Coefficients

The least-squares fit of the n parameters d_l of a function of the form of Equation 2, to $N + 1$ points $x(t_i)$, each point being taken with a weight w_i, is found by taking the partial derivative of the weighted L_2 norm with respect to its n coefficients. The resulting n simultaneous equations

in n unknowns are of the form $\mathbf{G}\vec{d}' = \vec{s}'$, where $\mathbf{G}$ is an $n \times n$ matrix with elements $G_{qr} = \sum_{i=0}^{N} w_i^2 g_q(t_i) g_r(t_i)$, and $\vec{s}'$ is an n-dimensional vector with elements $s'_r = \sum_{i=0}^{N} w_i^2 x_m(t_i) g_r(t_i)$. Here $x_m(t_i)$ is used to model $x(t_i)$, exactly as one might do in a numerical simulation. Simple inspection shows that if $g_1(t) = 1, g_2(t) = t$ and $g_3(t) = t^2$, and $s_r = \sum_{i=0}^{N} w_i^2 x_0(t_i) g_r(t_i)$, then $\mathbf{G}\vec{d} = \vec{s}$, where $\vec{d}$ is the vector of least-squares coefficients for the n functions, and $d_1 = d'_1 - a, d_2 = d'_2 - b$, $d_3 = d'_3 - c/2$ and $d_r = d'_r$ for $r \geq 4$. Thus for any set of weights, the fit exactly absorbs the deterministic part of the model function $x_m(t)$ and we have only to deal with the function $x_0(t)$. We can replace any difference between $x_m(t)$ and $x_p(t)$ (fit to the $x_m(t_i)$) with the difference between $x_0(t)$ and the $x_p(t)$ (fit to the $x_0(t_i)$). The fitting coefficients vector $\vec{d} = \mathbf{G}^{-1}\vec{s}$. The form of this equation is instantly quite revealing, for it shows that the least squares coefficients d_r only involve simple sums over the $x_0(t_i)$'s.

We will study all the effects of algorithm at t, reacting to the noise through the fitting procedure, by comparing $x_0(t)$ to the function fitted to the random part of the noise: $\vec{d} \cdot \vec{g}(t)$, where $\vec{g}(t)$ is the n-dimensional vector of the basis functions evaluated at t. This fitted function can in turn be transformed into a simple weighted sum over the $N+1$ of the $x_0(t_i)$'s: $\vec{d} \cdot \vec{g}(t) = \sum_{i=0}^{N} D_i(t) x_0(t_i)$, and $D_i(t) = w_i^2 \sum_{q=1}^{n} \sum_{r=1}^{n} (\mathbf{G}^{-1})_{qr} g_r(t_i) g_q(t)$.

The fitting algorithm and the noise model contribute a standard uncertainty [10] u_x in the least-squares fitted time which is $\langle [x_m(t) - x_p(t)]^2 \rangle$. This is the appropriate metric for judging the quality of the fit at t (relative to the set of $N+1$ specific fitting times $\{t_i\}$).

$$\begin{aligned} \left\langle [x_m(t) - x_p(t)]^2 \right\rangle &= \left\langle \left[x_m(t) - \vec{d}' \cdot \vec{g}(t) \right]^2 \right\rangle = \left\langle \left[x_0(t) - \vec{d} \cdot \vec{g}(t) \right]^2 \right\rangle \\ &= \left\langle \left[x_0(t) - \sum_{i=0}^{N} D_i x_0(t_i) \right]^2 \right\rangle = \sum_{i=-1}^{N} \sum_{j=-1}^{N} D_i D_j \langle x_0(t_i) x_0(t_j) \rangle , \end{aligned} \tag{5}$$

when t is labelled as t_{-1} and $D_{-1} = 1$ in the last equation. The sum over the $(N+2)^2$ autocorrelations simplifies since the autocorrelation depends on $\mid t_i - t_j \mid$. When the data are equally spaced, "only" $N+2$ different autocorrelations of $x_0(t)$ must be evaluated for a general value of t. The autocorrelations for our noise model can be evaluated using the expressions in the Appendix for $\mathcal{R}(\tau)$, if calculations can be done with sufficient numerical precision.

Constrained Least-Squares Fits

Fits constrained to go through one or more points can be considered as special cases in the above analysis through appropriate choices of weights. However, there is an interesting computational simplification which warrants explicit derivation: we can replace computations of the autocorrelation function of $x(t)$ with the simpler to compute S of Equation 13. Consider using a weighted least-squares fit constrained to go through one of the $x_m(t_i)$'s, for $i = i_c$ in general - although this might often be either end point: $i = 0$ or $i = N$ since we consider $\{t_i\}$ as being ordered so that $t_j < t_{j+1}$. For a constrained fit, $x_m(t_{i_c}) = \vec{d}'_c \cdot \vec{g}(t_{i_c})$, determined by substituting $\vec{g}(t_{i_c})$ for the first row of $\mathbf{G}$ to obtain $\mathbf{G}_c$, and by substituting $x_m(t_{i_c})$ for the first element of $\vec{s}'$ to obtain $\vec{s}'_c$. The weight w_{i_c} is assigned the value 0. The constrained fit is also unbiased by the deterministic part

of the noise model, satisfying the condition so that again we only have to consider the behaviour of $x_0(t)$ when evaluating the accuracy of the constrained fit. The fitted function is $\vec{d}_c \cdot \vec{g}(t)$, where $\vec{d}_c = \mathbf{G}_c^{-1}\vec{s}_c$ and $\vec{s}_c$ is just $\vec{s}_c'$ with $x_0(t)$ substituting for $x_m(t)$. For this constrained fit we can use the equality of x_m and x_p at t_{i_c}, or the equality $x_0(t_{i_c}) = 0$:

$$
\begin{aligned}
&\left\langle [x_m(t) - x_p(t)]^2 \right\rangle \;=\; \left\langle \left[\{x_m(t) - x_m(t_{i_c})\} - \left\{ \vec{d}_c \cdot (\vec{g}(t) - \vec{g}(t_{i_c}) \right\} \right]^2 \right\rangle = \\
&\left\langle \Big[\{x_0(t) - x_0(t_{i_c})\} \right. \\
&\left. - \left\{ \textstyle\sum_{q=1}^{n} \{g_q(t) - g_q(t_{i_c})\} \left\{ (\mathbf{G}_c^{-1})_{q1}\, x_0(t_{i_c}) + \sum_{r=2}^{n} (\mathbf{G}_c^{-1})_{qr} \sum_{i=0}^{N} w_i^2 g_r(t_i)\, x_0(t_i) \right\} \right\} \Big]^2 \right\rangle = \\
&\left\langle \left[\{x_0(t) - x_0(t_{i_c})\} - \left\{ \textstyle\sum_{i=0}^{N} (D_c)_i x_0(t_i) \right\} \right]^2 \right\rangle = \\
&\left\langle \left[\{x_0(t) - x_0(t_{i_c})\} - \left\{ \textstyle\sum_{i=0}^{N} (D_c)_i \left\{ \sum_{j=1}^{i} (x_0(t_j) - x_0(t_{j-1})) - \sum_{j=1}^{i_c} (x_0(t_j) - x_0(t_{j-1})) \right\} \right\} \right]^2 \right\rangle \\
&= \left\langle \left[\{x_0(t) - x_0(t_{i_c})\} - \left\{ \textstyle\sum_{j=1}^{N} (\hat{D}_c)_j \{x_0(t_j) - x_0(t_{j-1})\} \right\} \right]^2 \right\rangle \\
&= \left\langle \left[\left\{ \textstyle\sum_{j=0}^{N} (\hat{D}_c)_j (x_0(t_j) - x_0(t_{j-1})) \right\} \right]^2 \right\rangle ,
\end{aligned}
\tag{6}
$$

where $(\hat{D}_c)_j$ is just the reordered sum defined above for $j = 1$ to N, with $(\hat{D}_c)_0$ defined as 1.

Thus the constrained fit's standard uncertainty in time can be calculated more easily at a general time t, using only our expressions in the Appendix for $\mathcal{I}(\tau)$ rather than for $\mathcal{R}_x(\tau)$ to evaluate the standard uncertainty in time, which can done conveniently with 64-bit floating point calculations. Note that this is also the mean square of the time interval error over the time interval $[t_{i_c}, t]$.

Standard Uncertainty in Average Frequency

In a similar way we can calculate our model's estimate for the unconstrained least-squares fit's standard uncertainty [10] of the average frequency over an interval $[t, t+\tau]$, $u_y(t,\tau)$. We can calculate $u_y^2(t,\tau) = \left\langle [(x_m(t+\tau) - x_m(t)) - (x_p(t+\tau) - x_p(t))]^2 \right\rangle / \tau^2$ in terms of expressions using only the function $\mathcal{I}(T)$. For convenience, we will define $t_{-1} = t$ and $t_{N+1} = t+\tau$, neither restricting the value of t to be less than t_0 nor restricting the value of $t+\tau$ to be greater than t_N.

$$
\begin{aligned}
u_y^2(\tau)\tau^2 \;&=\; \left\langle [(x_m(t+\tau) - x_m(t)) - (x_p(t+\tau) - x_p(t))]^2 \right\rangle \;= \\
&\left\langle \left[\{x_0(t+\tau) - x_0(t)\} - \left\{ \vec{d} \cdot (\vec{g}(t+\tau) - \vec{g}(t)) \right\} \right]^2 \right\rangle \;= \\
&\left\langle \left[\textstyle\sum_{j=1}^{N+1} \{x_0(t_j) - x_0(t_{j-1})\} + \sum_{j=1}^{N} \check{D}_j \; \{x_0(t_j) - x_0(t_{j-1})\} \right]^2 \right\rangle \;= \\
&\left\langle \left[\textstyle\sum_{j=0}^{N+1} \tilde{D}_j \; \{x_0(t_j) - x_0(t_{j-1})\} \right]^2 \right\rangle
\end{aligned}
\tag{7}
$$

where $\check{D}_j = \sum_{i=j}^{N} w_i^2 \sum_{q=1}^{n} \sum_{r=1}^{n} (\mathbf{G}^{-1})_{qr} g_r(t_i)(g_q(t+\tau) - g_q(t))$ and $\tilde{D}_j = [\check{D}_j + 1]$ for $j = 0$ to N,

and where both $\tilde{D}_0$ and $\tilde{D}_{N+1}$ are equal 1. Again the uncertainty can be calculated conveniently with 64-bit floating point calculations from the expressions in the Appendix for $\mathcal{I}(\tau)$ rather than the more awkward $\mathcal{R}_x(\tau)$.

Multiple Calibration Runs

The time interval error that accumulates on an interval spanning m calibration runs is just the sum of time interval errors (the difference between the evolution of x and the evolution of the calibrated extrapolation x_p) on the interval $[t_{l^k}, t_{r^k}]$ around the k^{th} calibration run. The calibrated time scale is continuous, so that if the k^{th} fit yields to the next fit $(k+1)$ at a time $t_{r^k} = t_{l^{k+1}}$, then $\vec{d}^k \cdot \vec{g}^k(t_{r^k}) = \vec{d}^{k+1} \cdot \vec{g}^{k+1}(t_{l^{k+1}})$. Thus, across the m runs, the time interval error of the algorithm reacting to our noise model is

$$\sum_{k=1}^{m} E^k = \sum_{k=1}^{m} [x_0(t_{r^k}) - x_0(t_{l^k})] - \left[x_p^k(t_{r^k}) - x_p^k(t_{l^k})\right]. \tag{8}$$

To show the general form, consider the set of $\{t_\eta\}$, with the index η running in turn over the start time, the fitting times and the stop time for each interval, from 1 to M - the grand total of points (with N^k measurable subintervals and two extrapolated subintervals in the k^{th} interval). To analyze a fit on the k^{th} interval that extracts information from other intervals, we consider the k^{th} fit to be from t_{L^k} to t_{R^k}, with zero weight at the unmeasured times t_η where $\eta = l^j$ or r^j, $j = 1 \ldots m$. Equation 8 can then be combined over any overlap of the fits into a global weighted sum over the differences $[x_0(t_\eta) - x_0(t_{\eta-1})]$, weighted by $\mathcal{D}_\eta$ which sums over the fits which have used the η^{th} time interval. The mean square time uncertainty in the time scale algorithm over the total time interval $T = \sum_{\eta=1}^{M} [t_\eta - t_{\eta-1}]$ is $u_x^2(T)$, the mean square of the sum over E_i. Since for any algorithm of the class

$$\left\langle \left(\sum_{\eta=1}^{M} E_\eta \right)^2 \right\rangle = \left\langle \sum_{\eta=1}^{M} \sum_{\zeta=1}^{M} E_\eta E_\zeta \right\rangle = \sum_{\eta=1}^{M} \sum_{\zeta=1}^{M} \mathcal{D}_\eta \mathcal{D}_\zeta \left\langle [x_0(t_\eta) - x_0(t_{\eta-1})] [x_0(t_\zeta) - x_0(t_{\zeta-1})] \right\rangle. \tag{9}$$

The full evaluation of all terms of this $M \times M$ cross-correlation matrix, $\mathcal{E}$ (where $\mathcal{E}_{ij} = E_i E_j$) is possible for any particular set of calibration runs. With M^2 terms to evaluate, an efficient method for computing $\mathcal{S}$ is highly desirable. The method outlined in the Appendix will generally suffice. Fortunately, there can be very significant simplifications: $\mathcal{E}$ is symmetric, the main diagonal contributes most to the sum, the block-diagonal terms from the individual fits are the next most important parts (together these would contribute the quadrature sum of the standard time uncertainty u_x contributed across each individual interval, but neglecting interval-to-interval correlations), and generally the matrix elements far from the diagonal will not contribute significantly. Any regularity in the fitting times will also reduce the computational burden.

Examples

We will apply the above methods to one of the simplest algorithms that might be used: a linear fit constrained through two points, as shown in Figure 1. Here the mean square of the standard uncertainty in time is [11]

$$u_x^2 = 2\left[(1+\frac{\tau}{t_a})\mathcal{I}(\tau) + \frac{\tau}{t_a}(1+\frac{\tau}{t_a})\mathcal{I}(t_a) - \frac{\tau}{t_a}\mathcal{I}(t_a+\tau)\right] \tag{10}$$

This expression can be separated into the standard uncertainty u_x for each noise type, as illustrated in Figure 2. In this figure, u_x for each noise type has been normalized to equal 1 at the midpoint of the calibration interval t_a. The upper frequency cutoff ω is $100/t_a$, and the lower cutoff $\epsilon = 10^{-5}$ radians/second. To use a figure like this for a mixed noise model, recall that the standard uncertainties of the different noise types must be added in quadrature, with appropriate weights.

The above example shows how an rms residual at a time offset τ can be calculated, and that even in the simplest case it varies with position in different ways for the different types of noise. Using $\mathcal{I}(\tau)$ it is calculable in a perfectly straightforward manner.

Figure 3 shows the time interval error that develops from extrapolation to both earlier and later times than the calibration interval: from t_l earlier than the first point and to t_r later than the second fitted point. The standard uncertainty in frequency, $u_y(\tau)$, on this extended interval of τ is [14]

$$u_y^2 = \frac{2}{\tau^2}\left\{\mathcal{I}(\tau) + \left(\frac{\tau}{t_a}\right)^2\mathcal{I}(t_a) + \frac{\tau}{t_a}\mathcal{I}(t_r) - \frac{\tau}{t_a}\mathcal{I}(t_l+t_a) + \frac{\tau}{t_a}\mathcal{I}(t_l) - \frac{\tau}{t_a}\mathcal{I}(t_a+t_r)\right\} \tag{11}$$

As an interesting application of this simple algorithm, we can show how the local oscillator limit [12] might be circumvented for a pulsed cesium fountain. Consider the evaluation of different types of local oscillator for a pulsed cesium fountain that employs this algorithm and hyperfine phase differences [11] to span 0.01 s dead times between 0.99 s Ramsey times. The atomic polarization is measured for each pulse of atoms, and attributed (after calibration) to the discrepancy of the microwave phase compared to the primary hyperfine phase of the ensemble of cesium atoms. Thus the frequency of the local oscillator is known across the 0.99 s interval, with an uncertainty that may be limited by (among other things) atom shot noise in the ensemble. The average frequency from the preceeding and following Ramsey time is used to estimate the frequency of the local oscillator across the 0.01 s dead time between Ramsey pulses. If there is but a single ensemble "in flight" through the cesium fountain at any given time, then the simplest algorithm is the only one worth considering.

To analyze this system, we choose a symmetric interval centred on t_a, the active time, and $\mathbf{t}_l = \mathbf{t}_r = t_d/2$. Using the calibrated frequency from a single active interval for the two adjacent dead time half-intervals is equivalent to using the average frequency from two adjacent active intervals across any dead time. In this way we might hope to minimize the cross-correlations that need to be evaluated between neighbouring active times. We can use Equation 11 for the first estimate, but we should verify the size of the correction from neighbouring intervals. With a cycle time $t_a + t_d = T$,

and N equal-length cycles, Equation 9 becomes

$$\mathcal{E}_{j+l,j} = \mathcal{I}(lT+T) + \mathcal{I}(lT-T) - 2(1+[\tfrac{T}{t_a}]^2)\mathcal{I}(lT) + [\tfrac{T}{t_a}]^2 \{\mathcal{I}(lT+t_a) + \mathcal{I}(lT-t_a)\}$$
$$-2\tfrac{T}{t_a}\{\mathcal{I}(lT+\tfrac{1}{2}[T+t_a]) + \mathcal{I}(lT-\tfrac{1}{2}[T+t_a]) - \mathcal{I}(lT+\tfrac{1}{2}[T-t_a]) - \mathcal{I}(lT-\tfrac{1}{2}[T-t_a])\}. \quad (12)$$

Fortunately, for $l > 1$, the off-diagonal corrections are negligible. Depending on α, the adjoining term, $l = 1$, gives corrections to the diagonal terms of either sign and of up to 50% in magnitude. Thus correlations are significant in this case even though we tried apply our algorithm in a way to minimize the correlations. Figure 4 shows the standard uncertainty in frequency due to the local oscillator noise for three local oscillators. The curvature reflects some of the effects of correlations. The Allan deviation predicted [12] for the same three local oscillators in a conventional frequency servo is also shown (with a modulation frequency of 0.5 Hz). Clearly this hyperfine phase measurement plus post-processing algorithm is an interesting possibility as a replacement of the conventional servo, at least in some applications. It is crucial to have full confidence in the analysis of the complete effects of the interaction of the noise and the algorithm: our analysis is complete within the constraints of the noise model.

In the initial period of operation of a cesium fountain frequency standard, after having been evaluated as a standard, there might be a hour per day devoted to calibrating a time laboratory's secondary standards. In the case of NRC, this would mean the use of our new hydrogen masers [14], [15]. A crucial question in designing a cesium-fountain frequency-standard project is to choose the desired level of accuracy which might be used, either transferred to another laboratory or used to compare the frequency of two configurations of the cesium fountain which cannot coexist. The answer to this question can be obtained by our method for secondary frequency standards with well understood noise. As always, common mode noise between similar frequency standards is difficult to rule out - but the intent of these considerations is to establish a goal for an initial working standard that is not overbuilt, considering available frequency transfer characteristics. We present the analysis of a possible frequency-transfer budget at NRC.

The main question is the frequency transfer capabilities of our hydrogen masers. The Allan deviation of our two new masers has been measured with respect to each other. Attributing the noise equally to the two masers, we can calculate the best case for frequency transfer from one hour out to a day or so [14]. We again use the symmetric linear extrapolation from the end points. In this example, additional information will be available during the interval, but the "best case" for our noise types comes from using the end points [8]. By using the symmetric extrapolation, the frequency transfer will not be biased by any constant drift of the maser frequency [13].

Figure 5 shows an estimate for the standard uncertainty in frequency due to a hypothetical cesium fountain [11],[15], and the modelled Allan deviation one of the new NRC hydrogen masers [14]. The random-walk FM rise at large times is a somewhat pessimistic (or realistic ?) guess - the masers have not operated for a long enough time to properly evaluate their long-term stability $\sigma_y(\tau > 30$ days). Also shown is the calculated standard uncertainty of the average frequency of the fountain-plus-maser for the extrapolated frequency on an interval τ for a calibration run of 3000 s.

Using Equation 12, we can again evaluate cross-correlations for a regular series of runs (daily in this example). However, in this example there are no corrections larger than 1% to the $N^{-1/2}$ trend

line. At each end of τ we include a two-way time transfer uncertainty of 2 ns (1 - σ per transfer) as a realistic estimate of the state-of-the art (and including 0.2 ns as a worthwhile objective). The daily recalibration trend line gives the noise limit to the standard uncertainty for cesium fountain frequency transfer to another laboratory. It is encouraging in that the noise limit is well below 10^{-15} for an ambitious but realistic cesium fountain calibration schedule for the hydrogen masers. The estimated standard uncertainties for the as-transferred average frequency may facilitate the acceptance of the new standards, for example for the judging or steering of TAI. The predicted level of residual fluctuations will need to be experimentally verified to be fully confident that some significant common-mode noise source (random or deterministic) has not been missed, but even here having a baseline prediction will be very helpful in planning the evaluation level which is appropriate for any given frequency-transfer program.

Conclusions

We have developed, calculated and applied some useful metrics for judging the accuracy of algorithms extrapolating time and average frequency in the presence of noise that can be represented by a rather general noise model which includes all common types of power-law random noise as well as deterministic noise. For many algorithms the (rms) standard uncertainty in time, $u_x(t)$, and the (rms) standard uncertainty in average frequency across τ, $u_y(t,\tau)$, both can be calculated in terms of the noise model power law coefficients. This significantly enhances the attractiveness of standard uncertainties in time and frequency metrology where techniques for measuring the coefficients are widely used. The metrics can be used for guidance in the choice of algorithms and procedures (how often to recalibrate, and for how long), but a larger role can be played by these two "metrics", for rigorously judging the noise floor of different hardware and potential hardware used in novel ways.

The calculated standard uncertainty in time, $u_x(t)$, might also play a very useful role in the calculation of a reduced χ^2 for a set of experimental data for which one wishes to judge the adequacy of the fit and noise model. Conventionally, $\chi^2 = \frac{1}{N-n}\sum_{i=1}^{N}[x(t_i) - x_p(t_i)]^2/\{2[u_x(t_i)]^2\}$, where $N-n$ is used for the degrees of freedom: the number of (independent) data points minus the number of (independent) fitting parameters. Since we would be able to compute the cross-correlations between data points, it should also be possible to determine a better estimate of the degrees of freedom for the fit.

The procedures outlined here could be automated. Standard uncertainties could be used routinely in time and frequency metrology, for many common fits to a set of data points, in the presence of a general noise model. In all fields of metrology, where the use of standard uncertainties is now recommended [10], the very valuable techniques (such as the use of the modified Allan deviation) developed for analysing power law noise of frequency standards could be applied in other fields to give rigorous results for the standard way of reporting uncertainties.

Appendix: Autocorrelation Functions and General Interval Covariance for Power-law Noise Spectra

Consider two time intervals that may (or may not) overlap, and are not necessarily of equal duration. The general interval covariance of the random part of the time scale difference accumulated on the time interval $[t_1, t_2]$ with that accumulated on $[t_3, t_4]$ is

$$\begin{aligned} \mathcal{S} &= \langle [x_0(t_2) - x_0(t_1)]\,[x_0(t_4) - x_0(t_3)] \rangle &&= \langle x_0(t_2)x_0(t_4)\rangle + \langle x_0(t_1)x_0(t_3)\rangle \\ &&&- \langle x_0(t_2)x_0(t_3)\rangle - \langle x_0(t_1)x_0(t_4)\rangle\,. \end{aligned} \tag{13}$$

where $\langle x_0(t)x_0(t-\tau)\rangle$ is the autocorrelation function: $\mathcal{R}_x(\tau) = \int_0^\infty S_x(f)\cos 2\pi f\tau\, df$. We use the usual [5] upper-frequency limit $f_u = \omega/(2\pi)$ and low-frequency limit $f_l = \epsilon/(2\pi)$. The sharp upper cutoff is an artifice, although one which could be implemented with a digital filter applied to the output of a phase comparator. The lower frequency needs to be low enough so that it does not perturb the low-frequency rolloff supplied by the fitting function. A too-low value for ϵ will exacerbate the numerical precision problems in calculating and using $\mathcal{R}_x(t)$, which diverges as $\epsilon \to 0$. For the usual noise model of $S_x(f) = \sum_{\alpha=-2}^{2} \frac{h_\alpha f^{(\alpha-2)}}{(2\pi)^2}$, we have

$$\mathcal{R}_x(\tau) = \sum_{\alpha=-2}^{2} \frac{h_\alpha \tau^{(1-\alpha)}}{(2\pi)^{\alpha+1}} \int_{\epsilon\tau}^{\omega\tau} u^{\alpha-2}\cos u\, du = \sum_{\alpha=-2}^{2} R_\alpha(\tau). \tag{14}$$

It is possible to express $\mathcal{S}$ as a sum over the functions $R_\alpha(\tau)$'s, or as a sum over functions of similar form $I_\alpha(\tau)$'s that are less divergent as $\epsilon \to 0$:

$$\begin{aligned} \mathcal{S} &= -[\mathcal{R}_x(t_4 - t_1) + \mathcal{R}_x(t_3 - t_2) - \mathcal{R}_x(t_4 - t_2) - \mathcal{R}_x(t_3 - t_1)] \\ &= \mathcal{I}(t_4 - t_1) + \mathcal{I}(t_3 - t_2) - \mathcal{I}(t_4 - t_2) - \mathcal{I}(t_3 - t_1), \end{aligned} \tag{15}$$

where $\mathcal{I}(\tau)$ is

$$\mathcal{I}(\tau) = \sum_{\alpha=-2}^{2} \frac{h_\alpha \tau^{(1-\alpha)}}{(2\pi)^{\alpha+1}} \int_{\epsilon\tau}^{\omega\tau} u^{\alpha-2}\{1 - \cos u\}\, du = \sum_{\alpha=-2}^{2} I_\alpha(\tau). \tag{16}$$

With the help of mathematical tables or a symbolic algebra program such as Maple V, the integrals for $R_\alpha(\tau)$ and $I_\alpha(\tau)$ can be done analytically for our values of α.

$$R_2(\tau) = \frac{h_2}{(2\pi)^3}\omega\left\{\frac{\sin(\omega\tau)}{\omega\tau}\right\} \tag{17}$$

$$I_2(\tau) = \frac{h_2}{(2\pi)^3}\omega\left\{1 - \frac{\sin(\omega\tau)}{\omega\tau}\right\} \tag{18}$$

are the integrals for white phase noise ($\alpha = 2$). I_2 is proportional to the high frequency cutoff ω, as is the Allan deviation.

$$R_1(\tau) = \frac{h_1}{(2\pi)^2}\left\{Ci(\omega\tau) - Ci(\epsilon\tau)\right\} \tag{19}$$

$$I_1(\tau) = \frac{h_1}{(2\pi)^2}\left\{\gamma + \ln(\omega\tau) - Ci(\omega\tau)\right\} \tag{20}$$

are the integrals for flicker phase noise ($\alpha = 1$). $Ci(x)$ is the Cosine Integral function, which can be easily approximated by using numerical approximation as in Abramovitz [16] for arguments greater than 1, or by direct numerical integration of the expression between brackets (see end of this Appendix). The time interval τ tempers the logarithmic dependence on the frequency cutoff.

$$R_0(\tau) = -\frac{h_0}{2\pi}\tau\left\{\frac{\cos(\omega\tau)}{\omega\tau} + Si(\omega\tau) - \frac{\cos(\epsilon\tau)}{\epsilon\tau} - Si(\epsilon\tau)\right\} \tag{21}$$

$$I_0(\tau) = \frac{h_0}{2\pi}\tau\left\{\frac{\cos(\omega\tau) - 1}{\omega\tau} + Si(\omega\tau)\right\} \tag{22}$$

are the integrals for white frequency noise ($\alpha = 0$). $Si(x)$ is the Sine Integral function, which can be treated in the same way as the Cosine Integral for arguments greater than 1, or by direct numerical integration of the definition of $Si(x)$ for small arguments. This term depends linearly only on the time interval τ, and not on value of the high frequency cutoff.

$$R_{-1}(\tau) = h_{-1}\frac{\tau^2}{2}\left[\left\{\frac{-\cos(\omega\tau)}{(\omega\tau)^2} + \frac{\sin(\omega\tau)}{\omega\tau} - Ci(\omega\tau)\right\} - \left\{\frac{-\cos(\epsilon\tau)}{(\epsilon\tau)^2} + \frac{\sin(\epsilon\tau)}{\epsilon\tau} - Ci(\epsilon\tau)\right\}\right] \tag{23}$$

$$I_{-1}(\tau) = h_{-1}\frac{\tau^2}{2}\left[1.5 - \gamma - \ln(\epsilon\tau) - \left\{\frac{1-\cos(\omega\tau)}{(\omega\tau)^2} + \frac{\sin(\omega\tau)}{\omega\tau} - Ci(\omega\tau)\right\}\right] \tag{24}$$

are the integrals for flicker frequency noise ($\alpha = -1$), where γ is 0.57721..., Euler's constant. The logarithmic term looks as if it will diverge to infinity for a low frequency cutoff as $\epsilon \to 0$, but in combinations like in $\mathcal{S}$ normally the combination of terms using $I_{-1}(\tau)$ is such that the terms multiplying $1.5 - \gamma$ will cancel out, as will the logarithmic divergence with ϵ, leaving a term depending only on the square of the time interval and a geometric structure factor that depends on the logarithm of ratios of the time intervals concerned: for the Allan deviation this is ln(2) as expected, and for $\mathcal{S}$ it is $\ln\left(\frac{(t_4-t_1)(t_3-t_2)}{(t_3-t_1)(t_4-t_2)}\right)$.

$$\begin{aligned} R_{-2}(\tau) &= \pi h_{-2}\frac{\tau^3}{3}\left[\left\{\frac{\sin(\omega\tau)}{(\omega\tau)^2} + \frac{\cos(\omega\tau)}{\omega\tau} + Si(\omega\tau) - 2\frac{\cos(\omega\tau)}{(\omega\tau)^3}\right\}\right. \\ &- \left.\left\{\frac{\sin(\epsilon\tau)}{(\epsilon\tau)^2} + \frac{\cos(\epsilon\tau)}{\epsilon\tau} + Si(\epsilon\tau) - 2\frac{\cos(\epsilon\tau)}{(\epsilon\tau)^3}\right\}\right] \end{aligned} \tag{25}$$

$$I_{-2}(\tau) = \pi h_{-2}\frac{\tau^3}{3}\left[\frac{3}{\epsilon\tau} - \left\{2\frac{1-\cos(\omega\tau)}{(\omega\tau)^3} + \frac{\sin(\omega\tau)}{(\omega\tau)^2} + \frac{\cos(\omega\tau)}{\omega\tau} + Si(\omega\tau)\right\}\right] \tag{26}$$

are the integrals for random walk frequency noise ($\alpha = -2$) for a low frequency cutoff $\epsilon \rightarrow 0$. As expected, this expression will diverge to infinity in the limit $\epsilon \rightarrow 0$, but the combination of terms will cancel that divergence in many of our accuracy prediction problems, since our fitting procedures will act as an effective low-frequency rolloff for the residuals of the fit. In practice, the low frequency cutoff will depend on the observation (active) time and is always some value greater than zero. In our analysis of uncertainty, we should use the same value of ϵ that was used in determining h_{-2} from measured Allan variances of the standard in question.

Using standard expansions for the Sine and Cosine Integrals [16], it is easy to compute with enough accuracy the values needed for arguments greater than 1. For arguments smaller than one, numerical integration can be done easily for $Si(x)$. For $Ci(x)$, the following transformation is helpful, since the second integral is easy to do numerically for small arguments:

$$Ci(x) = \gamma + \ln x - \int_0^x \frac{1-\cos u}{u} du = \gamma + \ln x - \int_0^x \frac{\sin^2 \frac{u}{2}}{u} du. \tag{27}$$

In debugging any code written to evaluate $\mathcal{R}_x(\tau)$ or $\mathcal{I}(\tau)$, it is worthwhile noting that the two-sample variance or Allan variance is

$$\sigma_y^2(\tau) = \frac{-4\mathcal{R}_x(\tau)+3\mathcal{R}_x(0)+\mathcal{R}_x(2\tau)}{\tau^2} = \frac{4\mathcal{I}(\tau)-\mathcal{I}(2\tau)}{\tau^2}, \text{ where} \tag{28}$$

$$\mathcal{R}_x(0) = \frac{h_2}{(2\pi)^3}(\omega - \epsilon) + \frac{h_1}{(2\pi)^2}\ln\frac{\omega}{\epsilon} + \frac{h_0}{2\pi}(\epsilon^{-1} - \omega^{-1}) + \frac{h_{-1}}{2}(\epsilon^{-2} - \omega^{-2}) + \frac{h_{-2}2\pi}{3}(\epsilon^{-3} - \omega^{-3}),$$

which should reflect the traditional expressions [5] when the typographical errors have been corrected (their equations 101 through 105 have had the term $\gamma + \ln(2\pi f_h \tau)$, where γ is Euler's constant, incorrectly typeset as $2 + \ln(2\pi f_h \tau)$). Some approximations used to obtain their traditional expressions have not been made. The expressions given above for $I_\alpha(\tau)$ are correct for all α even for $\omega_c \tau < 1$, so that differences are to be expected if calculations are made for τ near to or less than $1/(2\pi f_c)$, where f_c is the upper frequency cutoff or bandwidth of the phase measurement system used to measure $x(t)$: the expressions above give the correct results.

References

[1] A. Clairon, C. Salomon, S. Guellati and W.D. Phillips: Europhys. Lett. **16** (1991) 165.

[2] K. Gibble and S. Chu: Phys. Rev. Lett. **70** (1993) 1771.

[3] A.A. Madej, K.J. Siemsen, J.D. Sankey, R.F. Clark and J. Vanier: IEEE Trans. Instrum. Meas. **IM-42**, (1993) pp 234-241.

[4] A.G. Mungall: IEEE Trans. Instrum. Meas. **IM-27**, (1978) pp 330-334.

[5] J.A. Barnes, A.R. Chi, L.S. Cutler, D.J. Healey, D.B. Leeson, T.E. McGunigal, J.A. Mullen Jr., W.L. Smith, R.L. Sydnor, R.C. Vessot and G.M.R. Winkler: IEEE Trans. Instrum. Meas. **IM-20** (1971) 105.

[6] D.W. Allan and J.A. Barnes: Proc. 35^{th} Ann. Freq. Control Symposium, (1981), pp 470-474.

[7] F. Vernotte, E. Lantz, J. Groslambert and J.J. Gagnepain: IEEE Trans. Instrum. Meas. **IM-42** (1993) pp 342-350.

[8] D.W. Allan: IEEE Trans. Ultrason., Ferroelec., Freq. Contr. **UFFC-34**, (1987) 647.

[9] N.J. Kasdin and T. Walter: Proc. 1992 IEEE Frequency Control Symposium (1992) pp 274-283.

[10] Guide to the Expression of Uncertainty in Measurement, First Edition, International Organization for Standardization, Technical Advisory Group on Metrology, Working Group 3 (1992).

[11] R.J. Douglas and J.-S. Boulanger: Proc. 1992 IEEE Frequency Control Symposium (1992) pp 6-26.

[12] C. Audoin, V. Candelier and N. Dimarq: IEEE Trans. Instrum. Meas. **IM-40**, (1991) 121-125.

[13] D. Morris: Proceedings of the Twenty-second Annual Precise Time and Precise Time Interval Planning Meeting (PTTI), (1990) 349.

[14] D. Morris, R.J. Douglas and J.-S. Boulanger: 1993 AFS-CQE Symposium, Nara, Japan; Submitted to Japanese Journal of Applied Physics.

[15] J.-S. Boulanger, D. Morris, R.J. Douglas and M.C. Gagné: Proceedings of the Twenty-fifth Annual Precise Time and Precise Time Interval Planning Meeting (PTTI), (1993).

[16] M. Abramovitz and I. Stegun: Handbook of Mathematical Functions, 9^{th} printing by Dover publications, New-York, pages 232 and 233: formulas 5.2.8, 5.2.9, 5.2.38 and 5.2.39.

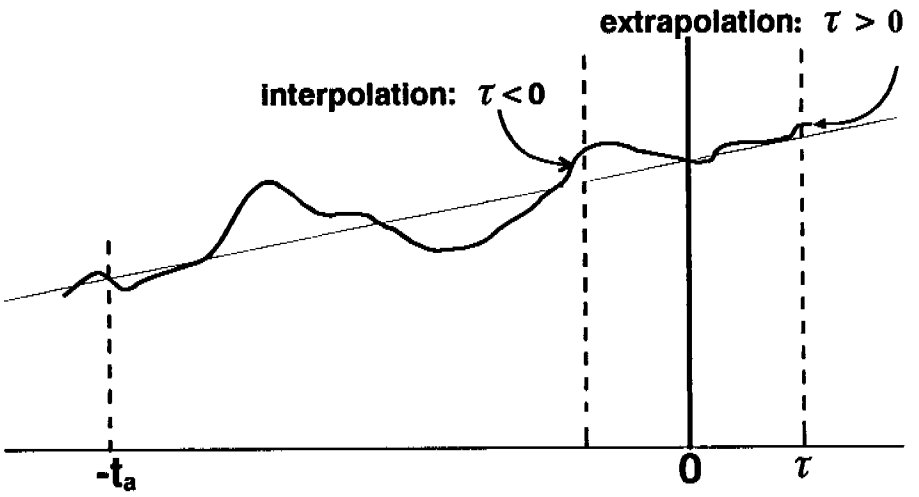

Figure 1. A simple algorithm: linear extrapolation or interpolation from a line constrained to lie through two points on x(t), separated by a time t_a. Its standard uncertainty in time, $u_x(\tau)$, can vary with extrapolation time τ, as shown in Figure 2.

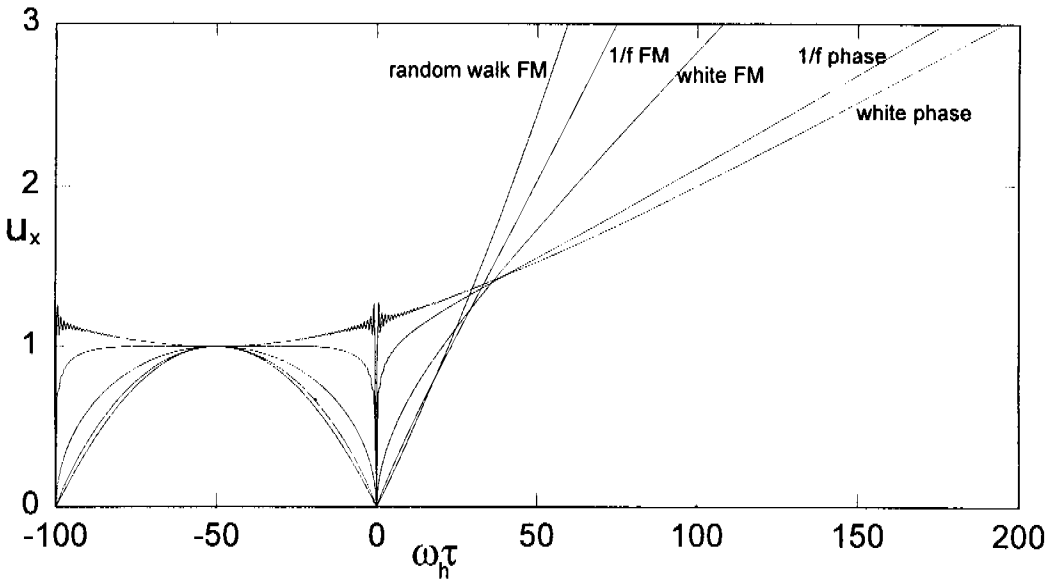

Figure 2. The standard uncertainty in time, $u_x(\tau)$, developed by the algorithm of Figure 1 for an extrapolation time τ. It is shown separately for the five types of noise, each normalized to 1 at the midpoint of the fitting interval t_a. In this example, $\omega_h = 100/t_a$.

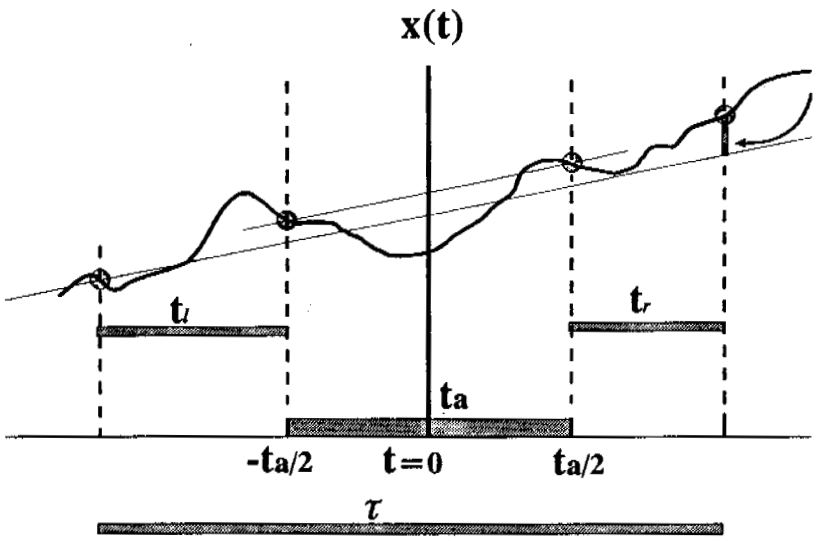

Figure 3. A simple algorithm: linear fit constrained to lie through two points on x(t). The variation in frequency from the calibration interval t_a is illustrated. With $t_l = t_r$, this is the basis for Figures 4 and 5.

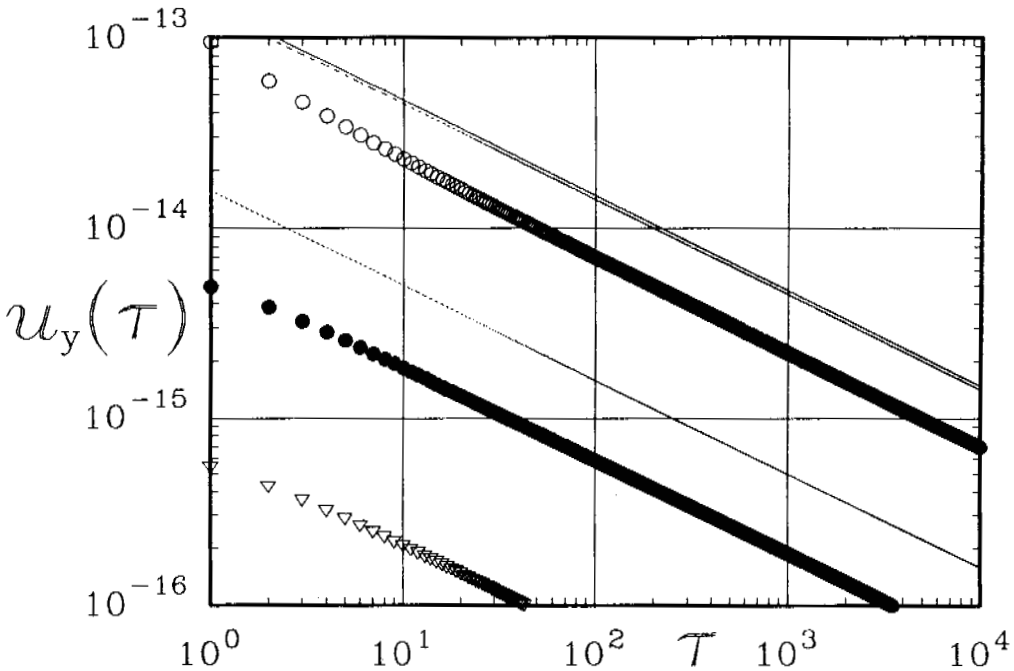

Figure 4. Local oscillator contribution to the standard uncertainty of the average frequency for a pulsed-ensemble cesium fountain, with a cycle time of 1 s and a dead time of 0.01 s. The light lines show the classical stability limit of the three oscillators, and the heavy symbols show the pulsed-ensemble result using linear extrapolation in phase to bridge the dead time. For each type of servo, the top curve is for an Oscilloquartz 8600-3 ($h_{-1} = 8\times10^{-26}$, $h_1 = 8\times10^{-27}$, $h_2 = 5.6\times10^{-29}$), the middle curve for a Wenzel 500-03475 100MHz & 5 MHz ($h_{-1} = 8\times10^{-26}$, $h_1 = 1.6\times10^{-30}$, $h_2 = 1.3\times10^{-34}$), and the bottom curve is for a JPL-type 77K sapphire X-band frequency discriminator ($h_{-1} = 1\times10^{-27}$). Optimally used, their noise corresponds to a shot noise of 700, 8×10^4 and 6×10^6 atoms/s respectively.

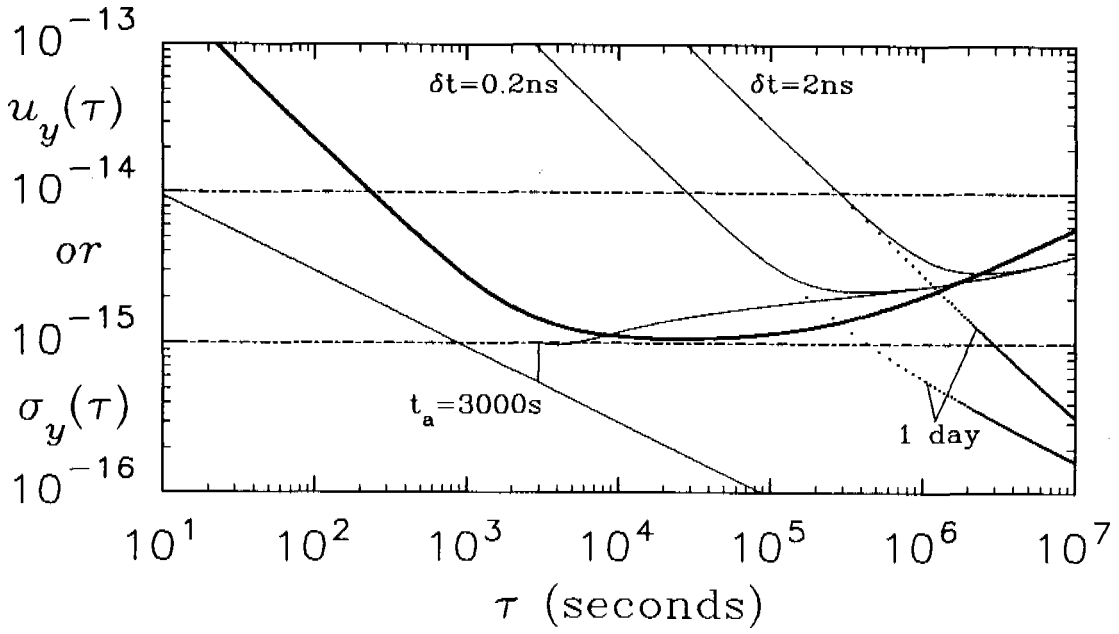

Figure 5. Cesium fountain's standard uncertainty of average frequency over an interval τ,due to random noise. The maser noise model 's Allan deviation is also shown as the heavy curve. A cesium fountain with a $u_y(t) = 10^{-14}\tau^{-1/2}$ is operated for 3000 s , calibrates a maser, the standard uncertainty of the extrapolation from one calibration is indicated by the light curve that rises abruptly at t = 3000 s . The other curves show what can be done with current (2 ns) and a possible future two-way time transfer at the ends of the interval τ. The dots show what can be done if a 3000 s calibration run is performed every day.