Characterization, Optimum Estimation, and Time Prediction
of Precision Clocks

by

David W. Allan
National Bureau of Standards
Boulder, Colorado 80303

"What then," asked St. Augustine, "is time? If no one asks me, I know what it is. If I wish to explain it to him who asks me, I do not know." We have learned a few things since St. Augustine, Einstein has taught us a lot. But still there are a lot of unanswered questions. In particular, how do you measure time? It intrigues me that we never measure time; we measure time differences, i.e. the time difference between two clocks. I know of no way to measure the time of a clock. I can measure the time of an event with reference to a particular clock. Another intriguing question is, if time cannot be measured, is it physical or is it an artifact? We conceptualize some of the laws of physics with time as the independent variable. We attempt to approximate our conceptualized ideal time by inverting these laws so that time is the dependent variable. The fact is that time, as we now generate it, is dependent upon defined origins, a defined resonance in the cesium atom, interrogating electronics, induced biases, and random perturbations from the ideal. Hence, at a significant level, time -- as man generates it by the best means available to him -- is an artifact. Corollaries to this are that every clock disagrees with every other clock essentially always, and no clock keeps ideal or "true" time except as we may choose to define it. Frequency or time interval, on the other hand, is fundamental to nature; hence, the definition of the second can approach the ideal. Noise in nature is also fundamental. Characterizing the random variations of a clock opens the door to optimum estimation of environmental influences and to the design of optimum combining algorithms.

Let's define some terms[1]: Beginning with a sine wave voltage with frequency $\nu(t)$ as the variable frequency output of a precision oscillator, we may write $V(t) = V_o \sin(2\pi \overline{\nu}(t) \cdot t)$, where we assume that amplitude fluctuations are negligible around V_o, and where $\overline{\nu}(t)$ is the average frequency from 0 to t. We can redefine this equation with ν_o being a constant nominal frequency and place all of the deviations in the phase, $V(t) = V_o \sin (2\pi\nu_o \cdot t + \phi(t))$. We then define a quantity $\overline{y}(t) = (\overline{\nu}(t) - \nu_o)/\nu_o$, which is dimensionless and which is the fractional frequency deviation of $\overline{\nu}(t)$ from the nominal value. We can integrate $\overline{y}(t)$ to get the time deviation, x(t), which has the dimensions of time. From the above we can write the time deviation of a clock as a function of the phase deviation: $x(t) = \phi(t)/2\pi\nu_o$.

Why do we have time deviations? We conceptualize two categories: systematics, such as frequency drift (D), frequency offset (y_o) and time offset (x_o); and then random deviations $\varepsilon(t)$, which are not considered to be deterministic:

$$x(t) = x_o + y_o \cdot t + 1/2 \ Dt^2 + \varepsilon(t). \qquad (1)$$

Note, the quadratic in the D term is because x(t) is the integral of $\overline{y}(t)$, the fractional frequency. In Figure 1 we have simulated two systematic cases: one a clock with frequency offset, and another case with a negative frequency drift. systematics caused by environmental influences are also very important. Figures 2 through 6 summarize some of the important systematic influences on precision oscillators. An important set of systematic deviations are modulation side bands, e.g. 60 Hz, 120 Hz, daily and annual dependences, those induced by vibrations, etc.

The random deviations of precision oscillators can typically be categorized by power law spectra, $S_y(f) \sim f^\alpha$ where f is the Fourier frequency and α takes on integer values, i.e. -2, -1, 0, 1, 2[1,2,3,4]. Figure 7 shows noise samples corresponding to these different power law spectra and Table 1 shows the nominal range of applicability of these power law models. Given a time deviation plot x(t) for the time difference between a pair of clocks or a clock against some primary reference, and some

sample time τ (see Figure 8 for example), the average fractional frequency for each interval is the time difference at the end of the interval minus that at the beginning of the interval divided by τ. We can thus construct a set of discrete frequency values over such a data set from this time deviation plot. We can calculate a classical standard deviation for these values, but one can show that for some kinds of power law spectra encountered in precision oscillators that the standard deviation is divergent[5], i.e., it does not converge to a well defined value, and it is a function of data length[2]. Hence, the standard deviation should not be used for characterizing clocks. An IEEE subcommittee has recommended $S_y(f)$ in the frequency domain and a measure, $\sigma_y^2(\tau)$ in the time domain[1]. The latter has come to be called the two-sample variance or the Allan variance. The convergence of $\sigma_y(\tau)$ has been verified[1,2,3,4] for the power law spectra of interest in precision oscillators. It is defined as follows[1]:

$$\sigma_y^2(\tau) = \frac{1}{2} \langle (\Delta y)^2 \rangle \tag{2}$$

where Δy is the difference between adjacent fractional frequency measurements each sampled over an interval τ and the brackets $\langle \rangle$ indicate the infinite time average or expectation value.

A pictorial description is shown in Figure 9 for a finite data set. A data set of the order of 100 points is quite adequate for convergence of $\sigma_y(\tau)$, though of course the confidence of the estimate will typically improve as the data length increases[6]. Given a discrete set of stored data, the value of τ can be varied in the software[7]. If τ_0 is the data spacing for the stored data set, $\bar{y}_i$, from the measurement system, one can change to $\tau = n\tau_0$ by averaging n adjacent values of $\bar{y}_i$ to obtain a new fractional frequency estimate with sample time τ as input to Equation (2). Hence, in a very convenient way one can calculate $\sigma_y(\tau)$ as a function of τ, which will be shown to be very useful in a moment. For a finite data set, Equation (2) then becomes

$$\sigma^2(\tau) \cong \frac{1}{2(M+1-2n)} \sum_{k=1}^{M+1-2n} \left(\bar{y}_{k+n}^{\tau} - \bar{y}_k^{\tau} \right)^2, \quad \text{and} \tag{3}$$

$$\sigma_y^2(\tau) \cong \frac{1}{2\tau^2(M+1-2n)} \sum_{k=1}^{M+1-2n} (x_{k+2n} - 2x_{k+n} + x_k)^2 \qquad (4)$$

where the y_k^τ are average fractional frequencies averaged over $\tau = n\tau_o$, starting at k, and the x_ks are the discrete time deviation measurements,

$$x_k = k\tau_o \sum_{i=1}^{k} \bar{y_i}.$$

Equation (3) is obtained from a first difference on frequency, and equation (4), from the second difference on the time; they are mathematically identical, which gives us the option of using frequency or time data.

For power law spectra, μ is constant for a particular value of α, where $\sigma_y^2(\tau) \sim \tau^\mu$. The relationship of the spectral density, $S_y(f) \sim f^\alpha$, exponent α versus μ is $\mu = -\alpha-1$ ($-2< \alpha \leq 1$) and $\mu = -2$ ($\alpha \leq 1$).

For example, $\sigma_y(\tau)$ is proportional to $\tau^{-1/2}$ (which is typical) for cesium, rubidium and passive hydrogen; then μ has the value of -1, and hence α has the value of 0 (white noise frequency modulation). This is the classical noise exhibited by most atomic clocks for τ's beyond a few seconds and in this case $\sigma_y(\tau_o)$ is equal to the classical standard deviation. Fortunately for most cases where $\tau \leq 1$ second the relationship $\mu = -\alpha-1$ is applicable.

We have an ambiguity at $\mu = -2$; we cannot tell whether we have flicker noise phase modulation (PM) or white noise PM. We can avoid this problem by realizing that in this region $\sigma_y(\tau)$ depends on the measurement bandwidth[2,3]. One can construct a variable software bandwidth, f_s, by realizing the following[8,9]. In any measurement system a hardware bandwidth, f_h, exits through which we measure the phase or the time difference between a pair of clocks; define $\tau_h = 1/2\pi f_h$. In other words τ_h is the sample time period through which we sample the data. If we average n time or phase readings, we increase the time period to $n\tau_h = \tau_s$. But $\tau_s = 1/2\pi f_s$, where $f_s = f_h/n$, i.e., we narrow the effective bandwidth,

f_s, in the software by n. In other words $f_s = f_h/n$ gets smaller as we average more values; i.e. increase n ($\tau = n\tau_o$). We can therefore construct a second difference composed of time deviations so averaged, and then define a modified $\sigma_y^2(\tau)$ that will remove the ambiguity through bandwidth variation:

$$\text{Mod.}\sigma_y^2(\tau) \approx \frac{1}{2\tau^2 n^2 (N-3n+1)} \sum_{j=1}^{N-3n+1} \left(\sum_{i=j}^{n+j} (x_{i+2n} -2x_{i+n} + x_i) \right)^2 \quad (5)$$

where N = M+1, the number of time deviation measurements available from the data set. And now if $\text{Mod.}\sigma_y^2(\tau) \sim \tau^\mu$, then $\mu' = -\alpha-1$ ($1\leq \alpha \leq3$)[9,10].

We typically employ $\text{Mod.}\sigma_y(\tau)$ as a subroutine to remove the ambiguity if $\sigma_y(\tau) \sim \tau^{-1}$ because the μ' dependence only approximates that given by equations (3 and 4) for $\alpha <1$. But for $\alpha = 2$ and 1, $\mu /2$ exactly equals -3/2 and -1 respectively, providing a clean differentiation between white noise PM and flicker noise PM.

Table 2 illustrates why one should not use the classical standard deviation to characterize clocks. For the different kinds of noise processes we list the classical standard deviation of the time deviations and the classical standard deviation of the fractional frequency deviations as a function of $\sigma_y(\tau)$ (the square root of the Allan variance). The divergent nature of either classical standard deviation is apparent, and even for classical white noise FM the standard deviation is apparent, and even for classical white noise FM the standard deviation of the time diverges as the square root of the data length i.e. the number of samples N[2].

Using $\sigma_y(\tau)$ or $\text{Mod.}\sigma_y(\tau)$ we can characterize typical power law processes. We then have the opportunity of determining optimum estimates of values by employing the statistical theorem that the optimum estimate of a white noise process is just the simple mean.

For example, consider the very common and very important case of white noise FM typically found on the signals from cesium standards, rubidium standards and passive hydrogen masers. The optimum estimate of the frequency is the simple mean frequency, which is equivalent to $(x_N - x_1)/M\tau_o$. It is still all too common within our discipline to see our colleagues erroneously determining the frequency for these kinds of oscillators by calculating the slope from a linear leasts squares fit to the time deviations and quoting the standard deviation around that fit as a measure of the clock performance. There are three problems in proceeding this way. First, the frequency estimate is not optimum in a mean square error sense and is equivalent to throwing away about 20% of the data, increasing the cost in the case of a calibration. Second, the standard deviation diverges as the square root of the data length and third, the standard is significantly dependent on the filter (linear leasts squares) as well as the clock deviations. On the other hand such a filter can be useful for assessing outliers. The optimum "end point" method outlined above has the risk that if either of the points is abnormal, i.e. the model fails, the result will, of course, be adversely effected, so such a filter is useful to assess whether there are, outliers -- paying especial attention to the end points.

There are other useful, and maybe not so obvious, optimum estimators at the conclusion of a data set: (1) Given white noise PM, the best time error estimate is the simple mean of the time deviations, the frequency estimate, then, is the slope from a linear least squares fit to the time deviations, and the frequency drift, D, is determined from a quadratic least squares fit to the time deviations per equation (1). (2) Given white noise FM, the optimum estimate of the time is the last value, the optimum frequency estimate is outlined in the previous paragraph and the optimum frequency drift estimate is from a linear least squares fit to the frequency. (3) Given random walk FM, the optimum time estimate is the last value, and optimum frequency estimate is obtained from the last slope of the time deviations, and the optimum frequency estimate is from the last slope from the time deviations, and the optimum frequency drift estimate is calculated from the mean second difference of the time deviations. Caution needs to be exercised here for typically there will

be higher frequency component noise in a real data stream, such as white noise FM, along with the random walk, and these can significantly contaminate the drift estimate from a mean second difference. If random walk FM is the predominant long-term, power-law process, which is often the case, then the noise can be handled by calculating the second difference from the first, middle and end time deviation points of the data. The flicker noise cases are significantly more complicated, though filters can be designed to approximate optimum estimation [11,12,13]. In the limit as the data length increases without limit, the time is not defined for flicker noise PM, an the frequency is not defined for flicker noise FM. This has some philosophical implications for the definitions of time and frequency unless some low frequency cutoff limits exist. If significant frequency drift exists in the data, it should be optimally subtracted from the data or it will bias the long term values of $\sigma_y(\tau)$:

$$\sigma_y(\tau) = \frac{D\tau}{\sqrt{2}} \ . \tag{6}$$

Once the power law spectra are deduced for a pair of oscillators, then one can also develop an optimum predictor. Table 3 gives both the optimum prediction values for the various relevant pure power law spectra, as well as, their assymtotic forms. Special forecasting techniques must be used for optimal prediction when combinations of these processes are present. To illustrate how these concepts relate to real devices, Figure 10 shows a $\sigma_y(\tau)$ diagram for some interesting state-of-the-art oscillators, and Figure 11 shows the rms time prediction errors for the same set of oscillators.

In conclusion it is clear that classical statistics does not allow characterization of common kinds of random signal variations found in precision oscillators. The two-sample or Allan variance provides an efficient and convergent measure of the power law spectral density models useful in characterizing most of these oscillators. Once characterized we can calculate optimum time and frequency estimates as well as predicted values. Characterizing the random variations also provides near optimum estimation of systematic effects, which often cause the predominant time and frequency deviations. For example, if we wanted to optimally determine the temperature dependence with the temperature set at two

51

different values, we would set at one temperature and measure the frequency against a reference for a time τ_m, corresponding to the τ for the minimum $\sigma_y(\tau)$ value. We would then change the temperature to the other value and repeat the measurement with the same criteria and note the resultant Δy between the two optimally determined frequency values. If these two steps are repeated several times, an arbitrarily good precision is achieved and is approximately given by $\sigma_y(\tau_m)/\sqrt{P}$, where P is the number of Δy values obtained from switching back and forth. Knowing the characteristics of both the random and the systematic deviations of precision oscillators clearly is useful to the designer, the planner, the user, as well as the vendor.

References

[1] Barnes, J. A., Chi, A. R., Cutler, L. S., Healey, D. J., Lesson, D. B., McGunigal, T. E., Mullen, Jr., J. A., Smith, W. L., Sydnor, R. L., Vessot, R. F. C., Winkler, G. M., "Characterization of Frequency Stability," Proc. IEEE Tran. on Instrumentation and Measurement, IM-20, 105, NBS Technical Note 394, 1971.

[2] Allan, D. w., "Statistics of Atomic Frequency Standards," Proc. IEEE 54, p. 221, 1966.

[3] Vessot, R., Mueller, L., Vanier, J., "The Specification of Oscillator Characteristics from Measurements Made in the Frequency Domain," Proc. of the IEEE, Special Issue on Frequency Stability, Vol. 54, No. 2, pp. 199-207, February 1966.

[4] Lesage, P., Audoin, C., "Characterization and measurement of time and frequency stability," Radio Science, Vol. 14, pp. 521-539, July-August 1979.

[5] Barnes, J. A., "Atomic Timekeeping and the Statistics of Precision Signal Generators," Proc. of the IEEE, Special Issue on Frequency Stability, Vol. 54, No. 2, pp. 207-220, February 1966.

[6] Howe, D. A., Allan, D. W., Barnes, J. A., "Properties of Signal Sources and Measurement Methods," Proc. 35th Annual Symposium on Frequency Control (SFC), May 1986, pp. 669, A1-A47.

[7] Allan, D. W., "The Measurement of Frequency and Frequency Stability of Precision Oscillator," NBS Technical Note 669, May 1975.

[8] Snyder, J. J., "An Ultra-High Resolution Frequency Meter," Proc. of the 35th Annual Frequency Control Symposium, USAERADCOM, Ft. Monmouth, NJ, May 1981, pp. 464-469.

[9] Allan, D. W., Barnes, J. A., "A Modified "Allan Variance" with Increased Oscillator Characterization Ability," Proc. of the 35th Annual Frequency Control Symposium, USAERADCOM, Ft. Monmouth, NJ, May 1981, pp. 470-475.

[10] Lesage, P., Ayi, T., "Characterization of Frequency Stability: Analysis of the Modified Allan Variance and Properties of Its Estimate," IEEE Transactions on Instrumentation and Measurement, Vol. IM-33, No. 4, December 1984, pp. 332-336.

[11] Barnes, J. A., Jarvis, Jr., S., "Efficient Numerical and Analog Modeling of Flicker Noise Processes," NBS Technical Note 604, June 1971.

[12] Barnes, J. A., "The Measurement of Linear Frequency Drift in Oscillators," Proc. of the 15th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Naval Research Laboratory, Washington, D.C., December 1983, pp. 551-582.

[13] Allan, D. W., Glaze, D. J., Machlan, H. E., Wainwright, A. E., Hellwig, H., Barnes, J. A., Gray, J. E., "Performance, Modeling, and Simulation of Some Cesium Beam Clocks," Proc. 27th Annual Symposium on Frequency Control, 309 (1972).

[14] Allan, D. W., Hellwig, H., "Time Deviation and the Prediction Error for Clock Specification, Characterization, and Application," <u>Proc. Position Location and Navigation Symposium (PLANS)</u>, 29 (1978).
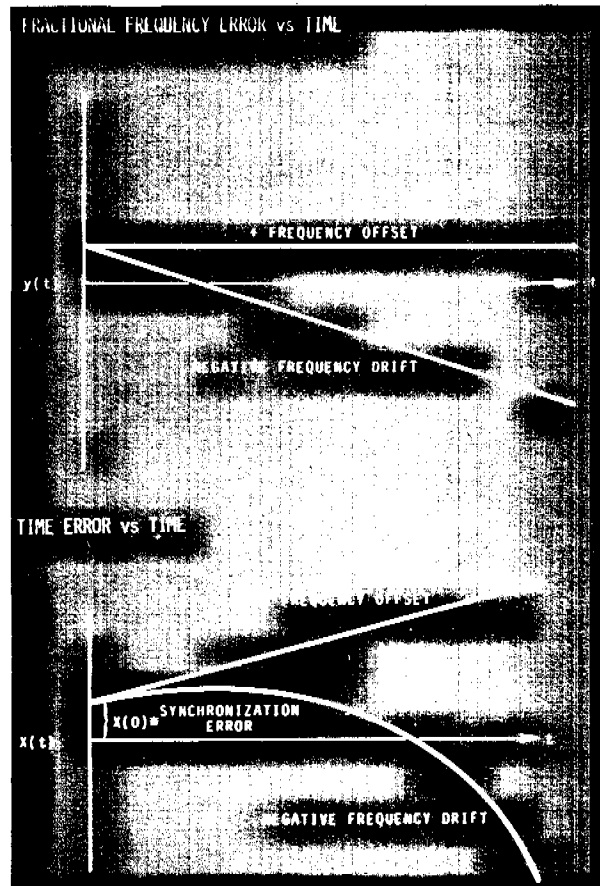
FRACTIONAL FREQUENCY ERROR vs TIME

+ FREQUENCY OFFSET

y(t)

NEGATIVE FREQUENCY DRIFT

TIME ERROR vs TIME

SYNCHRONIZATION
x(0)   ERROR

x(t)

NEGATIVE FREQUENCY DRIFT

Figure 1. Frequency, y(t), and time, x(t), deviations due to frequency
offset and to frequency drift in a clock.

OSCILLATOR
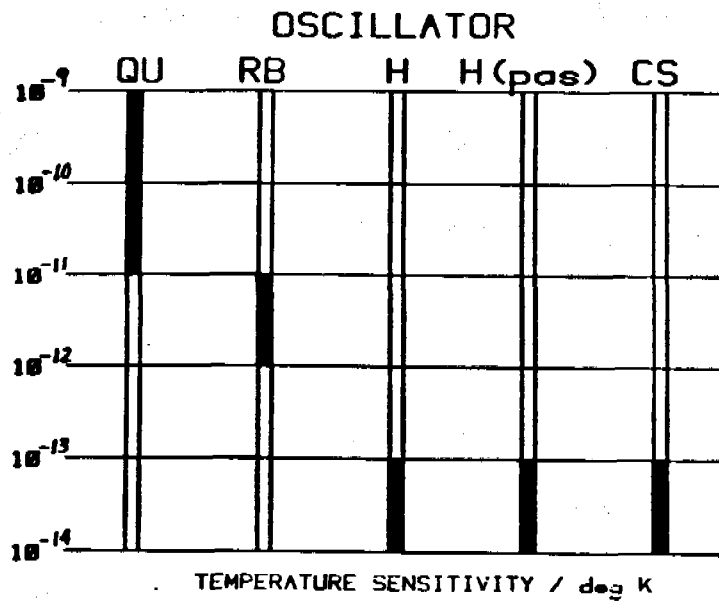
Figure 2. Nominal values for the temperature coefficient for the frequency standards: QU = quartz crystal, RB = rubidium gas cell, H = active hydrogen maser, H(pas) = passive hydrogen maser, and CS = cesium beam.
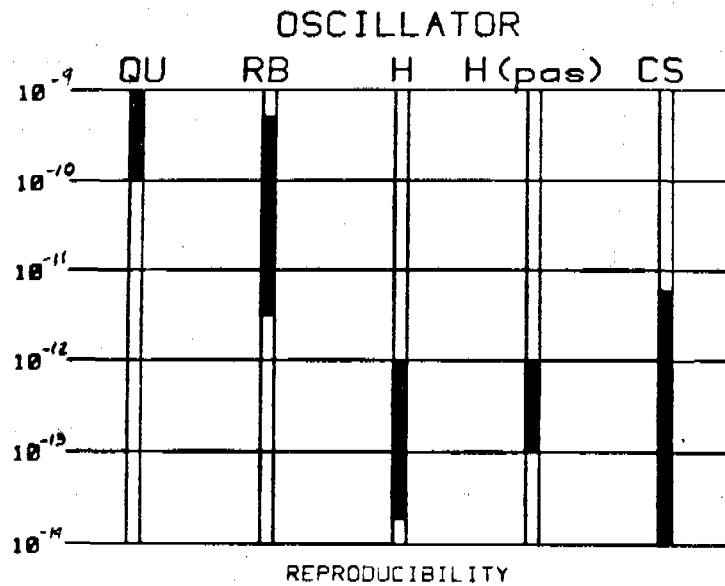
OSCILLATOR
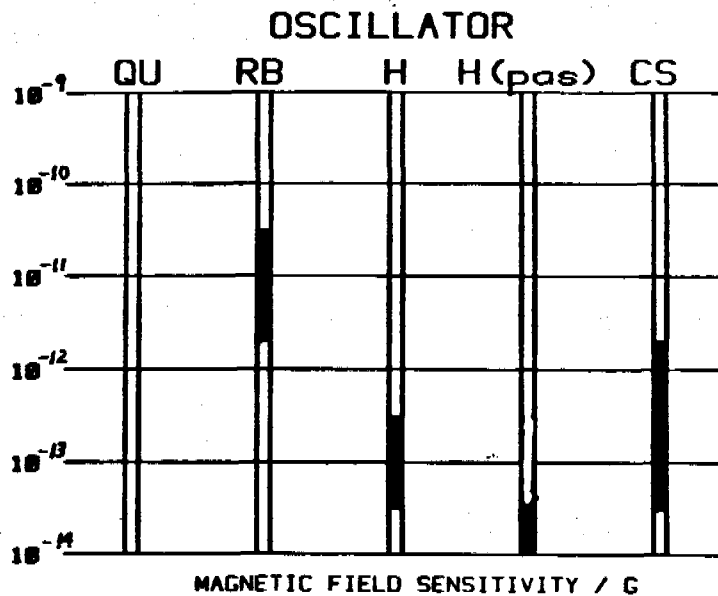
Figure 3. Nominal values for the magnetic field sensitivity for the frequency standards: QU = quartz crystal, RB = rubidium gas cell, H = active hydrogen maser, H(Pas) = passive hydrogen maser, and CS = cesium beam.

OSCILLATOR

Figure 4. Nominal capability of a frequency standard to reproduce the same frequency after a period of time for the standards: QU = quartz crystal, RB = rubidium gas cell, H = active hydrogen maser, H(pas) = passive hydrogen maser, and CS = cesium beam.
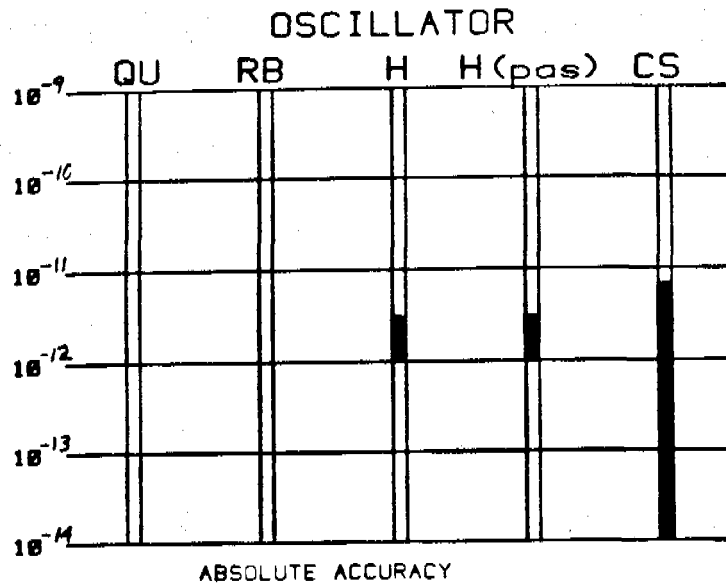
OSCILLATOR

Figure 5. Nominal capability for a frequency standard to produce a frequency determined by the fundamental constants of nature for the standards: QU = quartz crystal, RB = rubidium gas cell, H = active hydrogen maser, H(pas) = passive hydrogen maser, and CS = cesium beam.
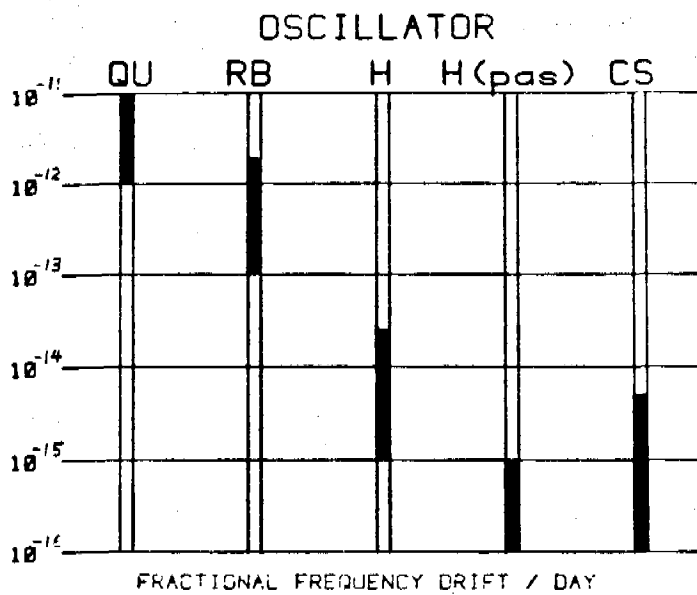
OSCILLATOR

Figure 6. Nominal values (ignoring the sign) for the frequency drift for the frequency standards: QU = quartz crystal, RB = rubidium gas cell, H = active hydrogen maser, H(pas) = passive hydrogen maser, and CS = cesium beam.
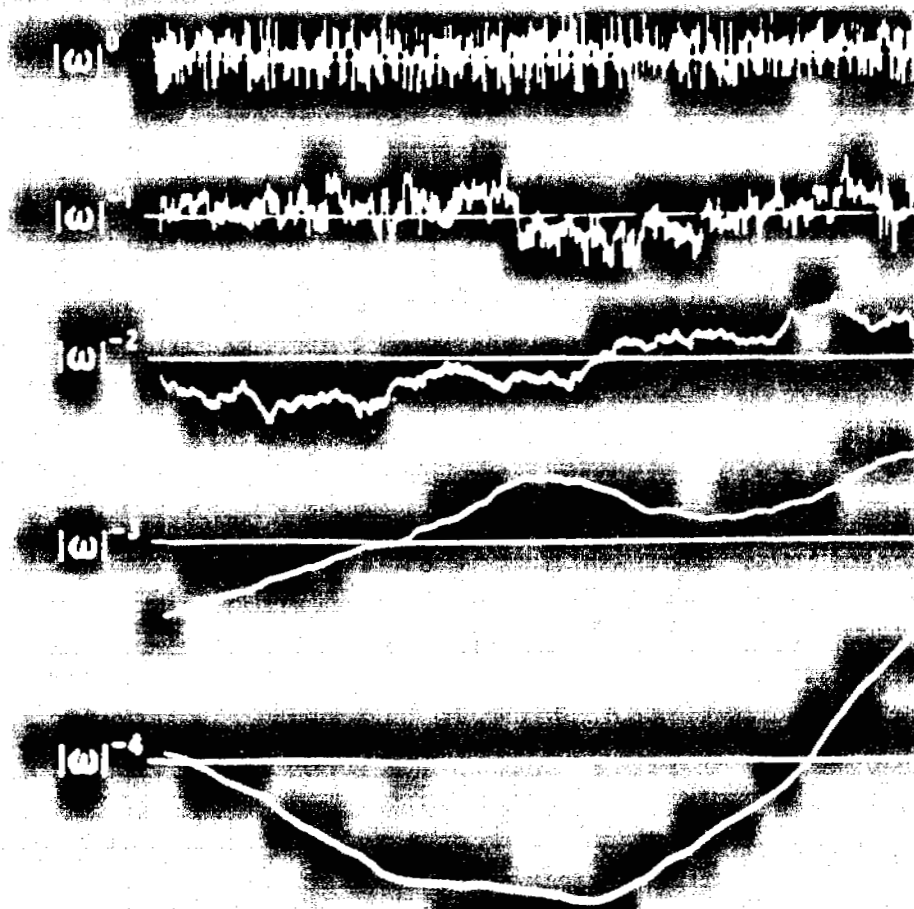
POWER LAW SPECTRA

$|\omega|^{0}$

$|\omega|$

$|\omega|^{-2}$

$|\omega|^{-3}$

$|\omega|^{-4}$

Figure 7. Simulated random processes commonly occurring in the output signal of atomic clocks.
Power law spectra $S(\omega)$, are proportional to ω to some exponent, where ω is the
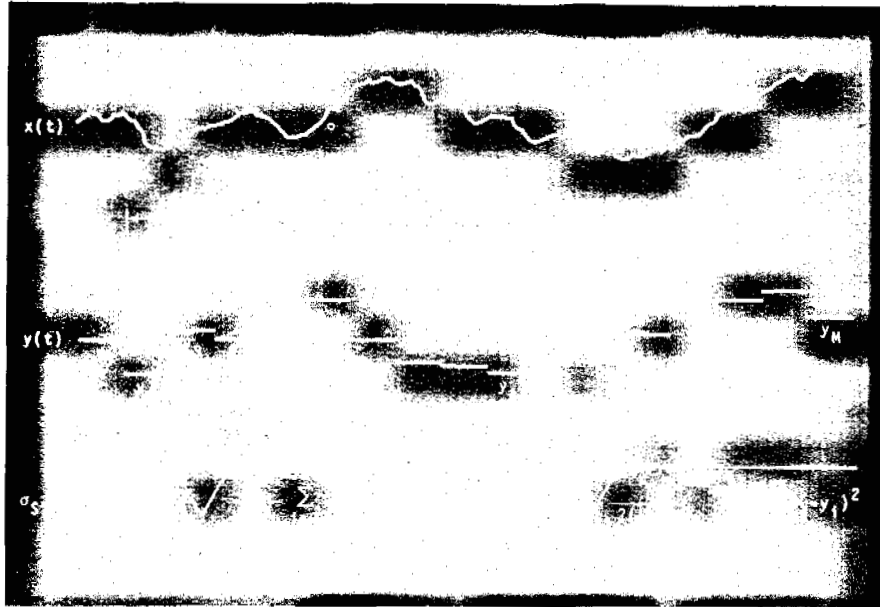Fourier frequency.

Figure 8. A simulated time deviation plot, x(t), with indicated ;sample
time τ over which each adjacent fractional frequency, y_i, is measured.
The equations are for the standard deviation and for the estimate of $\sigma_y(\tau)$
for a finite data set of M frequency measurements. It is often the case
that the standard deviation diverges as the data length increases when
measuring the long term frequency stability of precision oscillators,
wehereas $\sigma_y(\tau)$ converges.

'Allan variance' concept

difference in slope = $\Delta y = Y_2 - Y_1$
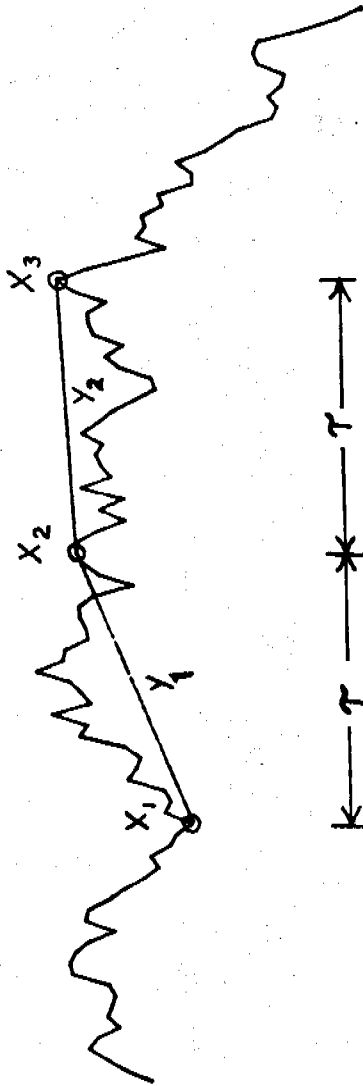
$$\sigma_y^2(\tau) = \frac{1}{2}\overline{(\Delta y)^2} \sim \tau^\mu$$

TIME

x = TIME DIFFERENCE

Figure 9. Pictorial of computation of Allan variance. The simulated time variations plotted are random walk. At a set sample time τ $\Delta y = (x_3 - 2_{x2} - x_1)/\tau$ are computed as the x's take on all possible values for the data set. Each Δy is squared and an average squared value determined, $\overline{(\Delta y)^2}$; taking 1/2 of this gives the Allan variance for that value of τ. τ can then be changed either in the hardware or the software to determine the Allan variance for another value of τ.
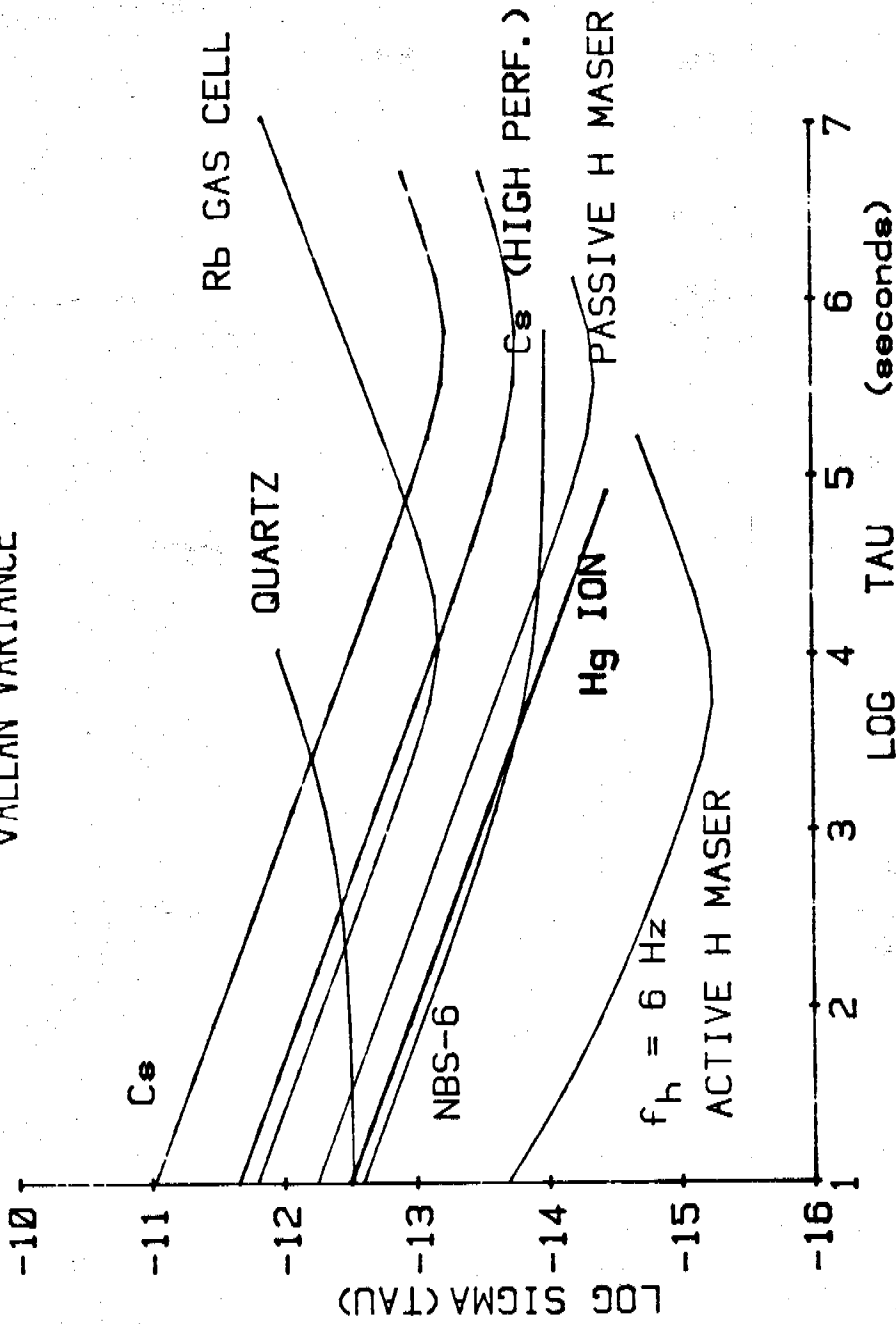
Figure 10. Square root of the Allan variance for a variety state-of-the-art precision oscillators including NBS-6, the NBS primary frequency standards.
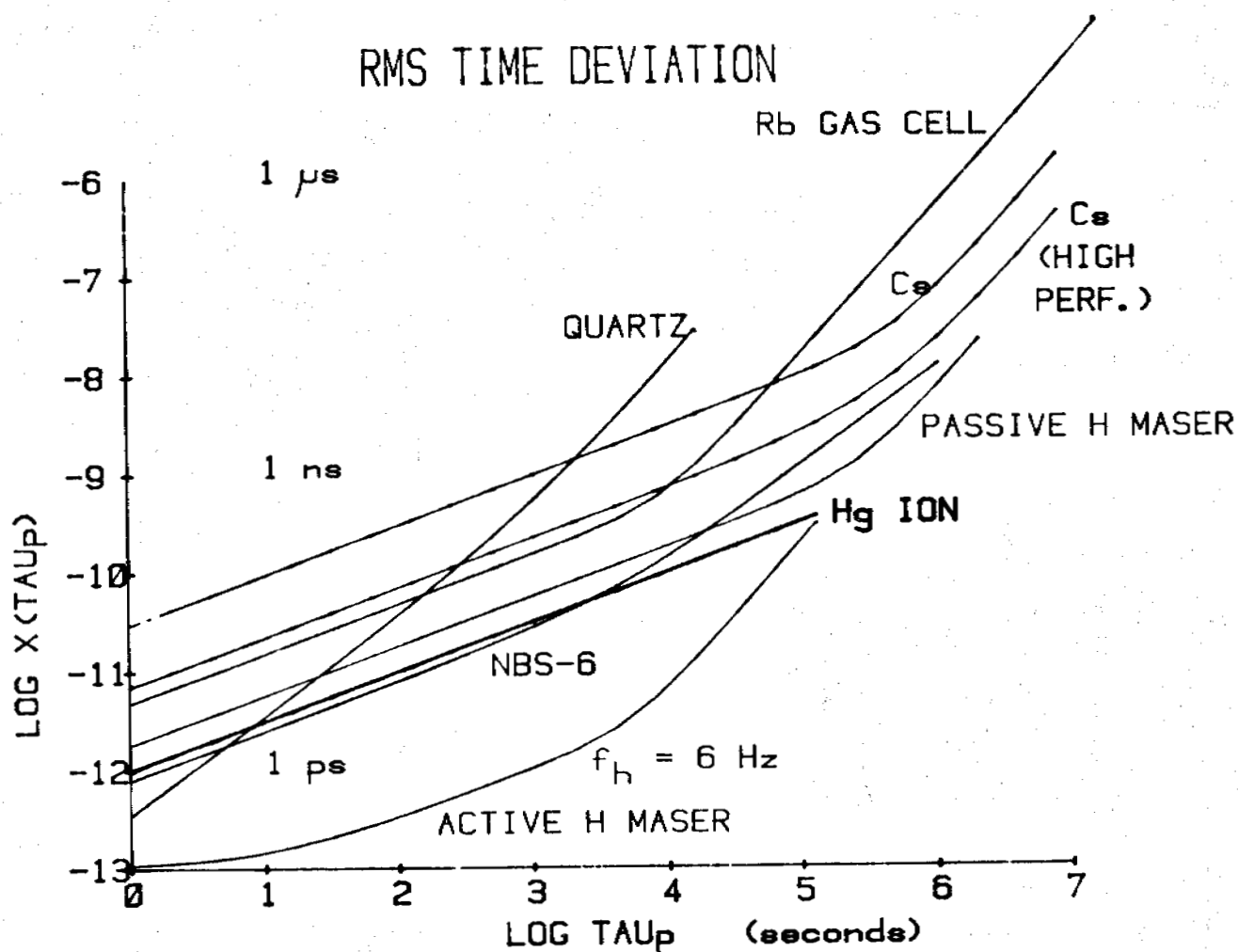
Figure 11. From the frequency stability characterization shown
in Figure 10, optimum prediction algorithms to minimize the time
error can be obtained. Based on optimum prediction procedures
the RMS time prediction error for a prediction interval τ_p can be
calculated for each of the oscillators shown in Figure 10, and
the corresponding values are plotted in Figure 11.

Table I.

Typical		Applicable Oscillators and Range of Applicability				
α, Noise Types	Cs	H-active	H-passive	Qu	Rb
2 White Noise PM		≤ 100 s		≤ 1 ms	
1 Flicker Noise PM				≤ 1 s	
0 White Noise FM	≥ 10 s	100 s $\leq \tau \leq 10^4$ s	≥ 1 s		≥ 1 s
-1 Flicker Noise FM	$\geq$ days	$\geq 10^4$ s	$\geq$ days	≥ 1 s	$\geq 10^4$ s
-2 Radom Walk FM	$\geq$ weeks	$\geq$ weeks	$\geq$ weeks	$\geq$ hours	$\geq$ days

Table II

Typical Noise Types	Classical Standard Deviation of x	Classical Standard Deviation of y
White Noise PM	$\tau \cdot \sigma_y(\tau)/\sqrt{3}$ (constant)	$\sigma_y(\tau)\sqrt{\dfrac{2(N+1)}{3N}}$
Flicker Noise PM	$\sim \tau \cdot \sigma_y(\tau)\sqrt{\dfrac{\ln M}{\ln 2}}$	$\sim \sigma_y(\tau)\sqrt{\dfrac{2(N+1)\ln\tau}{3N}}$
White Noise FM	$\tau_0 \cdot \sigma_y(\tau_0)\sqrt{\dfrac{M+1}{6}}$	$\sigma_y(\tau_0)$
Flicker Noise FM	undefined	$\sigma_y(\tau)\sqrt{\dfrac{N \ln N}{2(N-1)\ln 2}}$
Random Walk FM	undefined	$\sigma_y(\tau)\sqrt{N/2}$

Table III.

Typical Noise Types	Optimum Prediction $x(\tau_p)$ rms	Time Error: Asymptotic form
White Noise PM	$\tau \cdot \sigma_y(\tau)/\sqrt{3}$	constant
Flicker Noise PM	$\sim \tau \cdot \sigma_y(\tau) \sqrt{\dfrac{\ln \tau_p}{3\ln \tau_o}}$	$\sqrt{\ln \tau_p}$
White Noise FM	$\tau \cdot \sigma_y(\tau)$	$\tau_p^{1/2}$
Flicker Noise FM	$\tau \cdot \sigma_y(\tau)/\sqrt{\ln 2}$	τ_p
Random Walk FM	$\tau \cdot \sigma_y(\tau)$	$\tau_p^{3/2}$

QUESTIONS AND ANSWERS

PETER KARTASCHOFF, SWISS PTT:
Dave, I have a question which I will direct to you and to Gernot Winkler. We have measurements and we have recommended designations such a Y, F sub Y, sigma sub Y and so on. Ten years ago we had a recommendation in the CCIR, International Radio Consultative Committee, with an associated report which explains all the terms which you have so well presented here. Everybody uses them. Fifteen years ago there was a subcommittee of the IEEE which recommended these measurements. Do you, or anyone here, know if there is any other published international recommended standard, IEC or IEEE besides the CCIR recommendation? As far as I know, the CCIR recommendation still stands alone and I do not know what the IEEE is doing or what the IEC is doing about this.

MR. ALLAN:
As far as I know, there is nothing new beyond those that you have mentioned. There is an IEEE subcommittee in existence, but it is not active. I know of no other committee.

A COMMENT HERE BY MR. WINKLER, NOT INTO THE MICROPHONE AND NOT DECIPHERABLE.

MR. BEARD:
You mentioned environmental effects and such things as that on characterizing these clocks. Aren't the properties of the measuring system with which you make these measurements also important?

MR. ALLAN:
I really appreciate your mentioning that. Very often a process, in fact the millisecond pulsar is an example, is a measurement of the measurement noise. We think that we cannot get the real millisecond pulsar. This is a very good case in point, because all you are seeing is the measurement system. In many applications that can be a problem. This goes back to the point that Dr. Winkler made: one should make your phase or time measurement system as good as possible so that you don't have measurement noise contamination.