TUTORIAL:
CLOCK AND CLOCK SYSTEMS
PERFORMANCE MEASURES

David W. Allan
Allan's TIME

Introduction

This tutorial contains basic material — familiar to many. This will be used as a foundation upon which we will build — bringing forth some new material and equations that have been developed especially for this tutorial. These will provide increased understanding toward parameter estimation of clock and clock system's performance.

There is a very important ITU Handbook being prepared at this time which goes much further than this tutorial has time to do. I highly recommend it as an excellent resource document. The final draft is just now being completed, and it should be ready late in 1996. It is an outstanding handbook; Dr. Sydnor proposed it to the ITU-R several years ago, and is the editor with my assistance. We have some of the best contributors in the community from around the world who have written the ten chapters in this handbook. The title of the Handbook is, "Selection and use of Precise Frequency and Time Systems." It will be available from the ITU secretariat in Geneva, Switzerland, but NAVTEC Seminars also plans to be a distributor.

Definitions and Concepts

If we ask the very simple question, "What is a clock?" We discover that essentially all clocks can be considered two-part devices: a resonator or frequency source and counter or divider for keeping track of the number of oscillations. As an example, we have the definition of a second: when 9,192,631,770 oscillations occur of the photon associated with the quantum ground-state of the cesium-133 atom, we have a second. The electronics for counting are more sophisticated than for the ordinary clock, but the concepts are the same.

Given this concept, it is important to remember that the counter (divider) will always deteriorate the signal. In other words, the phase noise of the sine wave of the source, for example, will be more stable than the clock's output. We will come back to this when we talk about measurement noise and optimization algorithms.

Consider another very important concept: If we have two clocks, we feel obligated to ask the question, which is correct under the basic statistical theorem that every clock disagrees with

every other clock? The disagreement may be very small, but they will always disagree except at most an instant. In fact, you can make the very strong statement that it is impossible to have two clocks perfectly synchronized, except at an instant, because of noise. So, you can never have two clocks perfectly synchronized on a continuous basis.

In a fundamental sense, what is the difference between frequency and time? We can talk about frequency in an absolute sense as in the definition of the second. Time is not absolute. It is an artifact — depending on when we set the integrator as we started counting seconds, for example. Independently, a frequency standard built in Braunschweig, Germany will agree with one built in Boulder, Colorado within their accuracy limits because they are based on the same fundamental phenomena within physics.

Time, on the other hand, depends upon when you set the integrator for counting the seconds. Hence, two clocks may be arbitrarily different in their readings. There is no absolute time with which to compare a clock.

Figure 5 compartmentalizes the perceived causes of clock deviations into four areas:

- 1) How we process the data is very important. Clock performance can be made worse by improper processing.

- 2) The measurement noise limits our ability to see the true performance of a clock. A classic example is GPS SA. If it were turned off, the venetian blinds would go up, and we would be able to see the GPS satellite clocks very clearly. The measurement noise should always be considered; it may be negligible compared to the clock noise, but often it is not.

- 3) Every clock has intrinsic mechanisms which perturb its output time. By nature it is convenient to have two sub-categories for the intrinsic clock perturbations — typically denoted by the random variations and by the systematic variations. Random and systematic variations also occur in measurement systems as well as in the environmental perturbations. The environment perturbations can look like a random process, depending upon how it couples into the clock. And

- 4) the environmental perturbations often adversely impact the long-term performance of a clock. Clearly, it is desirable to design to minimize the environmental perturbing effects.

Let us now review some fundamental clock concepts (see Figure 6) A working definition of time is the apparent reading of a clock. Synchronization is to have two or more clocks with the same apparent time reading. In principle, this has to be within some level of uncertainty since every clock disagrees with every other clock except at most an instant. Syntonization is to have two or more clocks with the same apparent rate. In the telecommunication industry, they word "synchronization" is often used when, in fact, what is meant is "syntonization." These are two different useful concepts, and if used properly can help avoid a lot of confusion in specifications and in system performance descriptions. Syntonization also needs to be specified with an uncertainty.

Simultaneity doesn't really require a clock. Here two or more events occur at the same moment, as perceived in some reference frame.

The environmental elements which we often see that perturb clocks are listed in Figure 7. This list contains some of the more important ones. One of the fundamental breakthroughs with the HP model 5071A cesium clock is that by active electronic control and feedback the effect of the environment can be greatly reduced. This approach has improved the long-term stability by more than an order of magnitude over other models.

There is a new IEEE standard, 1193-1994, which gives guidelines for the measurement of environmental sensitivities. Later, we will use some of the messages from this standard. This standard is another good reference document. The place to order it is shown in Figure 8.

Measurement Models, Terminology and Concepts

Figure 9 shows the usual model for an oscillator's output. The frequency, $\nu(t)$, always varies with time; hence, its period, $\tau(t)$, also varies. One of our main goals in this tutorial is to clearly and in a parsimonious way quantify these variations so that there is good communication between the vendor and the user, and so that the designer or planner may work effectively and efficiently. The output need not always be a sine wave. The following characterization procedures have been kept general enough to work as well for square wave or for any periodic signal. But since sine waves are fundamental to nature, this is the common representation.

Using the sine-wave as a conceptual model, we usually have a nominal frequency at which the standard is designed to work, ν_0. By definition it does not vary with time. We then use $\phi(t)$ to denote all the phase variations around the nominal accumulated phase, $2\pi\nu_0 t$. The cycles of an oscillator are counted to create a clock. Again, the divider noise will degrade the signal from the oscillator. Hence, without some special filtering, the integrated clock noise will always be greater than the integrated oscillator noise.

As illustrated in Figure 10, what is measured in practice is never the time of a clock, since we have no absolute reference with which to measure it alone. What, in fact is measured is the time difference between two clocks. The time difference can be measured with arbitrary precision. Today there exists instrumentation which can measure time differences at the femtosecond level using the carrier phase.

Figure 11 shows the normalized representation of frequency offset $y(t)$. This is a dimensionless quantity which is simply defined as the free-running frequency, $\nu(t)$ of the clock, minus its nominal frequency, ν_0, all divided by the nominal frequency. Even though this is a conceptual value, in practice it is very useful because ν_0 can be the reference oscillator of the pair being measured. In addition, there is the big advantage that $y(t)$ is a small number compared to $\nu(t)$. Conceptually $y(t)$ represents the offset from the ideal. It is often referred to as parts in 10^{10}, for example, or equivalently 1×10^{-10}.

On the other hand, the time offset, $x(t)$, is the exact integral of the frequency offset, $y(t')$, integrated from 0 to t. It also can be written exactly in terms of the $\phi(t)$ shown in Figure 9, $x(t) = \frac{\phi(t)}{2\pi\nu_0}$. We often talk about time deviations or phase deviations interchangeably, and since

they are directly proportional this is okay.

Because of the integral relationship shown in Figure 11, the fractional frequency offset, $y(t)$ is the time derivative of the time offset. Hence, the slope on a phase plot is proportional to the frequency offset.

Figure 12 gives a simple parsimonious model for the time offset or time error of a clock. The first term represents the synchronization error, x_0. The second term contains the syntonization error, y_0. It gets multiplied by the running time to calculate its effect on the total time error. The third term contains the linear frequency drift. Its dimensions will be fractional frequency change per unit time interval, per second or per day, as examples. All of the rest of the deviations are included in $\epsilon(t)$. Here, we often hide a multitude of sins! This last term, for example, could represent all of the effects due to environmental perturbations while also containing the random noise deviations. In addition, it may contain side-band components due to diurnal effects, or to modulation or RFI.

If we subtract off the effects of the first three systematic terms, then $x(t) = \epsilon(t)$. Analyzing these residuals is very helpful in diagnosing the effects of the random and other perturbations on the clock. Once the level and kind of random perturbations are known, then optimum estimation procedures can be used to better estimate the systematic effects as well as being able to calculate optimum predictions, for example.

Taking the derivative of the model equation in Figure 12 yields: $y(t) = y_0 + Dt + \dot{\epsilon}(t)$. Writing the equation this way will be useful later as we get into optimum parameter estimation.

Frequency and Time Accuracy and Stability

Figure 13 shows an example of two very simple systematic situations: a positive frequency offset, and a negative frequency drift. The first drawing illustrates $y(t)$ and the second one its integral, $x(t)$. The constant frequency offset turns into a ramp for the time error, and the drift into a quadratic. We assume the same synchronization error (constant of integration) for both situations.

In 1988, IEEE Standard 1139-1988 was published providing a recommended set of measures for time and frequency characterization. Figure 14 gives some of those measures from this standard. The exception is $\sigma_x(\tau)$, as it had not been developed at that time. Subsequently, it has been adopted by the telecommunications community and by the ITU-R. As we need them, we will describe the functionality of some of these measures.

Figure 15 gives some time-domain definitions and some useful measures. For example, $y(t)$ is a direct indicator of frequency accuracy if the reference, ν_0, is the definition of an agreed upon standard. Similarly, $x(t)$ is a direct indicator of the time inaccuracy if it is taken with respect to UTC, which is the correct time by definition. The other three sigma measures shown are for determining the level and kind of instabilities, as will be shown later.

As was mentioned before, the design of the relatively new HP 5071A cesium-beam clock was for increased accuracy and improved immunity to environmental perturbations — resulting in

greatly improved long term stability. Figure 16 is a histogram of $y(t)$ for the 94 HP 5071A clocks contributing to TAI/UTC during 1994. Figure 17 shows the 311 total participating clocks during 1994 plotted with the same abscissa. It is apparent that the design goals have nicely been met. The accompanying paper gives more details as well as documenting the performance of TAI/UTC[1]. The introduction of the HP 5071A clocks, as Dr. Thomas has pointed out, is having a major impact toward improving the performance of International Atomic Time.

In both Figures 16 and 17 the mean is significantly larger than the standard deviation of the mean. So in both cases the standards would not be considered in statistical control. Hence, the need for primary standards, so that calibrations with same can provide frequency accuracy; i.e. agreement with the definition of the second.

Figure 18 a frequency stability diagram — using $\sigma_y(\tau)$ — showing the range of values available for most of the important clocks to our community. This stability diagram is taken from an ITU-R document giving the characteristics of these clocks[2]. QZ stands for quartz crystal oscillators, RB stands for rubidium-gas-cell frequency standards, CS stands for cesium-beam frequency standards, and HM stands for hydrogen-maser frequency standards. For CS stabilities, an extended line has been drawn in representing the improved long-term frequency stability of the relatively new HP 5071As.

Figure 19 is a plot of the time accuracy of three time scales over the last approximately 200 days: UTC(NIST), UTC(OP), and UTC(USNO-MC). These three time scales are taken with respect to UTC, the official time for the world. By definition, how well a clock agrees with UTC is a measure of its true time accuracy. All three have been within nominally 100 nanoseconds for about the last half year. The time accuracy of many of the worlds time scales have improved significantly over the last three years. This has been primarily driven with the introduction of the HP 5071As into these sundry time scales.

One of the most significant challenges that a timing center has toward time accuracy is in predicting where UTC will be at the current time, because UTC is calculated and distributed about one and one-half months after the fact. Each country maintains its own real-time estimate of UTC — denoted UTC(i) for the ith timing center. Clearly, if UTC were available in real-time, it would be far simpler to have a high-level of time accuracy. Through international cooperation, this direction is being pursued.

The limiting noise for the cesium clocks contributing to TAI/UTC is white-noise FM. The optimum RMS prediction error for this noise is $\tau_p \sigma_y(\tau_p)$, where τ_p is the prediction interval. For the USNO data over the last half of year, the RMS error is 6 ns. This is not the same as the standard deviation; the 6 ns is with respect to the truth, which is UTC by definition.

An RMS error of 6 ns with τ_p = about 45 days implies that $\sigma_y(\tau_p) \leq 1.5 \times 10^{-15}$, since any prediction algorithm cannot be better than optimum. If the USNO time scale and UTC were independent, then this number would be directly related to the square root of the sum of the variances from each scale. The weight of the USNO clocks contributing to TAI/UTC is about 40 percent. The effect of the bias of a time scale contributing is given approximately by 1/(1 - weight). Hence, we can conclude that either of the two time scales is equal to or better than 2.6×10^{-15} at τ = 45 days, and one of the scales is better than this number divided by $\sqrt{2}$, or

213

than 1.8×10^{-15}.

This level of stability represents a major advancement during the last three or four years. And again, it comes mainly as a result of the introduction of the HP 5071A clocks with their excellent environmental insensitivity. One would also conclude that the prediction algorithm used by USNO is very close to optimum.

Random Processes, Models and Measures

Characterizing the random deviations in a clock's performance allows us to determine the noise type. Knowing the type of noise then allows us to design optimum parameter estimation procedures. Figure 20 illustrates two very important types of noise. Since one flip of a coin is independent of the next flip, a series of flips generates a random and uncorrelated series. In other words, a flip of heads at one point in time has no bearing on whether the coin will come up heads or tails at another time. The spectral density of these flips is then a white-noise process.

We can integrate these flips by taking one step forward with heads and one step backwards with tails. Our displacement from the origin is now a random-walk process and has an f^{-2} spectral density. These same arguments are very analogous as to why the random time deviations out of most atomic clocks are a random-walk process. The atomic-clock servo hunts for the resonance frequency being limited by white noise in the search; the integral of these white frequency deviations generates a random-walk in the time deviations. Vice-versa, if a derivative or first difference is taken of random-walk time deviations, the process turns into one with a white-noise spectrum.

The HP 5071A is an excellent example of a clock with classical white-noise frequency spectrum over many decades of Fourier space. This kind of noise causes $\sigma_y(\tau)$ to go as $\tau^{-1/2}$, and for the high-performance model of this clock the white-noise behavior extends from about 10 seconds to as long as 10^7 seconds in some cases with a performance specification given by the top equation in Figure 21. Such behavior results in long-term stabilities well below 1×10^{-14}.

As also illustrated in Figure 21, whereas white-noise FM is the ideal classical noise for most atomic clocks, white PM is the ideal classical measurement noise. That measurement noise can, of course, contribute to $\epsilon(t)$ in the general model equation for the time error between two clocks.

As shown in Figure 22, typically five different noise processes are employed to model clocks, oscillators and measurement systems. These seem to be fairly basic in nature. Figure 23 gives the Fourier transformation relationships between the time-domain measures and the frequency-domain representation, as well as the region of applicability. Using these relationships and going back to Figure 17, one can see both the regions of applicability (from the different slopes corresponding to the f values) as well as the different levels of random variations.

Figure 24 gives the abbreviation, the name and the mathematical expression for each of these three time-domain measures. Their square roots are: ADEV, MDEV and TDEV, respectively. The first two measures are explained in detail in NIST Technical Note 1337[3] and all three

214

in the upcoming ITU-R Handbook. The transformation coefficients from the time-domain to the frequency-domain or *vice-versa* (preserving the noise type and level) may be found in reference[4] and for AVAR and MVAR in NIST TN 1337. TVAR $= \tau^2 \text{Mod}.\sigma_y^2(\tau)/3$ has been shown to be a very good measure for measurement system stability, network stability, and time dissemination stability. TVAR was developed after the publication of NIST TN 1337.

Note that the equations for AVAR, MVAR and TVAR in Figure 24 are all represented in terms of a second difference of the time deviations, x. In the case of MVAR and TVAR, the x values making up the second difference are each averaged over a separate, but sequentially adjacent interval τ — rather than being a time error measurement at a point as for AVAR. The effect of averaging the data in an appropriate way, applies a filter in the software so that it effectively modulates the bandwidth of the software processor. This bandwidth modulation removes the ambiguity associated with AVAR; i.e. AVAR has essentially the same slope ($\mu = -2$) value, for either white-noise PM or for flicker-noise PM. MVAR and TVAR can distinguish between white-noise PM and flicker-noise PM — having different slopes when plotted logarithmically versus τ.

Applications of Optimum Parameter Estimation and Prediction

As shown in Figure 25, optimum parameter estimation means that once a model parameter has been determined, the residuals around that parameter model have been minimized in a squared-error sense. Similarly, for prediction, the errors of prediction are minimized in a squared-error sense. Of course, both parameter estimation and prediction will depend upon the type and the level of the noise processes involved. Hence, knowing the noise type and level is essential for optimum parameter estimation and prediction.

The statistical theorem given in Figure 26 is important, as well as useful and simple. In particular, it is useful for parameter estimation and for prediction. Since nature gives us white PM and white FM, this theorem is directly applicable in these cases. In addition, the long-term performance of most clocks may be reasonably well modeled as a random-walk FM process; this is sometimes called white acceleration because the second derivative of x (d^2x/dt^2) has a white spectrum. Here again we may use the above theorem.

Two very important examples are the following. In Figure 21 it was pointed out that white FM is the classical noise for most atomic clocks, and white PM is the classical noise for an ideal time-difference measurement system. Hence, as illustrated in Figure 27, in the presence of white FM, AVAR is the optimum estimator of frequency change (or instability). This is true since each of the average frequencies, taken over an interval τ, is the optimum estimate of frequency over that interval. Comparing an optimum with an optimum causes the difference to be an optimum estimate of the change. AVAR then is an RMS computation of this optimal estimate of change over the interval τ. Similar arguments hold for TVAR in the presence of white-noise PM making it an optimal estimate, in an RMS sense, of the change in the time each averaged over an interval τ.

Flicker models also are very common; they are more arduous to deal with, but filters have been designed that turn flicker residuals into white noise — providing the opportunity of developing

optimum estimation and prediction procedures for $1/f$ type noise processes. These have only been partially developed because of their complexity.

Figure 28 gives the uncertainties associated with optimum estimation for three different circumstances. The first equation is an applicable model if the same clock signal is fed into both input ports of a time-difference measurement system. In the ideal case $\epsilon(t)$ would have a white-noise PM spectrum with mean zero. Hence, the mean value over a data set would be the optimum estimate of the time-delay difference, x_0, in the cable delays feeding the two input ports. The standard deviation of the measurements is given by $\sigma_x(\tau_0)$, and the uncertainty in this estimate is given by the standard deviation of the mean, where N is the number of measurements. Such a model may also be appropriate if two very good atomic clocks, remote from each other, were being compared using the GPS common-view technique. The day-to-day measurement noise is often characterized by white-noise PM, and if this noise is significantly higher than the clock noise at $\tau = 1$ day, then the simple mean gives the optimum estimate of the time difference between the remote clocks as averaged over the interval τ.

The second equation in Figure 28 would be a reasonable model if two independent clocks had negligible noise as compared to the measurement system's white PM level over the time of the measurement. This model also assumes there is no frequency drift between these two independent clocks. In this case, a linear regression provides the optimum estimate of the synchronization error, x_0, and the syntonization error, y_0, between the two clocks, since the residuals will have a white spectrum. The uncertainty is given at the right; notice that the confidence on the frequency-difference estimate improves as $N^{-3/2}$, whereas the confidence on the time-difference estimate only improves as $N^{-1/2}$. This is because we are estimating frequency in the presences of white-noise time residuals, and frequency and time are related by a derivative, $y = dx/dt$. We will show later that this $N^{-3/2}$ factor may be used to significant advantage in some frequency transfer experiments, such as with GPS and with Two-Way Satellite Time and Frequency Transfer.

In the third equation in Figure 28, the model, for example, could be for two clocks with relative frequency drift between them along with having time and frequency offsets. Again, the clock's random noise is negligible as compared to the white-noise PM measurement noise over the length of the measurement. This model could also be applicable for a clock with intrinsic white-noise PM, such as active hydrogen masers and quartz crystal oscillators have in the short-term. In this case, the quadratic regression line is the best fit, because the time residuals, $\epsilon(t)$, have a white spectrum. For similar arguments, the confidence of the estimate of the drift term improves as $N^{-5/2}$. That is, $4D = d^2x/dt^2$ is being optimally estimated in the presence of white-noise time residuals.

If we apply the second equation in Figure 28 to Dr. Mattison's experiment, reported in this conference[5], we get some very impressive results. With data taken once a second, having 100-picosecond white PM measurement noise, and having the satellite in view for 5 minutes, the uncertainty on the frequency measurement would be about 6.7×10^{-14}. Now if the data rate could be speeded up to a 1 ms rate, then the uncertainty becomes 2.1×10^{-15} — a factor of 30 improvement for the period of observation. The uncertainty expression at the right of the second equation in (28) is equivalent to 2 Mod.$\sigma_y(\tau)$, where τ is the observation interval.

216

Figures 29 through 31 apply to the case where classical white-noise is predominant as for most atomic clocks. The first equation in Figure 29 represents the true average normalized frequency over the interval τ_0 as determined from the time difference at the beginning and the end of the interval. Such a measurement is much like is done in a time interval counter over its gate time. The second equation is the definition of the average over the whole data length. Hence, if the first equation is substituted into the second equation, the result is the third equation. Therefore, the end point time-difference values yield the optimum estimate of the frequency in the presence of white-noise FM. The algorithm is extremely simple: the difference of the last point minus the first point divided by the data length. It is well to check either visually or statistically that neither of these two points is an outlier, which would contaminate the result. It is always good practice to check the data visually. Looking at the time residuals after subtracting the systematics is one of the most useful visual inspection techniques.

Since it is not uncommon for people to subtract a linear regression from the phase or time residuals to determine the frequency of their atomic clock from the slope thus derived, Figure 30 is a simulation showing the degradation in this estimate as compared with the optimum. This figure gives the results from a Monte Carlo analysis of 100 simulations of 100 points each. The mean frequency from the regression line slope was 72% worse than optimum. The standard deviation of the frequency residuals was 8.5% worse. The simulations were derived from a normally distributed set with unit variance for the white-noise FM frequency residuals. The column denoted "Mean x_0" is the average value of the synchronization term derived at the origin of each set and is zero by design in the optimum estimation procedure. The optimum value for time prediction is the last value, which is the value used in the optimum estimate of the frequency for the measurement period.

USNO has 40 HP 5071A clocks. They are well modeled by the first pair of equations in Figure 31 for τ values out to the 45 day prediction time needed to bring the UTC estimate forward to the current time. Using the white-noise model equation for the HP 5071A clock given in Figure 21, and the uncertainty relationship given at the right in (31), we obtain for the frequency measurement uncertainty, for $\tau = 45$ days and for 40 independent clocks, 6.4×10^{-16}. We previously deduced from the data an upper limit of 2.6×10^{-15} as derived from the actual prediction error in UTC(USNO-MC) as observed over the last half year.

The prediction upper limit is about a factor of four worse than optimum. From the previous analysis, we cannot tell whether the major contributor to the instability is TAI/UTC or USNO. It is possible that the white-noise FM model starts to break down for some of the clocks for τ values of the order of a couple of months. An other explanation for the disparity could be that optimum parameter estimation and prediction may not be used in the generation of TAI/UTC. In talking to personnel at USNO, it seems that their procedure is very close to optimum. Regardless, the results obtained are greatly improved over what they were a few years ago.

One may notice that the expression for the uncertainty at the right of the first pair of equations in Figure 31 is the same as the standard deviation of the mean. The second pair of equations in Figure 31 is the model for two clocks having relative frequency drift and where the predominant noise is white FM. In this case the linear regression on the frequency is the optimum estimator, because the residuals around that regression line are white. The uncertainty on that drift

217

estimate decreases as $N^{-3/2}$. This kind of regression analysis is often used in our community, and is obviously very useful.

Figure 32 considers the random-walk FM model as the predominant noise. This is often the model used for clocks for their long-term stability performance. The model in Figure 32 assumes the presence of frequency drift. The second difference of the $x(t)$ data has a white spectrum. Hence, from our statistical theorem the mean value of the second difference is an optimum estimator. This mean value is directly relatable to the drift as shown and which has an uncertainty given at the right. This uncertainty is equivalent to the standard deviation of the mean.

Unfortunately, as shown in Figure 33, life is not so simple. We almost never have single noise processes in a data set. But a filter can almost always be designed which will give white residuals. It may be a complex filter.

Figure 34 is an illustration useful to our community: the case of white PM and/or white FM with long-term random-walk FM. An appropriate filter may be designed to average down the white-noise PM and/or the white-noise FM, and then we can analyze the random-walk FM residuals. If the random-walk FM is the predominant noise in the long-term, as it often is, then a simpler algorithm for determining a near optimum estimate of the frequency drift is as follows. If the first, the middle, and the last time-difference points are used to compute the estimate of the frequency drift, D, this estimate has two distinct advantages. First, as a second-difference estimate it is optimum for random-walk FM. Second, the effect of the higher frequency noise processes (e.g. white-noise PM and white-noise FM) is diminished if the τ for half the data length is long compared to those τ values where the higher frequency noise processes predominate. If these higher frequency noise processes have been filtered, so much the better.

In the case of white-noise FM and frequency drift a linear regression to the frequency gives the optimum estimate of the drift. But in this case, if the second-difference estimator per equation Figure 34 were used, how much worse than optimum would it be? The uncertainty is given in Figure 35, and it is only 15% worse than optimum. However, in the case of random-walk FM, the three-point estimator is optimum and the linear regression is worse by some similar factor.

In telecommunications, very often the frequency drift of quartz oscillators as it affects the time-interval-error (TIE) is an important specification. Figure 36 gives a relationship between $\sigma_x(\tau) = $ TDEV, the frequency drift, D, and the corresponding TIE.

Figures 37 and 38 show the effect of modulation on $\sigma_y(\tau)$ and on $\sigma_x(\tau)$. In the latter case, a background noise of white PM is also included in the simulation. Notice that the effect of the modulation averages down as $1/\tau$.

If there is a need to estimate an effect due to temperature, pressure, humidity, etc., then the following procedure will be helpful. Suppose the clock has a $\sigma_y(\tau)$ diagram something like that shown in Figure 39. Denote τ_{floor} as the averaging time where the clock reaches its flicker floor. Now average the frequency for this length of time with the parameter in question fixed at some value. Switch the parameter to some new value, allowing for settling, and measure the frequency again for an interval τ_{floor}. Switch the parameter back to its original setting, again

allowing for settling, and measure the average frequency for the third time. Keep repeating the switching of the parameter setting as often as needed to get the uncertainty desired. In principle, the uncertainty in the size of the effect of this particular parameter on the frequency will decrease as 1 over the square root of the number of the independent switches. In this way, we are not limited by the flicker floor, and can determine the size of the effect arbitrarily well.

Now consider optimum procedures in using some of the clocks contributing to TAI/UTC. Figure 40 is a stability plot for 78 of the HP 5071As contributing to TAI/UTC during 1994. Clearly, there is not a single representative model for all of these clocks. The best possible stability obtainable from these clocks is given by the equation at the bottom right of Figure 41 and represented by the 'x's. These results are reasonably modeled by the equation given: $\sigma_y(\tau) = 8.7 \times 10^{-13}\tau^{-1/2}$, where τ is in units of seconds. The dots are the estimated stabilities under the assumption that the 78 clocks are all equal; square root of the average variance divided by the square root of the number of clocks. These two stability plots give an upper and a lower bound to the actual stability one could obtain using these clocks. The circles are the composite stabilities for the hydrogen masers used in TAI/UTC for this same period.

The above are only theoretical estimates since there is no clock good enough with which to measure this level of stability. In an effort to estimate the actual stability, a three-cornered hat experiment was performed between a time scale generated for each of three clock sets: the primary standards running as clocks, the hydrogen masers contributing to TAI/UTC, and 78 HP 5071A cesium clocks analyzed in Figure 40. The two plus '+' points were the resulting estimated stability for the HP 5071A clocks. The $\tau^{-3/2}$ slope would indicate we are only seeing measurement noise and are not limited by the clocks for these τ values.

Figure 42 is the time stability, $\sigma_x(\tau)$, and Mod.$\sigma_y(\tau)$ for several important time and/or frequency transfer techniques. Both can be plotted on the same graph since $\sigma_x(\tau) = \tau$ Mod.$\sigma_y(\tau)/\sqrt{3}$. Most of the plots are for state-of-the-art techniques except for Loran-C, which is plotted for comparison purposes.

The Two-way Satellite Time Transfer Technique has excellent short-term stability, but due to equipment delay variations to date it only reaches somewhat better than 1×10^{-13} before these variations significantly contaminate the estimation process. In the very long-term these instabilities start to average down again.

The enhanced GPS, EGPS, technique was developed to utilize the new multichannel GPS receivers and to over come to some degree the effects of the Selective Availability (SA) degradation present for the civil users of GPS. The degree to which the SA can be filtered away is a function of the quality of clock used with the multi-channel GPS receiver; i.e. quartz, rubidium or cesium. For example, if a very good quartz oscillator is properly used and servoed to GPS and the SA is optimally filtered, then the short-term stability will be that of the quartz oscillator, which is usually excellent, and the very long-term will be that of GPS. The intermediate-term stability will depend on the intermediate-term stability of the quartz oscillator, which is not as good as a rubidium gas-cell frequency standard; and the rubidium in turn is not as good as a cesium-beam frequency standard for the intermediate term.

The GPS carrier phase technique has outstanding frequency transfer capability — reaching

219

about 2×10^{-15} in 10^5 seconds (about one day). The data plotted here came from a comparison of hydrogen masers located in Goldstone, California and in Algonquin Park, Canada. The baseline distance is about 3.2 Mm (2,000 miles); the circumference of the Earth is about 40 Mm. Some 35 tracking stations were used to determine accurate orbits. Notice that the classical measurement noise only persists for about five minutes, then some random-walk errors start to come in. Notice also that the time instability averages down to below 10 ps; that is the time it takes a light signal to travel 3 millimeters! Clearly, Earth tides had to be included in this analysis. One also sees the power of these kinds of measurements to study plate tectonics for the Earth.

The GPS common view (GPS CV), which has been used since 1981 and still is the best operational means of comparing time and frequency standards remote from each other, starts at $\tau = 1$ day and integrates down to below the 10^{-14} level. This is the main means of time and frequency transfer for the clocks and frequency standards providing input into TAI/UTC.

If we go back to the second equation in Figure 28 — remembering that the confidence of the estimate of frequency improves as the degrees of freedom N to the 3/2s power — one can think of some very exciting opportunities with the new multi-channel GPS receivers. These receivers are able to take one second data. If the measurement noise is white PM at a level of 8 ns and four satellites could always be tracked in common-view with another site, then the frequency transfer uncertainty would be 1.4×10^{-16} for a one day's regression analysis. This technique is called the Advanced Common-view approach (GPS ACV).

Two eight-channel receivers were tested with common clock and common antenna to study instrumentation noise and to check the theory of the above paragraph. The results are plotted in Figure 42. A complex digital filter was developed to take advantage of all the degrees of freedom while increasing the averaging time to ten seconds in order to reduce the data rate. This digital filter explains the little hump at about 30 seconds. The curve generally follows the white-noise PM power-law spectral model with the data averaging as $\tau^{-1/2}$ down to a level of about 70 picoseconds. This data is taken from the accompanying paper[1], and is thanks to Dr. Robin Giffard. We next need to study the performance with separate antennas, as a function of temperature, and with the receivers located at sites remote to each other. The effects of the ionosphere, the troposphere and multipath can be measured and/or averaged and can be driven below the nanosecond level. Much work is yet to be done, but the GPS ACV technique appears to have the potential to be very practical and useful.

Notice the effects of diurnal and annual variations in the Loran-C stability. As better and better standards are being compared, it may be that in some cases temperature control will be necessary to avoid such variations as they may occur in other techniques as well. As we move time and frequency metrology forward, it is always well to keep in mind the basic concepts and methodologies for parameter estimation and prediction. Those presented here are not a complete set, but it is hoped that they will be useful to those interested in utilizing the powerful time and frequency resources and tools we now have available within our community.

References

[1] Our PTTI paper

[2] ITU-R $\sigma_y(\tau)$ Sydnor performance plot

[3] NIST TN 1337

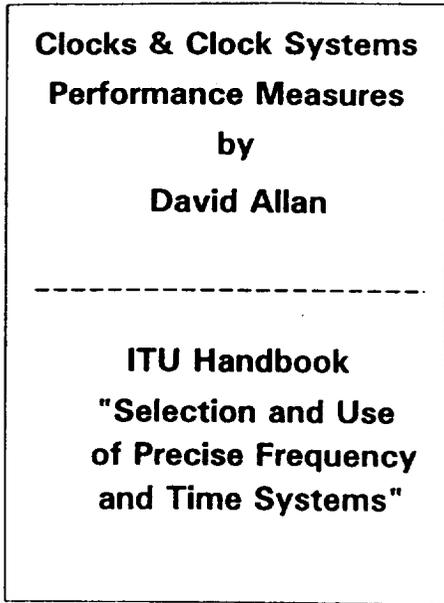[4] Allan, Weiss & Jespersen

[5] Mattison PTTI paper, '95

Clocks & Clock Systems

Performance Measures

by

David Allan

ITU Handbook

"Selection and Use of Precise Frequency and Time Systems"

Figure 1

Q: What Is A Clock?

counts the number of oscillations

frequency source

Cs - 133 ATOM

9 192 631 770 Hz

Fast Electronic Counter

Figure 2

A: an oscillator + a counter

Which is Correct?

Every clock disagrees with every other clock except at most an instant.

Disagreement is only a matter of degree.

hr | min | s | ms | μs | ns | ps | fs

Figure 3

Frequency VS. Time

Frequency

(from an atomic resonance) IS ABSOLUTE; hence is the basis of the definition of "the second."

Time

(clock reading) is an artifact of man; we define it to be what we want it to be. IT IS NOT ABSOLUTE.

Figure 4

Time is an artifact

Perceived Causes Of Clock Deviations

- Perception
- Measurement Noise
- Internal
- Environment

Environmental + Perturbations

Random And + Systematic Within Clock

= Measurement Noise

Perceived Performance

Figure 5

Clock Concepts

- **Time (in practice)**
  the apparent reading of a clock
- **Synchronization**
  two or more clocks have the same apparent time
- **Syntonization**
  two or more clocks have the same apparent rate
- **Simultaneity**
  two or more events occur at the same moment
  (does not require a clock)

Time is a definition

Figure 6

Model Elements Dependent On Environment

Operator Interference
Shock and Vibration
Supply Voltage
Magnetic Field
Temperature
Pressure
Humidity
Load
Etc.

All clocks affected by environment

Figure 7

IEEE Std 1193-1994

IEEE Guide for Measurement of Environmental Sensitivities of Standard Frequency Generators

Abstract: Standard frequency generators that include all atomic frequency standards and precision quartz crystal oscillators are addressed.
Keywords: environmental sensitivities, standard frequency generators

Figure 8

223

Figure 9

$$V(t) = V_0 \sin \Phi(t)$$
$$= V_0 \sin 2\pi\nu(t)t$$
$$= V_0 \sin(2\pi\nu_0 t + \varphi(t))$$

$$\nu(t) = \frac{1}{\tau(t)}$$
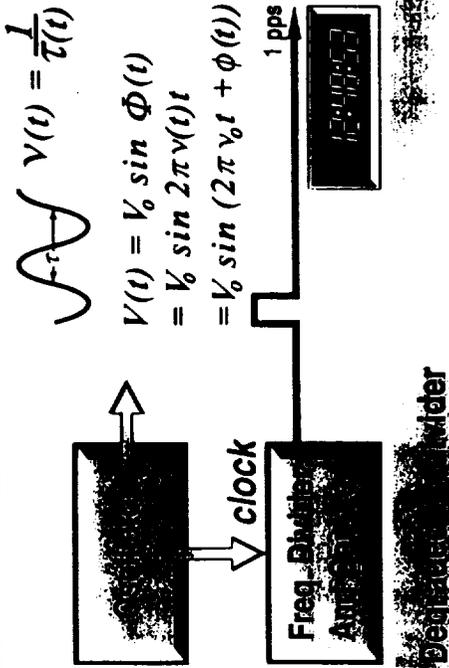
1 pps

clock

Freq. Divider

Figure 9

That Which Is Measured

We cannot measure the time of a clock against absolute time because absolute time does not exist.

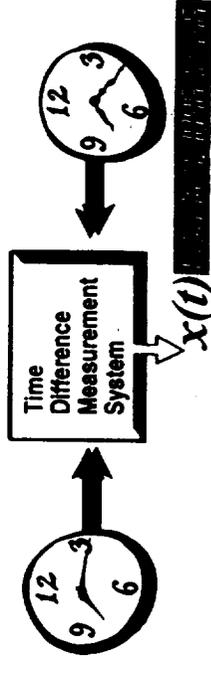We can measure the time difference between two clocks with great precision.

Time Difference Measurement System

$$x(t)$$

Time is not absolute

Figure 10

Define: Normalized Freq. & Time Residual

Dimensionless Normalized Freq.

Time Residual

$$y(t) = \frac{\nu(t) - \nu_o}{\nu_o}$$

$$x(t) = \frac{\varphi(t)}{2\pi\nu_o}$$

Then $$y(t) = \frac{dx(t)}{dt} \equiv \dot{x}(t)$$

And $$x(t) = \int_0^t y(t')\, dt'$$

Measures of departure

Figure 11

$$x(t) = x_o + y_o t + \tfrac{1}{2}D t^2 + \epsilon(t)$$

All the rest!
Frequency Drift
Syntonization Error (t=0)
Synchronization error (t=0)

Subtracting first three terms from data, then $x(t) = \epsilon(t)$

Figure 12

224

Figure 13

Time - Domain Measures

- **Frequency Accuracy:**
  The degree of conformity with a standard or a definition.

- **Frequency Instability:**
  Change, typically averaged over an interval, τ, with respect to another frequency.

- **Time Accuracy:**
  The degree of conformity with UTC or some agreed upon time-scale.

- **Time Instability:**
  Change, in residual readings, typically averaged over an interval, τ, with respect to nominal or other averaged interval(s).

$$y(t) \quad x(t) \quad \sigma_y(\tau) \quad \sigma_x(\tau) \quad Mod.\sigma_y(\tau) \quad UTC$$
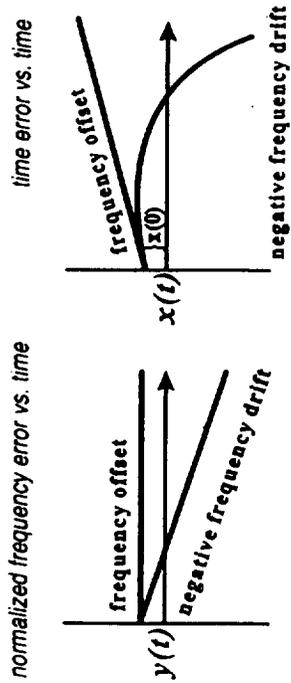
Figure 15

IEEE Recommended Measures

for clock/oscillator
$$\left\{ \begin{array}{l} S_y(f), \ S_\varphi(f) \\ \sigma_y(\tau), \ Mod. \ \sigma_y(\tau) \end{array} \right.$$

for measurement system or network
$$\left\{ \begin{array}{l} S_x(f), \ S_\varphi(f) \\ \sigma_x(\tau), \ Mod.\sigma_y(\tau) \end{array} \right.$$

$$\sigma_y(\tau), \sigma_x(\tau) \text{ optimum for classical noises}$$

Figure 14

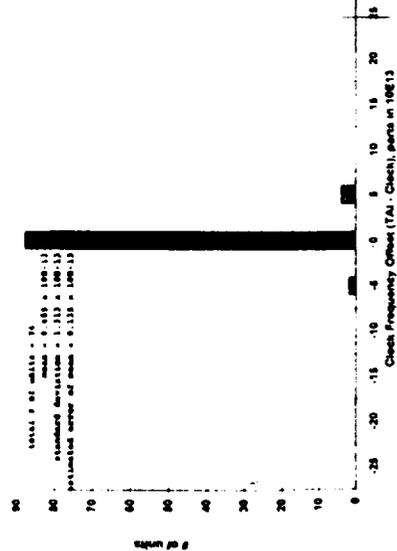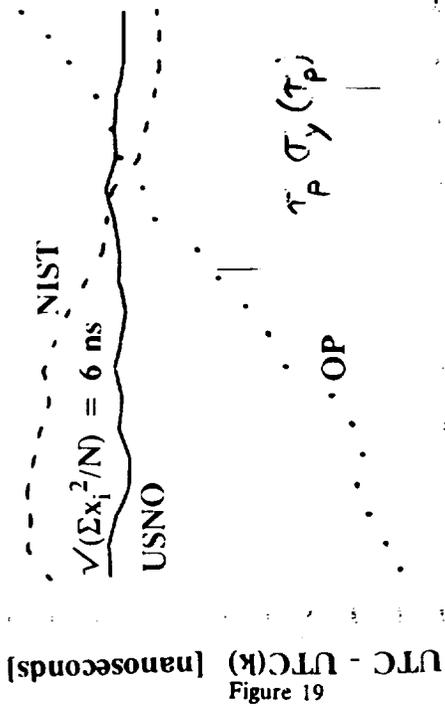HISTOGRAM OF FREQUENCY OFFSETS (TAI - CLOCK) FOR 1994
(ALL 6071Aa)

Figure 16

Figure 17

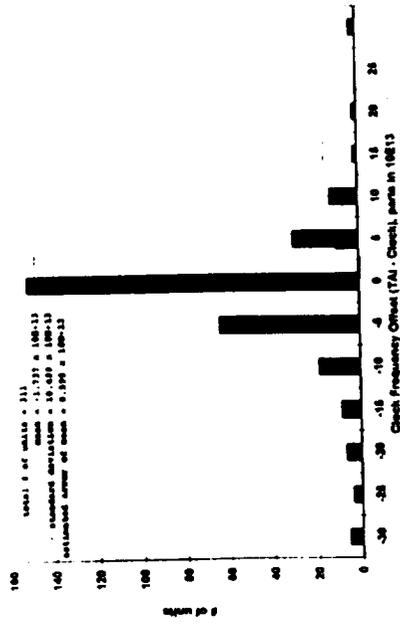**HISTOGRAM OF FREQUENCY OFFSETS (TAI - CLOCK) FOR 1994**
(ALL CLOCKS)

Clock Frequency Offset (TAI - Clock), parts in 10E13

Figure 18

$\log (\sigma_y(\tau))$

Log (τ), seconds

FIGURE 2
Stability Ranges of Various Frequency Standards

CS

RB

Cs

HM

TIME ACCURACY

NIST

$\frac{\sqrt{\Sigma x_i^2/N}}{} = 6$ ns

USNO

OP

$\tau_p \; \sigma_y (\tau_p)$

MJD - 49819

UTC - UTC(k) [nanoseconds]

Figure 19

Flip Coins For White Noise

Flip of a coin is a random, uncorrelated process:
(white noise spectrum)

$S_x(f) \sim f^0$

Integrating these flips generates a random-walk process:
heads = one step forward
tails = one step backward

$S_x(f) \sim f^{-2}$

After N flips of a coin, will be √N away from the origin
(ensemble rms)

$S_x(f) \sim f^{-1}$

Since $y = dx/dt$,

Taking a first difference of a random-walk process turns it
into a white process.

$y \sim \Delta t$

Figure 20

...and random walks

226

for clocks: (oscillators) white FM

$$\sigma_y(\tau) = 8 \times 10^{-13} \tau^{-1/2}$$

difference

for measurement: white PM

$$x(t) = x_o + y_o t + \tfrac{1}{2}Dt^2 + \epsilon(t)$$

Model for time difference

Figure 21

Time Deviations

α	β	
+2	0	White PM
+1	-1	Flicker PM
0	-2	R.W. PM or White FM
-1	-3	Flicker FM
-2	-4	R.W. FM

$\sim f^\beta$

Figure 22

α	Noise Type	Range of Applicability
+2	White PM	
+1	Flicker PM	
0	White FM	
-1	Flicker FM	
-2	Random Walk FM	

Quartz
Active Hydrogen
Passive Hydrogen
Rubidium
Cesium
Time/Frequency Dissemination Systems
Tel. Comm. Network

AVAR $\alpha = -\mu - 1$
MVAR $\alpha = -\mu' - 1$
TVAR $\beta = -\eta - 1$

$$\sigma_y^2(\tau) \sim \tau^\mu \quad Mod.\sigma_y^2(\tau) \sim \tau^{\mu'}$$
$$\sigma_x^2(\tau) \sim \tau^\eta \quad \eta = \mu' + 2$$

Figure 23

ABBREVIATION	NAME	EXPRESSION
AVAR	ALLAN VARIANCE	$\sigma_y^2(\tau) = \frac{1}{2}\langle(\Delta_y)^2\rangle$
MVAR	MODIFIED ALLAN VARIANCE	$Mod\sigma_y^2(\tau) = \frac{1}{2\tau^2}\langle(\Delta^2 x)^2\rangle$
TVAR	TIME VARIANCE	$\sigma_x^2(\tau) = \frac{\tau^2}{6}\langle(\Delta^2 x)^2\rangle$

Figure 24

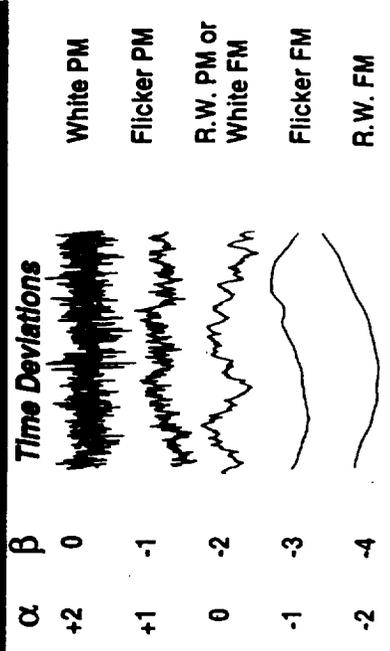## Optimum Estimation & Prediction

▼ Optimum means parameter estimation with
  ● *minimum* squared error residuals

  Goal: to minimize errors for parameter estimation
  and/or for time and frequency prediction.

Depends on noise type

Figure 25

Statistical Theorem:

*The optimum estimate of the mean of a process with
a white-noise spectrum is the simple mean.*

Naturally we have:
  White PM, White FM,
  R.W. FM (white acceleration, $\dddot{x}(t)$)

Flicker models are also very common.

Seek for white residuals

Figure 26

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

$$\bar{y} = \frac{x(t+\tau) - x(t)}{\tau}$$

is optimum for
white FM

$$\sigma_x^2(\tau) = \frac{1}{6}\left\langle (\Delta^2 \bar{x})^2 \right\rangle$$

$\bar{x}$ = avg. of x over τ
is optimum for
white PM

Figure 27

White PM

Uncertainty	
time: $\sigma_t(\tau,\ldots)$	
freq: $\frac{\sqrt{12}\,\sigma(\tau,\ldots)}{\tau_o}$	
drift: $\frac{12\sqrt{5}\,\sigma(\tau,\ldots)}{\tau_o^2}$	

$x(t) = x_o + \mathcal{E}(t)$
simple mean

$x(t) = x_o + y_o t + \mathcal{E}(t)$
linear regression

$x(t) = x_o + y_o t + \frac{1}{2}Dt^2 + \mathcal{E}(t)$
quadratic regression

Optimum estimate

Figure 28

Optimum Estimate For White FM

Given: Discrete $x(i)$ values spaced τ_o from a time difference series

Then $\quad y(i) = \dfrac{x(i) - x(i-1)}{\tau_o} = \dfrac{\Delta x(i)}{\tau_o}$

Average Frequency $\quad \bar{y} = \dfrac{1}{N}\sum_{i=1}^{N} y(i)$ Simple Mean

$\quad = \dfrac{x(N) - x(0)}{N\tau_o}$ Optimum Estimate for white FM

Optimum based on simple mean

Figure 29

GIVEN: CLASSICAL WHITE NOISE FM
(as from Cs or Rb)
THIS IS THE SAME AS RANDOM WALK PM
(100 simulations of 100 points)
Difference between end-point and regression for determining the frequency and time

	Mean Freq.	σ_y	Mean x_o
REGRESSION	-0.0005	0.0970	0.44
END POINTS	-0.00029	0.0894	0

Figure 30

White FM

Uncertainty

$x(t) = x_o + y_o t + \epsilon(t)$
$y(t) = y_o + \dot{\epsilon}(t)$
..mple mean

$x(t) = x_o + y_o t + \tfrac{1}{2}Dt^2 + \epsilon(t)$
$y(t) = y_o + Dt + \dot{\epsilon}(t)$
linear regression

freq:
$\sigma_y(\tau_o)/\sqrt{N}$

drift:
$\dfrac{\sqrt{12}\,\sigma_y(\tau_o)}{\tau_o\, N^{3/2}}$

Optimum estimate

Figure 31

Random Walk FM

Uncertainty

$x(t) = x_o + y_o t + \tfrac{1}{2}Dt^2 + \epsilon(t)$
$y(t) = y_o + Dt + \dot{\epsilon}(t)$
$z(t) = D + \ddot{\epsilon}(t)$
simple mean
$D = \dfrac{\overline{\Delta^2 x}}{\tau_o^2}$

drift:
$\dfrac{\sqrt{2}\,\sigma_Y(\tau_o)}{\tau_o\sqrt{N-1}}$

Optimum estimate

Figure 32

Life Is Not Simple

Any given time series may be modeled by more than *one* Noise Model

There will always be a filter which will produce white residuals

Optimum may be complex filter

Figure 33

Some Multiple Noise Type Examples

White PM and White FM and/or R.W. FM
Average white PM
Then analyze residuals

White FM and R.W. FM
Average white FM
Then analyze residuals

Optimum Freq. DRIFT estimation in both cases

$$D = \frac{x_N - 2\,x_{N/2} + x_o}{\tau^2} \;,\quad \tau = \frac{N}{2}\,\tau_o$$

Effective estimate for most clocks

Figure 34

Given White FM & Frequency Drift

Uncertainty Drift:

$$\frac{\sqrt{12}\;\sigma_y(\tau_s)}{\tau_s N}$$

$$\frac{4\sigma_y(\tau_s)}{\tau_s N}$$

$$\frac{4}{\sqrt{6}} = 1.15$$

linear regression to:

$$y(n) = y_o + Dn\tau_o$$

or using 3-point $\Delta^2 x$:

$$y(n)$$

$\bar{y}(1:\tfrac{N}{2})$
$\bar{y}(\tfrac{N}{2}:N)$

Optimum and near optimum

Figure 35

TIME INTERVAL ERROR due to DRIFT

Given:
$$x(t) = 1/2\; D\; t^2$$

Then:
$$TIE\,\big|_{drift} = \sqrt{6}/2\;\sigma_x(\tau)$$

$$= 1.2\; TDEV$$

Figure 36

230

$S_y(f) = 2.5 \times 10^{-15} (1 + 10^3 \delta(f - 60 \, Hz))$

$$\sigma_y(\tau) = \frac{x_{pp}}{\tau} \sin^2 (\pi f_m \tau)$$
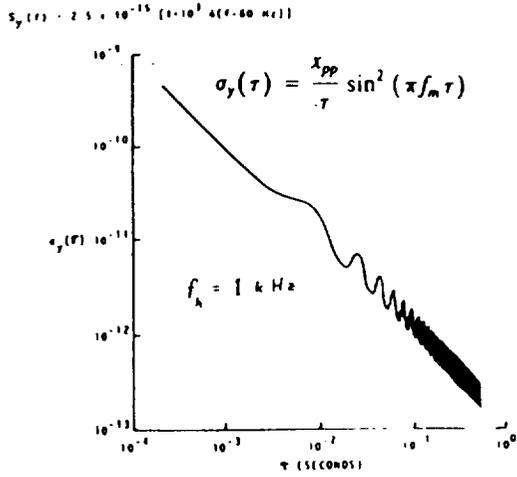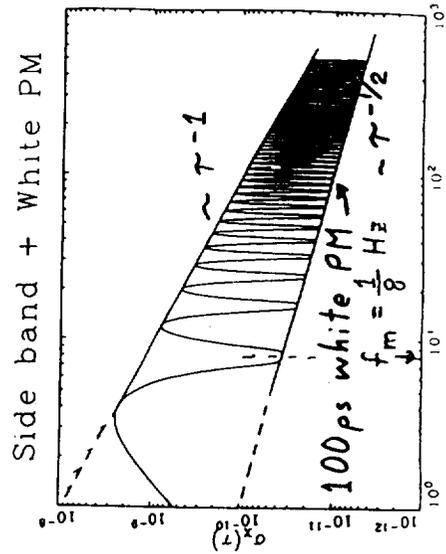
$f_h = 1 \, kHz$
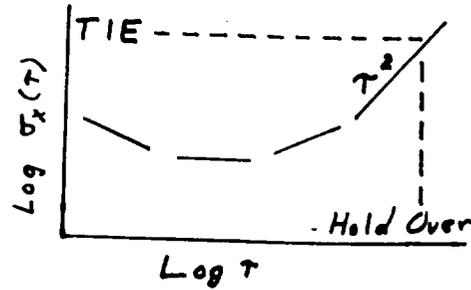
Figure 37

Figure 38

Given White FM and flicker-floor:

To Measure an environmental steady-state effect, hold the parameter constant and average the frequency for an interval such that the $\sigma_y(\tau)$ curve starts changing from $\tau^{-1/2}$ toward a flattening (flicker floor)   Then change the environmental parameter being evaluated and repeat the integration time to measure the frequency change. If "N" is the number of changes back and forth, then the confidence on the frequency change is the value of $\sigma_y(\tau_{floor})$ times $1/\sqrt{N}$
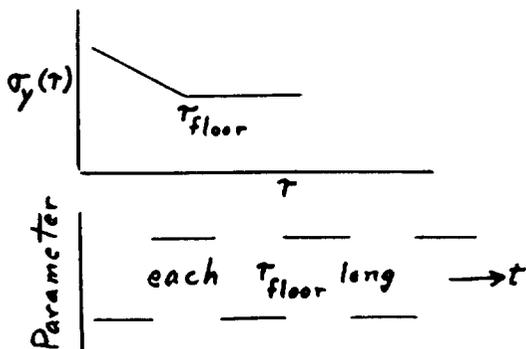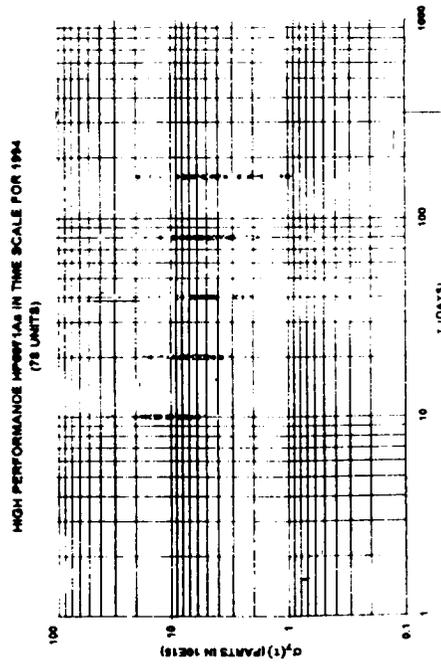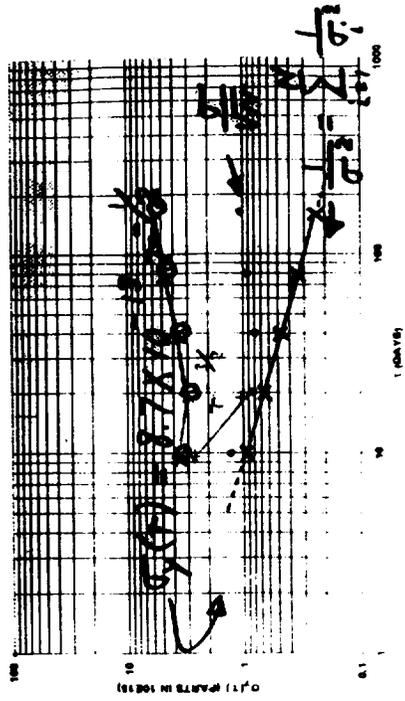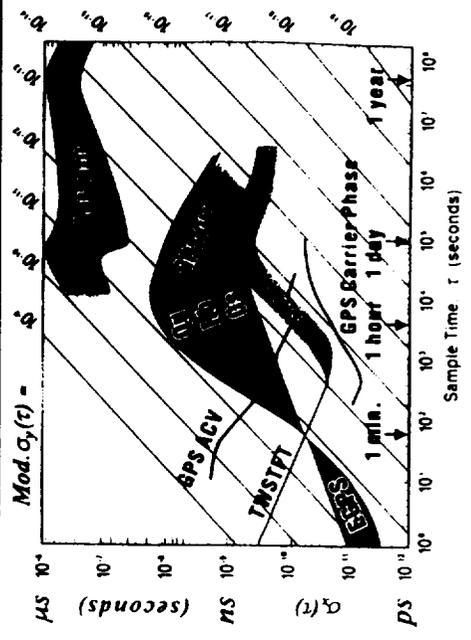

Figure 39

Figure 40

231

Figure 41

Figure 42

232

Questions and Answers

MICHAEL GARVEY (FREQUENCY AND TIME SYSTEMS): You showed the slide in which you were trying to pull an environmental sensitivity out of the noise; and you said "Wait until you hit the flicker floor." Is there any reason not to modulate the environmental effect at a faster rate?

DAVID ALLAN (ALLAN'S TIME): You have to wait for settling, so that increases the amount of time it takes to do the experiment.

MICHAEL GARVEY (FREQUENCY AND TIME SYSTEMS): I know. But if I wait for the flicker floor in a cesium standard, I might wait weeks.

DAVID ALLAN (ALLAN'S TIME): Yes, and if you can't hold the environmental parameter stable for that long, you should change it more often. For cesium and rubidium clocks the frequency averages as $\tau^{-1/2}$. If you can hold the parameter constant, then you're much better off to let the clock do the averaging because of the delay associated with settling for each switching time. If the parameter can be held sufficiently constant, then average all the way down to the flicker floor.

MICHAEL GARVEY (FREQUENCY AND TIME SYSTEMS): Rather than wait for the square root of N in the denominator.

DAVID ALLAN (ALLAN'S TIME): Yes, you buy information at the same rate except for the settling time. It is a trade-off between the parameter's instability and the length of the settling time. In the case of humidity effects in quartz, for example, the settling time can be very long.