A MULTI-VARIANCE ANALYSIS IN THE TIME DOMAIN

Todd Walter

W.W. Hansen Experimental Physics Laboratory
Stanford University, Stanford, CA   94305-4085

Abstract

Recently a new technique for characterizing the noise processes affecting oscillators was introduced [1] [2]. This technique minimizes the difference between the estimates of several different variances and their values as predicted by the standard power law model of noise. The method outlined in this paper makes two significant advancements: it uses exclusively time domain variances so that deterministic parameters such as linear frequency drift may be estimated, and it correctly fits the estimates using the chi-square distribution. These changes permit a more accurate fitting at long time intervals where there is the least information. This technique has been applied to both simulated and real data with excellent results.

I. Introduction

Stochastic noise processes are the dominant source of imprecision in high-performance oscillators. Better information about these noise sources leads to an improved understanding of the estimated stability of the oscillator. Furthermore, information about the level of contribution from each noise type can improve their theoretical description. Some of these processes are well understood (thermal noise, shot noise, etc.), but many are not adequately described by theory. For some processes, when they are understood, there is greater potential for their subsequent reduction. Rudimentary stability analysis of an oscillator is fairly straight-forward. However, placing confidence limits on the stability estimates requires detailed knowledge of the noise types affecting the precision. The processes being characterized are stochastic, and therefore we can only provide a rough statistical analysis. Additionally there may be many contributing noise sources that can have a tendency to obscure one another. Thus there has been a scarcity of good data on the precise contributions of individual noise sources. Recently a new method in noise analysis was introduced. This method, called multi-variance analysis [1], is capable of providing precise measurements of dominant noise sources.

Before examining the details of this new method, we must first understand how noise affecting oscillator precision is measured. The output of an oscillator can be represented by [3] [4]

$$V(t) = [V_0 + \varepsilon(t)] \sin[2\pi v_0 t + \varphi(t)] + V_I(t)$$

V_0 and v_0 are the respective nominal amplitude and frequency of the output, $\varepsilon(t)$ and $\varphi(t)$ are amplitude and phase fluctuations respectively and $V_I(t)$ is additive noise. Provided ε and V_I are much smaller than V_0, the instantaneous frequency of the oscillator output can be written

$$v(t) = v_0 + \frac{1}{2\pi} \frac{d\varphi(t)}{dt}$$

The instantaneous fractional frequency deviation from nominal may also be defined

$$y(t) \equiv \frac{v(t) - v_0}{v_0} = \frac{1}{2\pi v_0} \frac{d\varphi(t)}{dt}$$

It is the stability of the output frequency that is of primary concern. Stochastic processes that affect this stability will appear in y(t). Another useful quantity is the phase deviation, in units of time,

$$x(t) \equiv \frac{\varphi(t)}{2\pi v_0}$$

This quantity, the time integral of $y(t)$, is a measure of the time deviations of the oscillator. Averaged values of $y(t)$ can be obtained by differencing two phase measurements and dividing the result by the time interval between the measurements.

The effects of these noise processes manifest themselves in both $x(t)$ and $y(t)$. This paper is concerned with the characterization of these effects rather than their physical cause. Examining the disturbances to $y(t)$ is one of the more common means of characterizing the frequency stability of an oscillator. One method is to look at the power spectral density (PSD) of $y(t)$. It has been observed that the PSD of $y(t)$ often has integer slopes when plotted on a log-log graph. Empirically, five slopes are commonly observed. This has led to the standard power law model which may be written [3] [4]

$$S_y(f) = \sum_\alpha h_\alpha f^\alpha$$

where α is an integer that runs from -2 to 2 and the h_α's are the noise intensity coefficients. The five noise categories are respectively white phase modulation, flicker phase modulation, white frequency modulation, flicker frequency modulation and random walk frequency modulation.

This model adequately describes most of the observed noise processes. However the frequency domain is not necessarily the best place for analysis. Forming a PSD estimate from data discretely sampled in the time domain can lead to biases and distortions. Additionally the data is often affected by systematic effects such as frequency offset and linear frequency drift that, if not properly estimated and removed, will also distort the PSD estimate.

The process $y(t)$ is not the most convenient measure because it is not possible to measure the instantaneous frequency. Instead the frequency measurement takes place over a finite time interval τ. Also the measurement of $y(t)$ often involves some dead-time. This causes a reduction in the amount of information obtained. Therefore, the frequency stability is more easily specified through the characterization of $x(t)$ in the time domain.

II. Time Domain Variances

The most common time domain measure of oscillator stability is the Allan (or two-sample) variance. For the process $x(t)$ it is defined by

$$\sigma_y^2(\tau) \equiv \frac{1}{2\tau^2}\left\langle [x(t+2\tau) - 2x(t+\tau) + x(t)]^2 \right\rangle$$

The angle brackets denote an ensemble average or expected value. The Allan variance was chosen because it forms a convergent measure of the fractional oscillator stability as a function of time interval. It is possible to define other variances that meet this criterion. A less familiar measure is the Hadamard variance with binomial coefficients [5]-[8]. I will use a renormalized version given by

$$_H\sigma_y^2(\tau) \equiv \frac{1}{6\tau^2}\left\langle [x(t+3\tau) - 3x(t+2\tau) + 3x(t+\tau) - x(t)]^2 \right\rangle$$

This measure is convergent for $\alpha > -5$, unlike the Allan variance which is convergent for $\alpha > -3$. Thus it would be possible to use the Hadamard variance to probe for noise beyond random walk frequency modulation. Perhaps the most important feature of the Hadamard variance is that it is unaffected by linear frequency drift. This makes it an excellent tool for investigating noise types whose signatures are similar to and often confused with linear drift.

A new variance is introduced, which I call the alternate difference variance. It is defined by

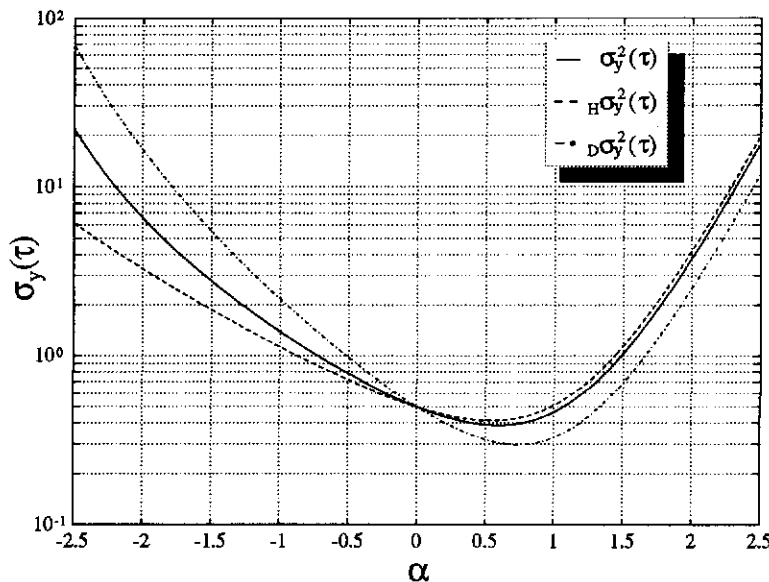$$_D\sigma_y^2(\tau) \equiv \frac{1}{2\tau^2}\left\langle [x(t+3\tau) - x(t+2\tau) - x(t+\tau) + x(t)]^2 \right\rangle$$

Its chief advantage is that it is affected to a greater degree by noise with α below 0 although it too is only convergent for $\alpha > -3$.

Table 1 shows the functional dependence of each variance on the different noise types. In addition Figure 1 plots these three variances as functions of α. Notice that for $\alpha > 0$ all three variances have similar responses, although the Hadamard variance is slightly above and the alternate difference variance is below the Allan variance. At $\alpha = 0$ the variances are equal by definition. For $\alpha < 0$ the three begin to diverge.

Figure 1. The Allan, Hadamard and alternate difference variances are plotted as functions of α ($h_\alpha = 1$, $\tau = 1$)

All three of these variances have difficulty distinguishing between white phase and flicker phase noise. To aid in this resolution the modified Allan variance was created [4]. It exploits the different dependencies of these two noise types on system bandwidth (f_h). Just as the modified Allan variance was created from the Allan variance it is possible to create the modified Hadamard variance and the modified alternate difference variance from their respective variances.

The time domain variances, as for the PSD, can only be estimated from the observed data. The noise processes will each have some underlying true variance that is unknown to the observer. We use the discretely sampled data of finite length to estimate this true variance. If we have N points each separated in time by τ_0, so that $x_k = x(k\tau_0 + t_0)$, the estimate of the Allan variance is given by

$$\hat{\sigma}_y^2(m\tau_0, N) = \frac{1}{2(N-2m)(m\tau_0)^2} \sum_{k=1}^{N-2m} (x_{k+2m} - 2x_{k+m} + x_k)^2$$

The $^\wedge$ specifically denotes the fact that this is only an estimate. This estimate has a specific uncertainty associated with it. Clearly as m approaches N fewer points will be included in the estimate and it will have greater uncertainty. For each specific noise type it is possible to calculate the variance of the variance estimate [4] [9] [10]. Similar estimates and uncertainties of estimates can be calculated for each of the other variances.

Noise Type	Allan Variance	Hadamard Variance	Alternate Difference
White Phase $\left(S_y(f) = h_2 f^2\right)$	$\dfrac{3}{(2\pi)^2 \tau^2} h_2 f_h$	$\dfrac{10}{3(2\pi)^2 \tau^2} h_2 f_h$	$\dfrac{2}{(2\pi)^2 \tau^2} h_2 f_h$
Flicker Phase $\left(S_y(f) = h_1 f\right)$	$\dfrac{3\gamma - \ell n2 + 3\ell n(2\pi f_h \tau)}{(2\pi)^2 \tau^2} h_1$	$\dfrac{10\gamma - 6\ell n2 + 15\ell n3 + 10\ell n(2\pi f_h \tau)}{3(2\pi)^2 \tau^2} h_1$	$\dfrac{2\gamma + 2\ell n2 - \ell n3 + 2\ell n(2\pi f_h \tau)}{(2\pi)^2 \tau^2} h_1$
White Frequency $\left(S_y(f) = h_0\right)$	$\dfrac{h_0}{2\tau}$	$\dfrac{h_0}{2\tau}$	$\dfrac{h_0}{2\tau}$
Flicker Frequency $\left(S_y(f) = h_{-1} f^{-1}\right)$	$2\ell n2 h_{-1}$	$(4\ell n2 - \tfrac{3}{2}\ell n3)h_{-1}$	$(\tfrac{9}{2}\ell n3 - 4\ell n2)h_{-1}$
Random Walk Freq. $\left(S_y(f) = h_{-2} f^{-2}\right)$	$\dfrac{2\pi^2 \tau}{3} h_{-2}$	$\dfrac{\pi^2 \tau}{3} h_{-2}$	$\dfrac{5\pi^2 \tau}{3} h_{-2}$
Linear Frequency Drift $\left(x(t) = \tfrac{1}{2}Dr\, t^2\right)$	$\dfrac{Dr^2}{2} \tau^2$	0	$2Dr^2 \tau^2$

Table 1. The functional dependencies of the three variances under the assumption $f_h \tau \gg 1$

III. Multi-Variance Analysis

The multi-variance method combines the power law model with the output of several variance estimates. Thus many observations with different variances must all agree within the predicted responses of the power law model. This greatly simplifies the analysis, as all of these observations are used to estimate the five noise intensity coefficients. A single-variance technique separates noise contributions by their differing dependencies on τ. The multi-variance method exploits those τ dependencies in addition to utilizing different responses of each variance for each particular value of τ. Thus the multi-variance method gains more resolution over analysis with a single variance. This powerful new technique was introduced recently by Vernotte et al. [1] [2].

The estimates used in previous work [1] [2], correspond both to the time domain (Allan and modified Allan variances) and to the frequency domain (Band-pass and High-pass variances) [6]. The frequency domain variances are a powerful analytical tool, but, as previously mentioned, they are more susceptible to biases and distortions. Systematic effects such as frequency offset and drift, if not properly accounted for, will also bias the spectral density estimates. We are interested in finding both the frequency offset and linear drift in our analysis of the time series. In order to correctly form the PSD estimate these effects *must be subtracted from the time series.* Often a least squares fit is performed on the data to estimate these parameters. Unfortunately, the noise is non-white. Non-white noise also has linear and quadratic components which will yield biased estimates and incorrect confidence limits on those estimates [4]. When these false estimates are used to detrend the data, some of the noise contribution will be subtracted out as well. A better approach would be to incorporate these systematic effects in the fitting process, or to estimate the parameters after the fitting process, so that the noise types will be known and can be correctly taken into account.

I have chosen to implement the multi-variance technique using exclusively time domain variances. The five variances used in this analysis are the Allan and modified Allan variances, the Hadamard and modified Hadamard variances, and the alternate difference variance. This change leads to a more robust estimator and allows systematic errors such as linear frequency drift to be solved for as part of the fit. The frequency offset is estimated after performing the fit when we have better knowledge of the noise shape. In the fit routine I present here, I use the five variances at different values of τ, to fit six parameters: the five noise intensity coefficients and linear frequency drift.

IV. Chi-Square Probability Distribution

Standard least squares fit methods return the maximum likelihood solution for estimates that are distributed normally. However, the variance estimates used in oscillator noise analysis follow a chi-square distribution. Therefore fit routines using the standard least squares method will not yield the best solution, particularly when the estimates have few degrees of freedom. Figure 2 shows the chi-square distribution for two different values of the number of degrees of freedom (ν). It is evident from this figure that for low degrees of freedom the distribution is quite different from a normal distribution.

The mechanics of performing a fit on chi-square distributed variables are similar to fitting normally distributed variables. One possible source of confusion is that least-squares fitting of data with normally distributed noise is often referred to as chi-square fitting. This is because the cost function (parameter to be extremized) is chi-square distributed (the sum of the squares of normally distributed variables). I define a new fitting routine, for chi-square
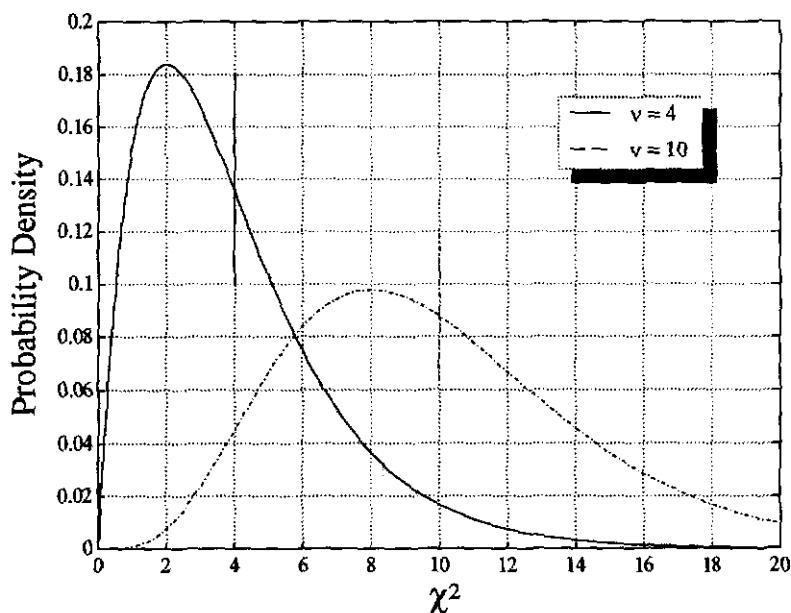
Figure 2. The chi-square probability distribution is plotted for two different values of the number of degrees of freedom (v). The vertical lines indicate the mean value or 50% point

distributed variables, in which the cost function will have a different distribution. I have termed this type of fit *xi-square fitting*.

The variance estimate contains both stochastic and deterministic effects. It is the stochastic contribution which follows the chi-square distribution. We must subtract out the systematic effects and then properly scale the estimate. The subtraction of systematic components takes place as part of the fit. They do not have to be estimated and subtracted before the variance estimates can be formed, as in the case of the frequency domain measures. After the fit, when the systematic effects have been estimated, they can be subtracted from the raw data and this detrended version can be refit for comparison.

In order to perform the fit we must construct our chi-square variables. Just as in the case of the PSD, the expected value of the kth variance at time interval τ_i can be treated as the sum of five noise contributions

$$\sigma_k^2(\tau_i) = \sum_\alpha h_\alpha \; {}_\alpha \phi_k^2(\tau_i)$$

The functions ${}_\alpha \phi_k^2(\tau_i)$ are the known functional dependencies for a specific noise type (see Table 1). The uncertainty for each noise type $\sigma^2[{}_\alpha \phi_k^2(\tau_i)]$ is also calculable [10] and can be used to construct the variance of the variance estimate

$$\sigma^2[\hat{\sigma}_k^2(\tau_i)] = \sum_\alpha h_\alpha^2 \; \sigma^2[{}_\alpha \hat{\phi}_k^2(\tau_i)]$$

This variance of the variance estimate permits us to calculate the number of degrees of freedom for that variance estimate [4]

$$v_{k,i} = \frac{2[\sigma_k^2(\tau_i)]^2}{\sigma^2[\hat{\sigma}_k^2(\tau_i)]}$$

The number of degrees of freedom is important for two reasons: the definition of our chi-square variable depends explicitly upon it and it determines the shape of the probability distribution.

418

With the number of degrees of freedom and the fit function in hand we have only to subtract off the systematic effects to create our chi-square variable. It is defined in the following manner [4]

$$\chi_{k,i}^2 \equiv v_{k,i} \frac{\hat{\sigma}_k^2(\tau_i) - [Dr_k(\tau_i)]^2}{\sigma_k^2(\tau_i)}$$

$Dr_k(\tau_i)$ is the contribution from deterministic effects on the variance estimate. $\chi_{k,i}^2$ is a random variable that is distributed according to the standard chi-square distribution.

Figure 2 shows the standard chi-square distribution for two different degrees of freedom. It is evident that although the expected value of $\chi_{k,i}^2$ is $v_{k,i}$, the estimate is more likely to be found at the distribution peak that occurs at $v_{k,i} - 2$. For small degrees of freedom this is a significant difference. The estimates of the variances at longer values of τ are formed with fewer data points and consequently have lower degrees of freedom. Thus the most likely estimate will be biased below the true value of that variance. Xi-square fitting correctly accounts for this effect.

Another result apparent from Figure 2 is that the distribution is skewed about this maximum likelihood point. The estimate is more often found to the right of the peak than to the left. This can be corrected by multiplying the cost function by an appropriate factor when the estimate is found to be to the left of the peak. Also chi-square distributed variables have zero probability of being zero or negative. The variance of a chi-square distributed variable is twice the number of degrees of freedom. Thus, lower degrees of freedom lead to narrower, steeper peaks. Putting these ideas together leads to the following definition of the cost function.

$$\xi^2 \equiv \sum_{k,i} \frac{[\chi_{k,i}^2 - v_{k,i} + 2]^2}{2 v_{k,i}}$$

Now it is ξ^2 that must be minimized. It is a non-linear function of the five h_α's and any deterministic parameters we choose to include. One must remember that not only is $\chi_{k,i}^2$ a function of h_α, but $v_{k,i}$ is as well (note that $\chi_{k,i}^2$ is also a function of $v_{k,i}$). The deterministic parameters are found only in $\chi_{k,i}^2$. The minimization of ξ^2 can be accomplished in nearly the same fashion as for a non-linear least squares problem.

Non-linear fitting routines require initial values of the parameters being fit. They attempt to step from one set of values to a better set in an effort to minimize the cost function. The fitting routine described in this paper is not excessively sensitive to the initial guesses. It will converge to the same solution as long as the initial guesses are roughly of the right order of magnitude. It has been observed that it is better to overestimate the magnitude of the parameters and have the routine shrink their value down than to start at too small of a value and try to have it grow out to the correct solution. Thus, to initialize this routine, assume that certain variance estimates are caused entirely by one noise source. Because we know the functional dependencies of the variances on the noise sources we can then estimate the noise intensity coefficients. The same can be done for deterministic parameters. This insures that the guesses are exaggerated but not exceedingly distorted.

V. Results

The method outlined in this paper was first tested against simulated data. With real data there is no way of knowing the true parameters. The routine might consistently converge on the wrong answer without our knowledge. It is therefore crucial to test routines such as this one with computer simulated data with a known truth model. Correctly simulating power law noise can be a difficult task. It is not enough for the noise to have the correct shape (e.g., 1/f), but it must be distributed about that shape in the correct fashion. The criteria by which simulated noise is judged and generated is beyond the scope of this paper. There are, however, several good references on the subject [11]. The noise generated for the truth models used in this analysis came from a routine described in [12].

The routine is able to fit the data very quickly. It often takes more time to form the estimates than to fit them. The one drawback is that the calculation of the variance of the variance estimates is very time intensive (roughly an hour for N on the order of a thousand). Fortunately these values need only be calculated once and then can be stored for subsequent use. When taking data, one can attempt to take the same number of points from run to run. Clearly, an area that warrants further investigation is finding simpler functional approximations for these uncertainties as has been done for the Allan variance [4]. These approximations would permit faster calculation and more flexibility in data taking.
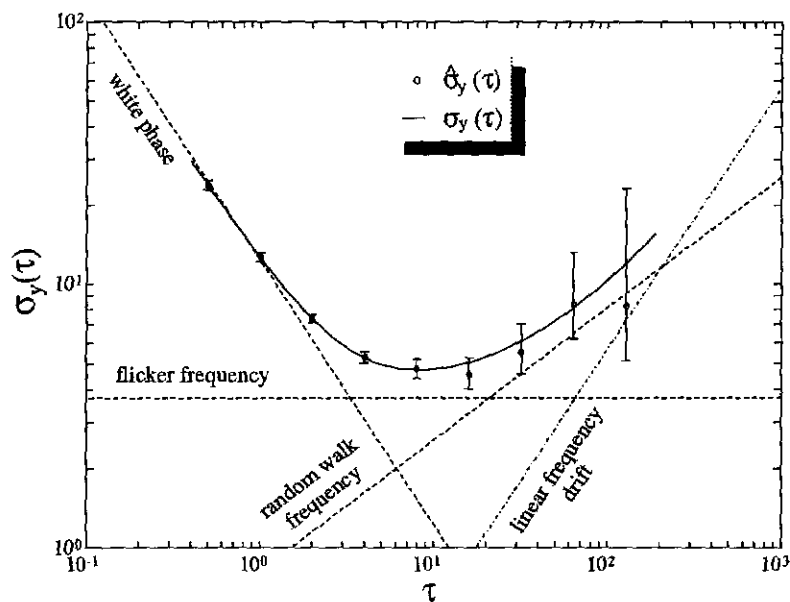

Figure 3. The Allan variance estimates are fit for a particular simulated data set of 1024 points. The solid line is the fit variance. The 90% confidence limits correspond to the fit results. The dashed lines represent the true noise levels and the dash dot line is the true contribution from frequency drift.

Va. Simulated Noise

Various magnitudes of the noise intensity coefficients and linear drift were simulated in combination. There are many combinations in which the contribution of a noise of a certain type is overwhelmed by other noise sources. Also, some noise types may not be observable because a sufficiently long data record was not taken or because the sampling rate was not sufficiently fast. These effects cause such noises to fall below the limits of measurability. Unless one takes an inordinately large amount of data, and none of the noise types completely obscure each other, it will not be possible to precisely determine each noise intensity coefficient and each deterministic parameter. Thus the routine is not always able to resolve each parameter. Obscuration effects are also discussed by Vernotte et. al [1] [2]. When one of these situations occurs, it is important to determine that the parameter has not been well estimated. In these cases, the confidence limits on the estimate are orders of

magnitude larger than the estimate itself. Thus the routine correctly identifies those noise types that have little or no contribution and weights them accordingly. When a noise type is observable, this routine is often capable of correctly estimating the values to within 10% or better.

Parameter	Estimate	Uncertainty	Truth Model	Percent Error
h_2	436	290	500	12.8%
h_1	0.00693	452	0	100%
h_0	25.7	103	0	100%
h_{-1}	9.95	6.17	10	0.5%
h_{-2}	0.0917	0.0662	0.1	8.3%
Dr	0.0752	0.0623	0.08	6.0%
y_0	22.5	5.95	25	10.0%
x_0	-151	17.1	-150	0.667%

Table 2. Best fit coefficients and parameters for the simulated data.

For one particular example of simulated noise, Table 2 lists the estimated parameters, uncertainties, truth model and percent error. Notice that for the two absent noise types (flicker phase and white frequency) the estimated parameters are low and the uncertainties are high. For the noise types that were present, the parameters were well estimated. The comparatively large error for the white phase coefficient is a result of not sampling often enough. This can be seen in Figure 3. More estimates at shorter time intervals are necessary to better resolve this parameter. The fit variance is in excellent agreement with both the estimates and the true variance within the observed time intervals.
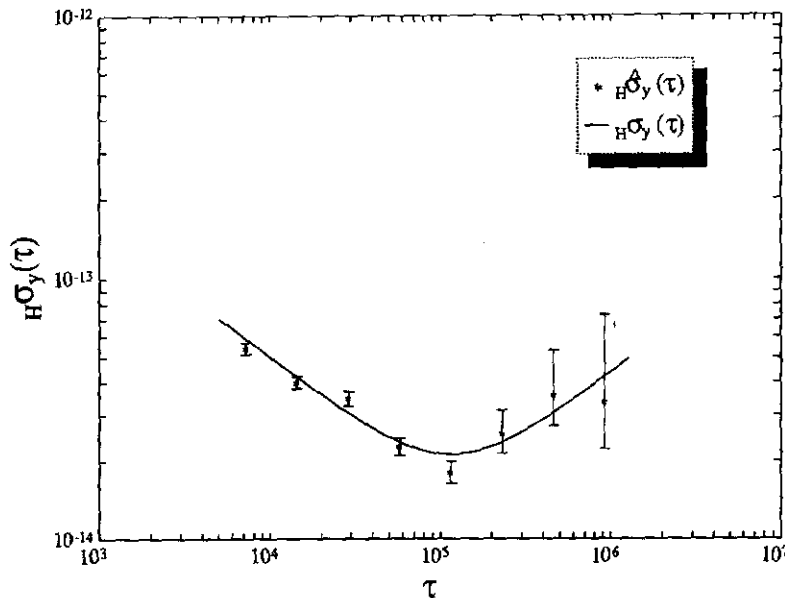

Figure 4. This plot shows the Hadamard variance estimates and fit values from the raw rubidium data. Again the error bars correspond to 90% confidence limits obtained from the fit.

Unfortunately, the confidence intervals on the parameter estimates are excessively large. They are nearly a factor of five too large in the example of Table 2. Some of this error is because the xi-square residuals of the fit are treated as though they are chi-square distributed. While this is a reasonable approximation, more study needs to be done on the true statistics of the residuals to obtain better, stricter estimates of the uncertainties. Another factor is that the different variances are not statistically independent. Probably the best way to place reasonable confidence limits on the parameter estimates would be through computer simulation. By simulating

Parameter	Estimate	Uncertainty
h_2	1.39e-22	1.11e-18
h_1	7.76e-24	2.43e-18
h_0	4.42e-23	4.70e-23
h_{-1}	2.65e-32	2.53e-28
h_{-2}	5.17e-34	3.17e-34
Dr	-2.41e-19	5.9e-20
y_0	2.182617e-10	4.67e-14
x_0	1.18e-09	1.66e-09

Table 3. Best fit coefficients and parameters for the rubidium oscillator vs. AT1 data.

many data sets in the region of the predicted parameters, one could obtain a better feel for what the real confidence limits might be.

This routine is particularly adept at picking out small values of linear drift even when completely buried in the noise. Notice in Figure 3 that the last data point has a large uncertainty and occurs where the linear drift level is still below the random walk frequency contribution. Yet the routine still estimated the drift parameter to better than 10%. If the second difference method of estimating drift [4] [13] had been applied to this data, it would have obt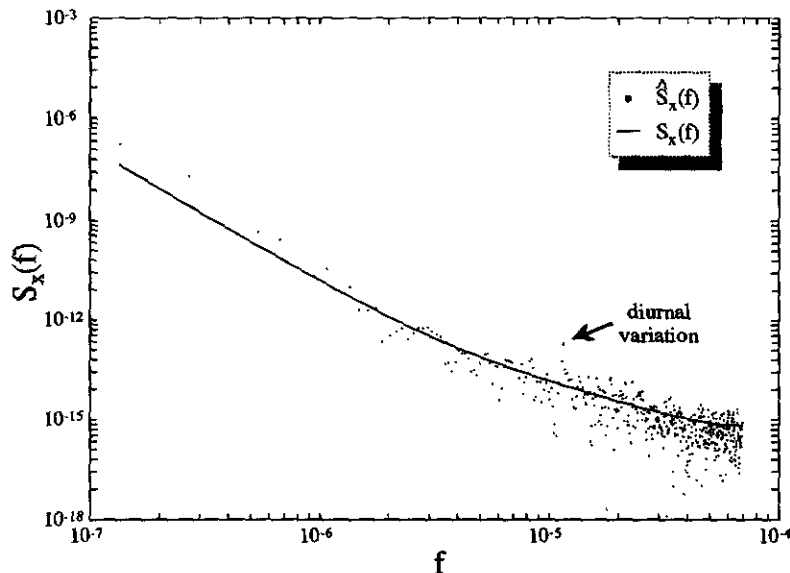ained a value of -0.015. That method is incapable of estimating drift when it is so far buried in the noise. For larger relative values of linear frequency drift, the second difference method yields estimates comparable to, and sometimes better than, those found with this routine.

Also, notice in Figure 3 that the last point dips well below the true variance or even below the contribution just from the random walk frequency noise. For this point the number of degrees of freedom is predicted to be 3.135 and the value of the chi-square variable corresponds to 0.731 or right near the distribution peak. Thus the estimate has less than one third the value of the true variance. Because of the scarcity of data for this time interval, the variance estimate is not very good. If the chi-square distribution were not correctly applied to this case, one would obtain an overly optimistic prediction of the stability.

Vb. Rubidium Data

The routine was also tested on real data from an EG&G rubidium oscillator. This oscillator was measured against ensemble time (AT1) at NIST [14]. The raw data was fit reasonably well by this routine. Because of the large drift that was present in this oscillator the Hadamard variance is a good measure of the stability, Figure 4. Unfortunately it can be observed that the fit values lie outside the 90% confidence limits for some of the estimates. Such an effect could have a number of causes: noise that does not follow the standard power law model, environmental effects or other deterministic effects such as periodic modulation of the data. Because the most

Figure 5. The power spectral density of the partially detrended data is plotted as a function of Fourier frequency. The diurnal variation is clearly visible.
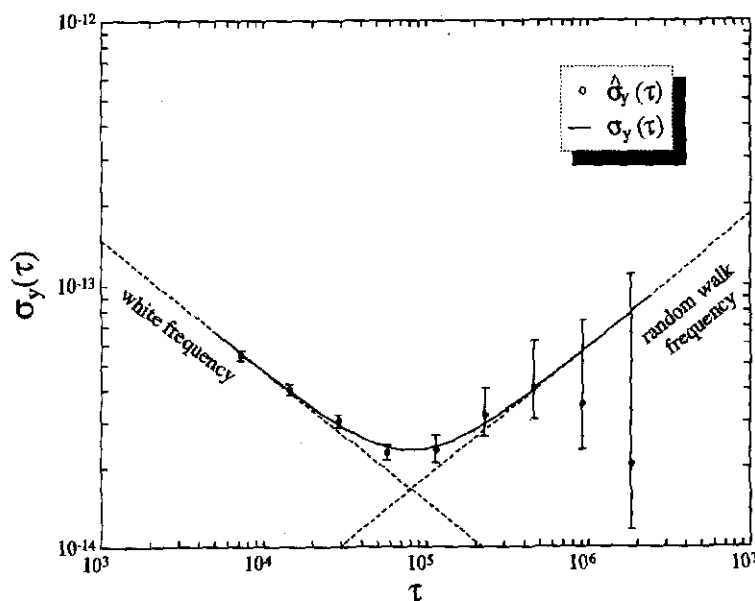
Figure 6. The Allan variance estimates and fit variances are shown for the fully detrended rubidium vs. AT1 data. The dashed lines indicate the estimated level of contribution from each noise type.

affected point corresponds to a time interval of nearly half a day, diurnal variations were suspected. Using the estimates of the drift and frequency offset, these effects were removed from the data. Investigation of the PSD of the detrended data confirmed the suspicion (Figure 5). By fitting a sine-wave response to the PSD, a diurnal modulation with an amplitude of 0.35ns was found and subtracted from the data. The magnitude of the variation is reasonable for an oscillator with a temperature coefficient of 2×10^{-12} and temperature variations on the order of a tenth of a degree Celsius. When this detrended data was refit, the fit variances were in excellent agreement with the estimates (see Figure 6). Thus it was this periodic modulation that had corrupted the initial fit. It is evident that only two significant noise types affect the data. This can also be seen in the estimated noise coefficients in Table 3. Only the estimated coefficients for white frequency modulation and random walk frequency modulation are not well below their confidence limits. Again the confidence limits in this case could probably be reduced significantly. Nevertheless the presented routine was not only able to fit the real data, but it was also able to indicate the periodic modulation that was present.

VI. Acknowledgments

The Author is grateful to the Time and Frequency Division at NIST, particularly Marc Weiss and David Allan, for providing the rubidium data. I am also grateful to Bill Riley at EG&G for providing information about their rubidium oscillator. Finally, I also wish to thank Professor John Turneaure and Dr. Jeremy Kasdin for their many helpful comments and suggestions.

VII. References

[1]   F. Vernotte, E. Lantz, J. Groslambert, J. J. Gagnepain, "A New Multivariance Method for the Oscillator Noise Analysis," Proceedings of the 1992 Frequency Control Symposium, pp 284-289, 1992.

[2]   F. Vernotte, E. Lantz, "Time Stability: An Improvement of the Multi-Variance Method for the Oscillator Noise Analysis," Proceedings of the 6th European Frequency and Time Forum, pp 343-347, 1992.

[3] J. A. Barnes, et. al., "Characterization of Frequency Stability," IEEE Trans. Instr. Meas., Vol. IM-20, pp. 105-120, May 1971.

[4] D. B. Sullivan, D. W. Allan, D. A. Howe and F. L Walls, Characterization of Clocks and Oscillators, NIST Technical Note 1337, 1990

[5] R. A. Baugh, "Frequency Modulation Analysis with the Hadamard Variance," Proceedings 25th Frequency Control Symposium, pp 222-225, 1971.

[6] J. Rutman, "Characterization of Phase and Frequency Instabilities in Precision Frequency Sources: Fifteen years of progress," Proceedings of the IEEE, vol. 66, no. 9, pp. 1048-1075, Sept 1978.

[7] J. A. Barnes, "Atomic Timekeeping and the Statistics of Precision Signal Generators," Proceedings of the IEEE, Vol 54, pp. 207-220, February 1966.

[8] W. C. Lindsey, C. M. Chie, "Theory of Oscillator Instability Based Upon Structure Functions," Proceedings of the IEEE, Vol 64, pp. 1652-1666, December 1976.

[9] P. Lesage, C. Audoin, "Characterization of Frequency Stability: Uncertainty due to the Finite Number of Measurements," IEEE Trans. Instr. Meas., Vol. IM-22, pp. 157-161, June 1973.

[10] T. Walter, "Characterization of Frequency Stability: A Continuous Power Law Model with Discrete Sampling," submitted to IEEE Trans. Instr. Meas.

[11] N. J. Kasdin, "Discrete Simulation of Colored Noise and Stochastic Processes and $1/f^{\alpha}$ Power Law Generation," submitted to Proceedings of the IEEE.

[12] N. J. Kasdin, T. Walter, "Discrete Simulation of Power Law Noise," Proceedings of the 1992 Frequency Control Symposium, pp 274-283, 1992.

[13] M. A. Weiss, D. W. Allan, D. A. Howe, "Confidence on the Second Difference Estimation of Frequency Drift," Proceedings of the 1992 Frequency Control Symposium, pp 300-305, 1992.

[14] D. W. Allan, J. Levine, "A Rubidium Frequency Standard and a GPS Receiver: A Remotely Steered Clock System with Good Short-Term and Long-Term Stability," Proc. of the 44th Ann. Symp. on Freq. Cont., pp 151-160, 1990.

QUESTIONS AND ANSWERS

Question: I imagine it is hard to estimate drift in the presence of random walk. They start looking the same. I am wondering if I have not seen anybody do this, but can you imagine a way to estimate drift from random walk, and remove it that would not automatically bias the variance low. That would at least give some deviation ....

T. Walker, Stanford University: I can imagine a way that would work. I think a slightly better way to implement it, rather than subtracting the drift the way I do, would be to actually detrend the data after every step. This would be computationally impossible or not worth doing but if you actually detrend the data as you go along, optimize that way rather than subtracting, there is cross terms that can get in there. But you can expect from random walk frequency that the last point will be biased low because you essentially have only degree of freedom, or something very low that will be biased very low. What you have to realize ....

Questioner: Just because it is Chi-Square?

T. Walker: Well I think that the statistics for the Allan variance estimate is almost worthless for just one estimate. You really do not have enough information to say anything meaniful about what the stability is.

Questioner: Why bias low?

T. Walker: Well it could be high but is more likely to be found at that.

D. Allan, Allan's Time: It is interesting to look back at classical statistics for these low frequency processes. They turn out to be incredibley sensitive to low frequency. In fact they diverge but we know they diverge as a function of a number of samples and if you look at that independence you can get estimates of some of these low frequency properties for the very low frequency components; ie: one cycle per data length. The standard deviation is a very good measure and it is sensitive to the number of samples and the kind of parallel processes. I wonder if we could exploit this some to help us. It is a measure we have kind of forgotten, that has information in it.

T. Walker: I think that you could, providing you have a model, like this, where you assume the noise type. You certainly can do a lot with the statistics in analyzing what you are seeing.

H. Fliegel, The Aerospace Corporation: I may be wrong. I am trying to remember something from a long time ago. I wonder if it is mathematically even possible to separate linear drift from random walk. The only way I ever thought it might be handled is through the arc sine law. If you have a fantastic amount of data, then the number of zerocrossings you get from a pure random walk is predictable. You could use that to estimate roughly where your line should go. I do not know if that is practical.

T. Walker: Right, it is probably not practical and I think what you are saying is correct. What I attempted to do in this, is use some of the information by fitting them at the same time. Then you do have some information on each level. You may be able to subtract them again. You need to detrend the data after you find the drift and verify that has been correctly done.

R. Keating, USNO: I just want to compliment you on a fine and interesting paper. It is not very often that we get comments from the people that ask questions. What I would like to know, what are your plans for the future, what are you going to do now?

T. Walker: That is a perfect question. I am a graduate student at Stanford right now and I am finishing up my thesis, hopefully in the spring. I actually do not have plans beyond that.

I am working on the gravity B project, and I may be continuing with that. I definately would like to stay involved with some of this work. I already see some things that could potentially be done in extending this. So I would have to say my plans are not set. If anyone has any offers for plans, I would gladly accept them.