A Historical Perspective on the Development of the Allan Variances and Their Strengths and Weaknesses

David W. Allan and Judah Levine, *Member, IEEE*

***Abstract*—Over the past 50 years, variances have been developed for characterizing the instabilities of precision clocks and oscillators. These instabilities are often modeled as nonstationary processes, and the variances have been shown to be well-behaved and to be unbiased, efficient descriptors of these types of processes. This paper presents a historical overview of the development of these variances. The time-domain and frequency-domain formulations are presented and their development is described. The strengths and weaknesses of these characterization metrics are discussed. These variances are also shown to be useful in other applications, such as in telecommunication.**

***Index Terms*—Allan variances (AVARs), atomic clocks, nonstationary processes, precision analysis, time series analysis.**

I. Introduction

NATURE gives us many nonstationary and chaotic processes. If we can properly characterize these processes, then we can use optimal procedures for estimating, smoothing, and predicting them. During the 1960s through the 1980s, the Allan variance (AVAR), the modified Allan variance (MVAR), and the time variance (TVAR) were developed to this end for the timing and the telecommunication communities. Since that time, refinements of these techniques have been developed. The strengths and weaknesses of these variances will be discussed in the following text. The applicability of these variances has been recognized in other areas of metrology as well. Knowing the strengths and weaknesses is important so that they can be properly used.

Prior to the 1960s, before atomic clocks were commercially available, quartz-crystal oscillators were used for timekeeping. The largest contribution to the long-term-frequency instability of these oscillators was frequency drift. It was generally recognized that the stochastic contribution to the long-term performance could be modeled by flicker-noise frequency modulation (FM), which is a nonstationary process, because it has a power-spectral-density proportional to $1/f$, where f is the Fourier frequency. The integral of the power spectral density for this type of process diverges logarithmically, so that the data cannot be characterized by a well-defined classical variance. In real experiments, the upper and lower limits on the integral of the power spectral density are bounded by frequencies of order $1/\tau$, the reciprocal of the averaging time between measurements, and $1/\mp$, the reciprocal of the total elapsed time $\mp$ of the data set, respectively. The variance is thus a function of the number of data points that are used in the computation, since this number is on the order of $\mp/\tau$.

In 1964, Barnes developed a generalized autocorrelation function that was well behaved for flicker noise [1]. Barnes' work was the basis for his Ph.D. thesis, and it also gave Allan the critical information that he needed for his master's thesis, which was based on Barnes' results and the work of Lighthill [2]. Allan studied the dependence of the classical variance of frequency-difference measurements as a function of various parameters of the measurement process.

The frequency-difference of a device under test is measured as the evolution of the time-difference between it and a second, standard device over some time interval. The result is the average frequency-difference over that averaging time interval. In addition, the early version of the hardware that was used to make these time-difference measurements could not make measurements continuously, and required some "dead time" between measurements. Allan studied the estimate of the classical variance as a function of the averaging time τ the number of samples that were included in the variance N, the dead time between frequency averages $T - \tau$, where T was the time between the beginning of one measurement to the beginning of the next one, and the dependence of the variance on the measurement-system bandwidth f_h.

Barnes and Allan developed a set of spectral density, power-law noise models that covered the characterization of the different kinds of instabilities that were observed in clock data. The models, which proved to be very useful, included the noise of the measurement systems, the frequency fluctuations of the clocks, and any environmental influences. These results were published as a Technical Report of the National Bureau of Standards (which became NIST, the National Institute of Standards and Technology in the 1980s), and in a Special Issue of the Proceedings of the IEEE on "Frequency Stability" [1], [3].

II. Modeling Clocks With Power-Law Noise Processes

If the free-running frequency of a clock at some time is $\nu(t)$, and its nominal frequency is ν_o, then the normalized dimensionless frequency deviation at that time is $y(t) = (\nu(t) - \nu_o)/\nu_o$.

Manuscript received October 6, 2015; accepted January 29, 2016. Date of publication February 12, 2016; date of current version April 1, 2016.

D. W. Allan is with Allan's TIME, Fountain Green, UT 84632-0066 USA (e-mail: david@allanstime.com).

J. Levine is with the Time and Frequency Division, National Institute of Standards and Technology, Boulder, CO 80305-3328 USA (e-mail: Judah.Levine@nist.gov).

Digital Object Identifier 10.1109/TUFFC.2016.2524687

The time-deviation (TDEV) of a clock $x(t)$ is the integral of $y(t)$. The frequency-domain spectral densities of the time and frequency fluctuations are $S_x(f)$ and $S_y(f)$, respectively. The spectral densities for many clocks and oscillators can be represented as a power of the Fourier frequency, $S_y(f) \sim f^{\alpha}$ and $S_x(f) \sim f^{\beta}$, where the exponents are small integers, and $\alpha = \beta+2$, since the frequency is the evolution of the time-difference over some averaging time. The statistical models of clocks, their measurement systems, and their distribution system can generally be modeled with integer values of alpha from -2 to $+2$. For most clocks and oscillators, the value of α in the noise model becomes more negative as the averaging time is increased. That is, the dependence of the power spectral density on Fourier frequency diverges at low Fourier frequencies.

Fig. 1(a)–(e) displays examples of the visual appearance of data that can be modeled by the different power-law spectra. In each case, the data set is computed so that the two-sample Allan deviation is nominally the same value for an averaging time of 1 s—the time interval between individual samples. As one can see in the figure, the appearances of these power-law spectra are very different, so that it is often possible to estimate the exponent of the power spectral density by a simple visual examination of the data. This visual examination is often difficult in practice, because most data sets cannot be characterized only by a single noise type.

III. Time-Domain Representation

In his master's thesis [4], Allan had studied the divergence of the classical variance for the power-law noise processes described above as a function of the number of data points taken. The divergence depends upon both the number of data points in the set as well as upon the kind of noise. In other words, the classical variance was data-length-dependent for all of the power-law noise models being used to characterize clocks except for classical white-frequency noise. Hence, the classical variance was not useful in characterizing atomic clocks, because more than just white-frequency noise models were needed.

The two-sample variance, which is typically called the AVAR, may be written as follows:

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

where the brackets denote the average over the ensemble of observations, and the "2" in the denominator normalizes it to be equal to the classical variance in the case of classical white-frequency noise. By using the results of Lighthill and Barnes referenced earlier, Allan showed that this variance is well behaved and convergent for all the interesting power-law spectral density processes that are useful in modeling clocks and measurement systems.

The bias function $B_1(N)$, which is the ratio of the classical N-sample variance to the AVAR as a function of N [5], is defined as

$$B_1(N) = \sigma^2(N)/\sigma_y^2(\tau_0). \tag{2}$$

It is a function of N in all cases except for classical white noise. One can turn this dependence to an advantage and use it to characterize the kind of noise by estimating the value of the bias function for the data set being studied and comparing it to the values expected for the different noise types.

The two-sample or AVAR is defined above without dead time. In other words, the frequency measurements are sequentially adjacent. For example, the ith frequency deviation taken over an averaging time $\boldsymbol{\tau}$ may be computed from the TDEVs as $y_i = (x_i - x_{i-1})/\boldsymbol{\tau}$. This equation gives the average frequency deviation over that interval, but it may not be the optimum estimate of frequency. If the average is taken over the whole data set, then all the intermediate-differences cancel, and one is left with the average frequency deviation over the data set: $y_{\text{avg}} = (x_N - x_0)/N\boldsymbol{\tau}$. This is one of the benefits of no-dead-time data. For white-frequency noise, $\sigma_y^2(\tau)$ is an optimum-variance estimator of the change of frequency over any averaging time $\boldsymbol{\tau}$ and is equal to the classical variance for the minimum data-spacing τ_0.

Barnes also showed that $\sigma_y^2(\tau)$ is an unbiased estimator for the level of the power-law noise process of interest in modeling atomic clocks and that it is Chi-squared distributed. The value of $\boldsymbol{\tau}$ in the software analysis can take on values for all $\boldsymbol{\tau} = n\,\boldsymbol{\tau}_o$ for any integer $\boldsymbol{n = 1}$ to $\boldsymbol{n = N/2}$. The confidence of the estimate is best at $\boldsymbol{\tau = \tau}_o$ decreasing to $\boldsymbol{\tau = (N/2)\tau}_0$, where there is only one degree-of-freedom for the confidence of the estimate. The Chi-squared-distribution function has a most probable value of zero for one degree-of-freedom. Even though it is unbiased, the probability of small values is significant. Therefore, in a plot of $\boldsymbol{\sigma_y^2(\tau)}$ as a function of τ, one often observes too small values for $\boldsymbol{\sigma_y^2(\tau)}$ as the value of $\boldsymbol{\tau}$ approaches half the data length, because the number of degrees-of-freedom is too small for a good confidence on the estimate. This problem was addressed many years later by David Howe as we discuss below [14], [15].

For the noise types commonly found in time and frequency applications, the simple power-law dependence of the spectral density on the Fourier frequency results in a corresponding power-law dependence of the two-sample AVAR on the averaging time. That is, if $S_y(f) \sim f^{\alpha}$, then $\sigma_y^2(\tau) \sim \tau^{\mu}$. Fig. 2, based on Lighthill's work referenced above [2], shows the relationship between μ and α for the noise types commonly found in clock and oscillators. The slope of a log-log plot of the AVAR as a function of averaging time can be used to estimate both the kind of noise and its magnitude in the time domain. The relationship between α and μ can also be used to estimate the power spectral density of the noise for any averaging time.

From Fig. 2, we can see that there is an ambiguity problem for the simple AVAR at $\mu = -2$. The relationship between α and μ is no longer unique at that point, and one cannot tell the difference in the time domain between white and flicker phase noise processes. This problem was a significant limitation in clock characterization for the time and frequency community for 16 years after the simple AVARs was developed. Even though there is an ambiguity in the $\boldsymbol{\tau}$ dependence in this region, it was known that it could be resolved because the variance also depended on the measurement bandwidth. Since it was inconvenient to modulate the measurement-system bandwidth

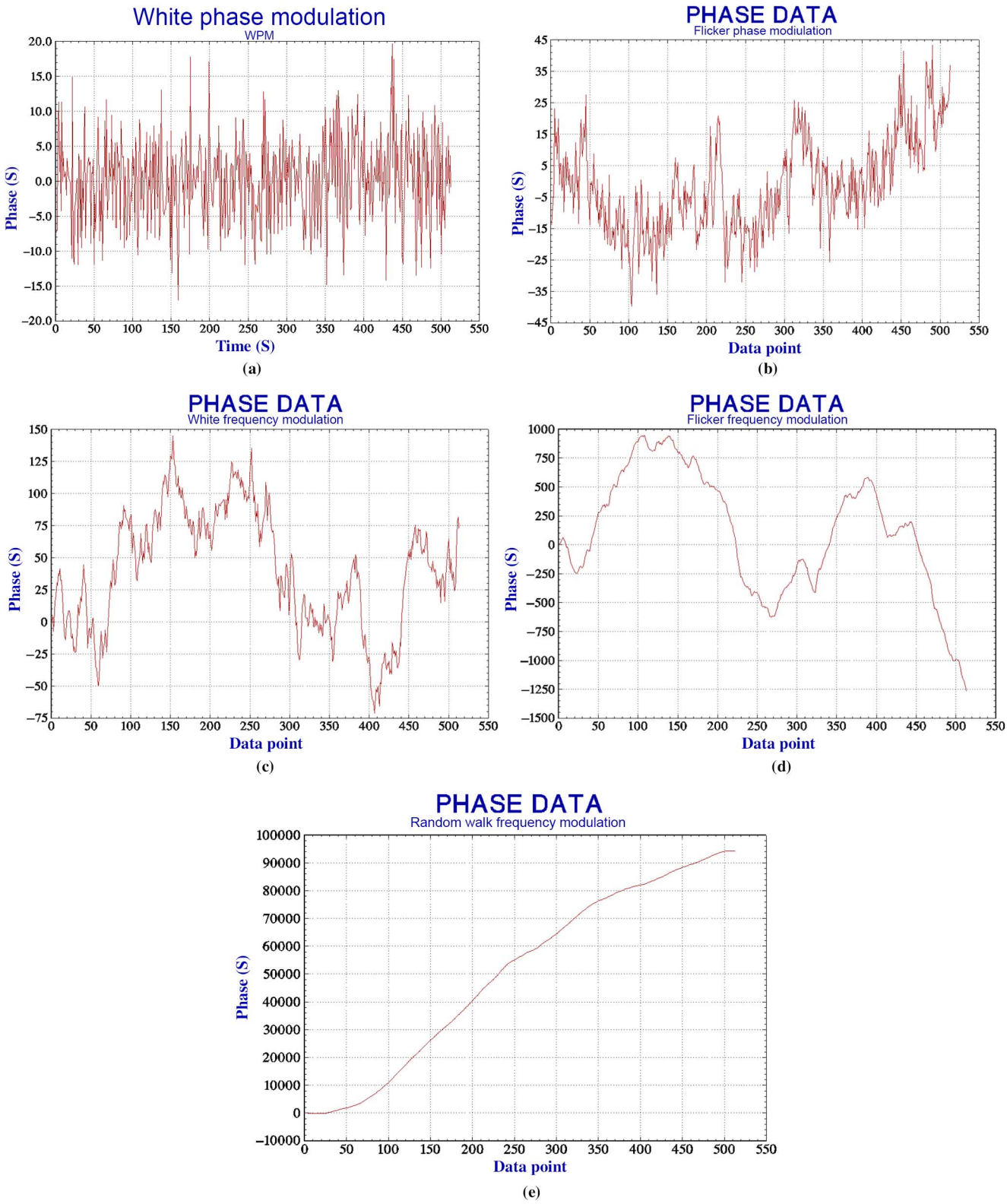

Fig. 1. (a)–(e) Appearance of typical time-difference data that can be modeled by the five common power-law spectra. Each plot shows 513 points computed with the same value for the Allan deviation at an averaging time of 1 s—the interval between data points. The x- and y-axes are in units of seconds.

to distinguish between white and flicker phase noise types, this approach never became useful. But, in 1981, a way was discovered to effectively modulate the bandwidth as part of the estimation process, and this was the breakthrough needed. [6]. This gave birth to MVAR, the modified Allan variance, and the concept of varying the bandwidth by averaging is illustrated in Fig. 3.

One can think of software bandwidth modulation in the following way. There is always a finite measurement-system bandwidth. Call it the hardware bandwidth f_h. Let $\tau_h = 1/f_h$.

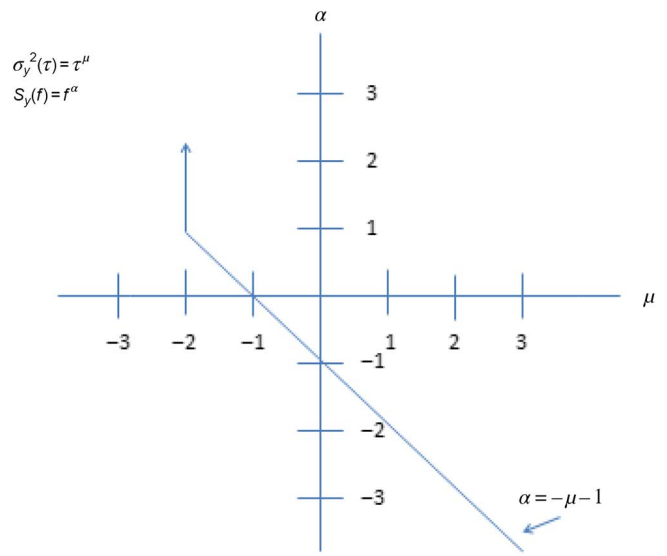

Fig. 2. Relationship between the exponent of the power spectral density α and μ, the exponent of the dependence of the two-sample AVAR on averaging time for each of the common noise types. Note the elegant relationship between the exponent of the dependence of the power spectral density on Fourier frequency α and the exponent of the dependence of the AVAR on averaging time μ is given by the simple equation $\boldsymbol{\alpha = -\mu - 1}$.

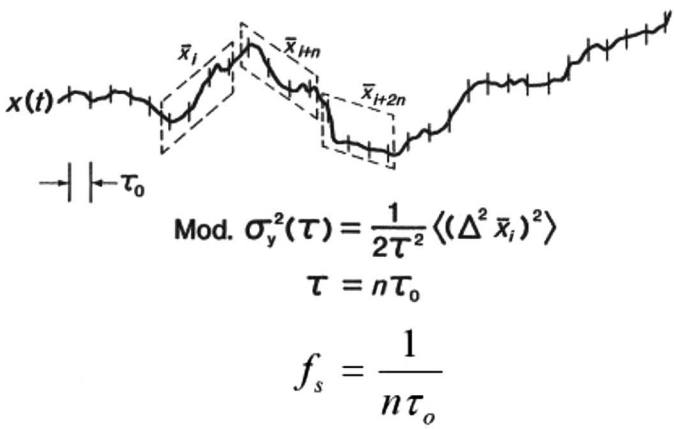

Fig. 3. Software-bandwidth modulation technique used in the modified AVAR to resolve the ambiguity problem at $\boldsymbol{\mu = -2}$ and allows us to characterize all the power-law spectral density models from $\alpha = -3$ to $\alpha = +2$, which includes the range of useful noise models for most clocks. Illustrated in this figure is the case for 4-point averages, or $n = 4$. The averaging parameter n takes on values from 1 to *N*/3, where *N* is the total number of data points in the data set with a spacing of τ_o.

Then, every time a phase or time reading is added to the data, it inherently has a $\boldsymbol{\tau}_h$ sample-time window. If n of these samples are averaged, the sample-time window has been increased by n, $\boldsymbol{\tau}_s = n\boldsymbol{\tau}_h$. Let $\boldsymbol{\tau}_s = 1/f_s$. Now if we increase the number of samples averaged as $\boldsymbol{\tau}$ is increased, then the software bandwidth is decreased by the reciprocal of the number of samples averaged, or $1/n$. Modulating the bandwidth in this way removes the above ambiguity and maintains validity for our simple Fourier transform equation over all the power-law noise processes of interest: $\alpha = -\mu - 1$.

In the later part of the 1980s, the telecommunications industry in the United States came to Allan and asked for help in developing a metric for characterizing telecommunication networks. Allan and Dr. Marc Weiss worked on this problem, and analyzed a lot of telecommunications data sent to them to find the best metric. Out of this work they developed the time variance, TVAR. It is defined as follows: $\mathrm{TVAR} = \boldsymbol{\tau}^2$ MVAR/3. The "3" in the denominator normalizes it to be equal to the classical variance in the case of white-noise phase modulation, or WPM. One can show that for white-noise PM, TVAR is an optimum estimator of change in the phase or time residuals in a variance sense.

The three variances, AVAR, MVAR, and TVAR, became international IEEE time-domain measurement standards in 1988 [7]. There are three general regions of applicability for time and frequency systems:

1) AVAR for characterizing the performance of frequency standards and clocks;
2) MVAR for characterizing the performance of time- and frequency-distribution systems;
3) TVAR for characterizing the timing errors in telecommunication networks.

Following the development of each of these three variances, many other areas of applicability have arisen. The TDEV, which is the square-root of TVAR, has no dead-time issues and has become a standard metric in the international telecommunications industry. All three have application capability in many other areas of metrology. If you conduct a Web search for "AVAR," you will find about 50 000 results.

IV. Summary of the Definitions of the Variances

The equations for computing AVAR, MVAR, and TVAR from *N* measurements of the time deviations are, respectively,

$$\begin{aligned}
\sigma_y^2(\tau) &= \frac{1}{2\tau^2(N-2n)} \sum_{i=1}^{N-2n} (x_{i+2n} - 2x_{i+n} + x_i)^2 \\
\mathrm{mod}.\sigma_y^2(\tau) &= \frac{1}{2\tau^2 n^2 (N-3n+1)} \\
&\quad \sum_{j=1}^{N-3n+1} \left(\sum_{i=j}^{n+j-1} (x_{i+2\,m} - 2x_{i+n} + x_i) \right)^2 \\
\sigma_x^2(\tau) &= \frac{1}{6n^2(N-3n+1)} \\
&\quad \sum_{j=1}^{N-3n+1} \left(\sum_{i=j}^{n+j-1} (x_{i+2\,m} - 2x_{i+n} + x_i) \right)^2 \qquad (3)
\end{aligned}$$

where x_i are the measured time deviation data separated by a time interval $\boldsymbol{\tau}_o$, and $\boldsymbol{\tau} = n\boldsymbol{\tau}_o$. For MVAR and TVAR, the computation involves a double sum. Although a simple evaluation of these variances would require a computation time that increased as N^2, which would be a problem for large data sets, one can employ some computation tricks, such as simple drop–add averaging, to make the time linear in *N*. The software references cited later include these computation techniques [8].

The following equations show how the three time-domain variances may be derived from frequency-domain information.

TABLE I
COEFFICIENTS RELATING THE POWER SPECTRAL DENSITY OF THE FACTIONAL-FREQUENCY FLUCTUATIONS $S_Y(f)$ AND THE RESIDUAL TIME FLUCTUATIONS $S_x(f)$ TO THE AVAR AND THE TVAR, RESPECTIVELY, FOR THE FIVE COMMON NOISE TYPES

NOISE TYPE	$S_y(f)$	$S_x(f)$
White PM	$\frac{(2\pi)^2}{3f_h}[\tau^2\sigma_y^2(\tau)]f^2$	$\frac{1}{\tau_o f_h}[\tau\sigma_x^2(\tau)]f^0$
Flicker PM	$\frac{(2\pi)^2}{A}[\tau^2\sigma_y^2(\tau)]f^1$	$\frac{3}{3.37}[\tau^0\sigma_x^2(\tau)]f^{-1}$
White FM	$2[\tau^1\sigma_y^2(\tau)]f^0$	$\frac{12}{(2\pi)^2}[\tau^{-1}\sigma_x^2(\tau)]f^{-2}$
Flicker FM	$\frac{1}{2\ln 2}[\tau^0\sigma_y^2(\tau)]f^{-1}$	$\frac{20}{(2\pi)^2\,9\ln 2}[\tau^{-2}\sigma_x^2(\tau)]f^{-3}$
Ramdom Walk FM	$\frac{6}{(2\pi)^2}[\tau^{-1}\sigma_y^2(\tau)]f^{-2}$	$\frac{240}{(2\pi)^4\,11}[\tau^{-3}\sigma_x^2(\tau)]f^{-4}$

$A = 1.038 + 3\ln(2\pi f_h\tau)$.

One cannot do the reverse—derive the spectral densities from time-domain analysis. It is often very useful to analyze the data in both the frequency and time domains:

$$\tau = n\tau_0$$

AVAR:

$$\sigma_y^2(\tau) = \int_0^\infty 2\left[\frac{\sin^4(\pi f\tau)}{(\pi f\tau)^2}\right] S_y(f)\,df.$$

MOD AVAR:

$$\mathrm{Mod}\sigma_y^2(\tau) = \int_0^\infty 2\left[\frac{\sin^3(\pi f\tau)}{(n\pi f\tau)\sin(\pi f\tau_0)}\right]^2 S_y(f)\,df. \quad (4)$$

TVAR:

$$\sigma_x^2(\tau) = \frac{8}{3n^2}\int_0^\infty \left[\frac{\sin^3(\pi f\tau)}{\sin(\pi f\tau_0)}\right]^2 S_y(f)\,df.$$

Allan *et al.* [9] have computed the effective Fourier windows using the transfer functions $H(f)^2$ for each of these three variances for $n = 1$, 2, 4, 8, 16, 32, 64, 128, and 256. The transfer function between the time-domain and the frequency-domain for any of the two-sample variances is approximately equivalent to observing the Fourier components over a nominally square window in Fourier space if the tau values used are incremented by 2^n, where $n = 0$, 1, 2, 3,... up to that allowed by the data length. Allan *et al.* shows that the transfer-function window is different for each of the three different variances. From experience, one can often observe low-frequency Fourier components in the time domain better than in the frequency domain. TVAR is especially sensitive to low-frequency components. We have used this advantageously on several occasions.

Table I shows the conversion relationships between the AVAR and power spectral density for the five common noise types.

V. ESTIMATION, SMOOTHING, AND PREDICTION

There is a simple and powerful statistical theorem that is useful for estimation, smoothing, and prediction. It is that the optimum estimate of the mean value of a stochastic process with a white-noise spectrum is the simple mean. As examples, in the case of white-noise phase modulation, the optimum estimate of the phase or the time is the simple mean of the independent phase or time-residual readings added to a systematic value, if necessary. In the case of white-noise frequency modulation, WFM, the optimum estimate of the frequency is the simple mean of the independent frequency readings, which is equivalent to the last time-reading minus the first time-reading divided by the data length, if there is no dead time between the frequency measurements. Thus, the best estimate of the average frequency is given by $y_{\mathrm{avg}} = (x_N - x_0)/N\tau$.

Using the above theorem for optimum prediction, if the current time is "*t*," and one desires to predict ahead an interval τ, then the optimum time prediction, for a for a clock having WFM and an average offset frequency y_{avg} given by the above equation, is given by a simple linear extrapolation

$$\hat{x}(t+\tau) = x(t) + y_{\mathrm{avg}}(t)\,\tau. \quad (5)$$

The even-powered exponents are directly amenable to this theorem, but the flicker-noise (odd exponents) are more complicated. However, there is a simple prediction algorithm for flicker frequency modulation using what is called the second-difference predictor. It is very close to optimum and is simple. A prediction of τ seconds in the future can be obtained by the following equation:

$$\hat{x}(t+\tau) = 2x(t) - x(t-\tau) + \left(\Delta^2 x\right)_{\mathrm{avg}} \quad (6)$$

where *t* is the current time and $(\Delta^2 x)_{\mathrm{avg}}$ is the average value of the second-difference of the TDEVs—spaced by τ—over the past available data. The first two terms on the right side of (6) are simply the prediction based on the assumption of a constant frequency offset, and the value of $(\Delta^2 x)_{\mathrm{avg}}$ will be nonzero if frequency drift is present. If there is no drift, it will tend to zero for most of the common noise processes. In general, the time predictability is given approximately by $\tau\sigma_y(\tau)$.

VI. SYSTEMATICS

A good model for the time deviations in a clock is $x(t) = x_o + y_o t + \frac{1}{2}\,Dt^2 + \varepsilon(t)$, where x_o and y_o are, respectively, the synchronization error and syntonization error at $t = 0$, *D* is the frequency drift, and $\varepsilon(t)$ represents the remaining random errors in addition to the first three systematic terms. It is important to subtract the systematics from the data, so that the random effects can be viewed visually and then analyzed with better insights.

In addition, if frequency drift *D* is present in a clock, then it adds a bias to AVAR, MVAR, and TVAR. For AVAR and MVAR, it increases the variance estimates by $D\tau/\sqrt{2}$. For TVAR, the increase is $D\tau^2/\sqrt{6}$. If there is frequency drift, the values of $\sigma_y(\tau)$ in that region where the drift is affecting the plot will lie very close to the τ^{+1} line. If there is random noise present, then the values will not fit tightly to this line.

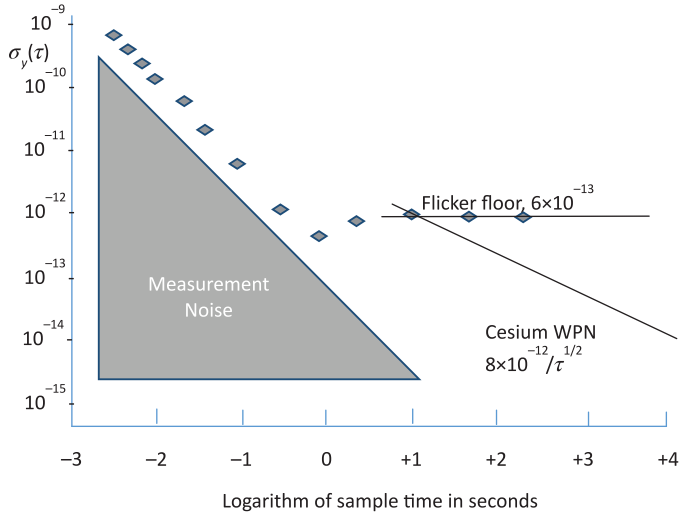

Fig. 4. Sigma-tau plot of the frequency instabilities between a precision free-running, quartz-crystal oscillator, and a commercial cesium-beam atomic clock. The shaded region shows the contribution of the measurment noise to the variance. The points are the measured values, and the fitted lines display the estimated noise type as discussed in the text.

The Hadamard variance is an estimate based on the second-difference of the fractional frequencies or, equivalently, the third-difference of the time-difference data. Therefore, it is not sensitive to a constant frequency drift [10].

If there is a frequency modulation with frequency f_m present in the data, then there is a systematic bias to the calcuated value for ADEV given by

$$\sigma_y(\tau) = \frac{x_{\rm pp}}{\tau}\sin^2(\pi f_m \tau) \tag{7}$$

where $x_{\rm pp}$ is the peak-to-peak amplitude of the modulation. Both MVAR and TVAR are affected as well. The effect of the modulation is exactly cancelled by aliasing for averaging times of $\boldsymbol{\tau} = n/f_m$, where n is any positive integer. Recognizing this null effect allows these three variances to be used as low-frequency spectrum analysis techniques for bright Fourier-frequency lines in the data. It is our experience that the low-frequency spectral lines can often be observed using this null approach in the time domain better than can be observed in the frequency domain.

VII. Example

In Fig. 4, we show a sigma-tau plot of the frequency instabilities between a precision free-running, quartz-crystal oscillator, and a commercial cesium-beam atomic clock. For sample times shorter than 1 s, we see a $\boldsymbol{\tau}^{-1}$ behavior due to the measurement noise. This plot was made before MDEV, the square-root of MVAR, was developed, so the noise type is ambiguous. The Allan deviation, ADEV, is the square root of AVAR. Whenever a $\boldsymbol{\tau}^{-1}$ behavior occurs in an ADEV plot, one should then analyze the data using MDEV to resolve the ambiguity regarding the kind of noise modulation present in the data.

The rise in the value of $\sigma_y(\boldsymbol{\tau})$ as the sample or averaging time approaches 10 s is due to the attack time of the cesium-beam locking its quartz-crystal-slave-oscillator to the cesium resonance. Over the next decade, one observes a $\boldsymbol{\tau}^{-1/2}$ behavior or $\boldsymbol{\mu} = -1$, which means $\alpha = 0$ from our simple "super-fast-Fourier" transform relationship. This is classical white-frequency noise being measured for this cesium-beam atomic clock. For the longest averaging times, one observes a $\boldsymbol{\tau}^{0}$ behavior, which then corresponds to $\alpha = -1$, and this is due to the flicker-noise FM of the precision, quartz-crystal oscillator.

In 1965, Bob Vessot brought his hydrogen maser from Boston, MA, USA, Harry Peters brought his hydrogen maser from NASA Goddard, Beltsville, MD, USA, and Len Cutler brought his Hewlett Packard commercial-cesium-beam atomic clock. These devices were compared to the primary frequency standard at NBS/Boulder. This grand-clock-comparison effort resulted in an interesting 12-author paper [11]. This was one of the first papers illustrating the usefulness of ADEV for clock characterization. The AT1 time-scale algorithm [12], which was also developed by Allan at that time, was a major application of the AVAR. Although it has been modified several times since then, the basic algorithm is still used to generate the official time scales at NIST. The Echelle Atomique Libre (EAL) time scale used the BIPM, the International Bureau of Weights and Measures, as the first step in the computation of international atomic time (TAI) and coordinated universal time (UTC) [13] was based on similar principles. In both cases, the weight of a clock in the ensemble average is determined by its prediction error over some previous averaging time—a quantity directly related to the AVAR.

VIII. Conclusion

With 50 years of experience in the time and frequency community, the AVAR and its cousins (MVAR and TVAR) have matured significantly. Over those years, one observes the use of these variances increasing in several other areas of metrology—especially in navigation and telecommunications.

A number of important milestones stand out in the improvements and changes over the years. First, for 16 years, the time and frequency community lived with the ambiguity problem with the AVAR when ADEV, its square-root, varies as $\boldsymbol{\tau}^{-1}$, which means that it is not possible to distinguish between white-noise phase modulation and flicker-noise phase modulation. That ambiguity problem was resolved in 1981 with the development of the modified AVAR, which allowed us to modulate the bandwidth in the software. We see this ambiguity problem persisting, unnecessarily, in other areas of metrology. In the case of quantization errors, MDEV allows to average the noise down as $\boldsymbol{\tau}^{-3/2}$, which then allows the observation of other noise types and instability problems more quickly. MDEV is also the optimum averaging technique for such errors. In addition, if one desires to estimate the average frequency with white-noise residuals present, then a linear regression on the slope improves the knowledge of the slope as $N^{-3/2}$, where N is the number of data points in the regression analysis and is the optimum estimate of the slope for exactly the same reason that MDEV improves as $\boldsymbol{\tau}^{-3/2}$. Second, after the quantization errors are averaged down, ADEV works well and is an efficient

metric for characterizing the intermediate and long-term instabilities. But because the AVAR is Chi-squared distributed, when there are too few degrees-of-freedom for the longest averaging times available from the data [5], then the ADEV values are often too small. This problem has, in large measure, been solved by David A. Howe and his group at NIST, Boulder, CO, USA, with "Theo BR" used in calculating ADEV for extended averaging times [14], [15]. Their work, in a clever way, adds the needed degrees-of-freedom and results in a more efficient use of the data. Third, systematic errors are often hard to deal with. It is generally useful to subtract the systematics from the data, as much as is reasonable, before analyzing the residuals for their noise characteristics. (Diurnal frequency fluctuations, usually driven by the variation in the ambient temperature, are a common systematic that is often removed before a statistical analysis is performed.) This practice is usually done after the fact, but can be done in real time with proper filter functions and prediction algorithms that will estimate and remove the systematics. Since optimum estimation procedures depend upon the kind of noise, this problem can be solved recursively or from some prior knowledge of the noise characteristics of a given system. The principle of parsimony dictates that we use the simplest and most efficient metric in our noise analysis. ADEV satisfies that requirement in many areas of metrology, and this seems to be the main reason for it becoming as widely used as it is.

Finally, we should mention the development of the dynamic AVAR [16], which is useful when the time-differences are only approximately stationary. This method is based on a computation of the variance of consecutive blocks of data and is particularly useful in detecting time steps in the data or changes in the amplitude of the noise.

REFERENCES

[1] J. A. Barnes and D. W. Allan, "Effects of long-term stability on the definition and measurement of short-term stability," in *Proc. IEEE-NASA Symp. Definition Meas. Short-Term Freq. Stab.*, 1964, pp. 119–123.

[2] J. M. Lighthill, *Introduction to Fourier Analysis and Generalized Functions*. Cambridge, U.K.: Cambridge Univ. Press, 1958.

[3] J. A. Barnes and D. W. Allan. (1965). "Two papers on the statistics of precision frequency generators," Nat. Bur. Stand., Boulder, CO, USA, Tech. Rep. 8878 [Online]. Available: http://tf.nist.gov.

[4] D. W. Allan, "Statistics of atomic frequency standards," M.S. thesis, Univ. Colorado Lib., Boulder, CO, USA, Call Number T1965.Al52, 1965.

[5] C. A. Greenhall, "Recipes or degrees of freedom of frequency stability estimators," *IEEE Trans. Instrum. Meas.*, vol. 40, no. 6, pp 994–999, Dec. 1991.

[6] D. W. Allan and J. A. Barnes, "A modified Allan variance with increased oscillator characterization ability," in *Proc. IEEE 35th Annu. Freq. Control Symp.*, 1981, pp. 470–475.

[7] *Standard Teminology for Fundamental Frequency and Time Metrology*, IEEE Standard 1139-1988, 1988.

[8] W. J. Riley. (2008). *Handbook of Frequency Stability Analysis*, National Institute of Standards and Technology Special Publication 1065, Boulder, CO, USA. [Online]. Available: http://tf.nist.gov.

[9] D. W. Allan, M. A Weiss, and J. L. Jesperson, "A frequency-dmain view of time-domain charaterization of clocks and time and frequency distribution systems," in *Proc. IEEE 45th Annu. Symp. Freq. Control*, 1991, pp. 667–678.

[10] D. A. Howe, "A total estimator of the Hadamard function used for gps operations," in *Proc. 32nd Precise Time Time Interval Plann. Appl. Meeting*, Nov. 29, 2000, pp. 255–268.

[11] R. E. Beehler *et al.*, "An intercomparison of atomic standards," *Proc. IEEE*, vol. 54, no. 2, pp. 301–302, Feb. 1966.

[12] J. Levine and D. W. Allan, "The steering of a real time clock to UTC (NBS) and to UTC," in *Proc. 3rd Int. Time Scale Algorithm Symp.*, 1988, pp. 239–254.

[13] J. Azoubib, M. Granveaud, and B. Guinot, "Estimation of the scale unit duration of time scales," *Metrologia*, vol. 13, pp. 87–93, 1977.

[14] D. A. Howe, "An extension of the Allan variance with increased confidence at long term," in *Proc. IEEE Int. Freq. Control Symp.*, 1995, pp. 337–346.

[15] D. A. Howe, "Total variance explained," in *Proc. IEEE Joint Meeting. Int. Freq. Control. Symp. and Eur. Freq. Time Forum (EFTF) Conf.*, 1999, pp. 1113–116.

[16] L. Galleani and P. Tavella, "The dynamic allan variance," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 56, no. 3, pp. 450–464, Mar. 2009.

[17] D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls. (1999). *Characterization of Clocks and Oscillators*, NIST Tech. Note 1337 [Online]. Available: http://tf.nist.gov.

[18] C. Hackman and D. B. Sullivan, "Time and frequency measurement," *Amer. J. Phys.*, vol. 63, pp. 306–317, 1995.

David W. Allan was born September 25, 1936, in Mapleton, UT, USA. He received the B.S. degree in physics from Brigham Young University and the M.S. degree in physics from the University of Colorado.

His master's thesis is the basis for the Allan variance and was published in the February PROCEEDINGS OF THE IEEE in 1966. During his 32 years in Boulder, CO, USA, he helped develop the United States official time and frequency standards and with colleagues published over 100 papers in that process. He also consulted for the APOLLO, GPS, and JPL deep-space timing programs. He believes that cooperation rather than competition brings about the greatest good and feels greatly blessed to have worked with such outstanding colleagues. He retired from the government in 1992, and has remained active in the time and frequency community as an independent researcher and consultant.

Judah Levine (A'02–M'09) received the Ph.D. degree in physics from New York University, New York, NY, USA, in 1966.

He is a Fellow of the National Institute of Standards and Technology and is the leader of the Network Synchronization Project in the Time and Frequency Division, located in Boulder, CO, USA. He is responsible for the design and implementation of the time scales AT1 and UTC (NIST), which provide the reference signals for all of the NIST time and frequency services. In addition, he designed and built the servers that support the Automated Computer Time Service (ACTS) and the Internet Time Service, which provide time and frequency information in a number of different digital formats. These services receive about 16 500 million requests per day.

Dr. Levine is a Fellow of the American Physical Society.