The Generalized Autocovariance: A Tool for Clock Noise Statistics

C. A. Greenhall[1]

The generalized autocovariance (GACV) function is a natural extension of the autocovariance function from stationary processes to a class of nonstationary processes used as models of clock noise. This article gives a theory of the GACV and shows how to use it to carry out useful statistical calculations.

I. Introduction

Despite Hamming's maxim that the purpose of computing is insight, not numbers, the purpose of this article is to explain how to perform certain calculations on models of noise and instability in frequency standards and signal-processing equipment. Random processes with stationary nth increments, whose correlation theory was developed by Yaglom [1], have been recognized as appropriate models for clock noise [2]. Using these models, the author has been able to exploit a systematic method for computing means and variances of clock-noise statistics, including estimators of frequency stability and drift [3–8]. The present article is intended to explain this method in sufficient depth and detail that other workers might be able to use it for similar calculations.

All persons who deal with theoretical aspects of random noise know about stationary processes (see Section II.C), along with the familiar concepts of spectral density and autocovariance (ACV, also called autocorrelation). Loosely speaking, we can regard a process $x(t)$ with stationary nth increments as the nth integral of a stationary process $z(t)$, and derive properties of $x(t)$ in terms of properties of $z(t)$. Often, though, it is simpler to deal with $x(t)$ standing on its own, as it were; for example, we might ascribe to $x(t)$ a spectral density of form $a|\nu|^{\beta}$, say, where ν is frequency and $\beta \leq -1$. Although this spectrum is non-integrable in a neighborhood of $\nu = 0$, certain variance and covariance calculations based on this spectrum still can be performed in the frequency domain if they involve filters whose frequency responses cancel the spectral divergence at low frequencies. For the case of a stationary $x(t)$, whose spectrum is integrable, we could do equivalent calculations in the time domain using the ACV, which is the inverse Fourier transform of the spectrum. Faced with a non-integrable spectrum, we must inquire carefully whether there might be some function of one time variable that plays a similar role. The author argues that the natural candidate for such a role is the generalized autocovariance (GACV), obtained by applying a certain form of generalized inverse Fourier transform to the spectrum [see Eqs. (2) and (21)]; this transform uses a reduced form of the complex exponential that also cancels the low-frequency divergence. Although this concept really belongs to the theory of tempered Schwarz distributions [9,10], the necessary results are developed here without calling on that theory.

[1] Tracking Systems and Applications Section.

At this point, a note of caution is appropriate. Without a consistent theory for guidance, one can be seriously misled by uncritical use of Fourier transform tables in an attempt to produce an ACV that corresponds to a non-integrable spectrum. One table ([11], citing [12]) gives the transform of $|2\pi\nu|^{-1}$ (a flicker noise spectrum) as $\sqrt{2\pi}\,|t|^{-1}$ in the Fourier-transform convention adopted here. According to the theory given below, the GACV for this noise is actually $-(1/\pi)\ln|t|$ [in a limiting sense because of high-frequency spectral divergence; see Eq. (46)]. Having been unable to find a derivation of the tabulated transform, the author concludes that either the entry is simply a mistake or there is more than one reasonable way to extend the Fourier transform beyond its initial domain of integrable functions. Although this is the only example of this kind that the author has found, it is especially worthy of note because its presence in a well-known reference work is a trap for the unwary person interested in "$1/f$" noise. For this reason, and because the GACV itself has no obvious physical interpretation as yet, the author feels justified in subjecting the reader to a mathematical exposition that may seem plodding and pedantic, the goal being to establish a set of results that one can depend on. Since there are few surprises here, the reader can skip most of the proofs without sacrificing insight.

Another candidate for replacing the ACV is the nth structure function of Lindsey and Chie [13] [called increment variance here; see Eq. (47)]. Indeed, the spectrum of a process with stationary first increments is determined by either the first or the second increment variance. Recently, however, it was shown that knowledge of the second increment (Allan) variance of a process with stationary second increments does not determine the spectrum [14], and the situation for higher-order increments must be regarded as an open problem. In contrast with this situation, the GACV and spectrum determine each other, and knowledge of either one permits a calculation of all covariances of increments of sufficiently high order.

Following is a summary of the rest of the article. After going through some preliminary definitions and review material, we reach our true starting point, Yaglom's frequency-domain structure theorems (Theorems 2 and 3). From there we develop the time-domain covariance structure theory, derive the GACVs of power-law processes, and give some applications of the theory to clock statistics. Even though all the examples are real-valued, the theory is presented for complex-valued processes, mainly because doing so enforces greater care and consistency. Some of the more pedestrian arguments have been put in the Appendix. The section on applications assumes that the reader is familiar with clock stability measures; all applications but the last are treated briefly without motivation and carried out only to formulas for the answers in terms of the GACV. The last application, estimation of Allan variance with drift removal, is treated at greater length and carried out to numerical results for power-law clock noises.

II. Preliminaries

A. Notations and Conventions

The conjugate of a complex number z is written $\overline{z}$; the adjoint of an operator H is written H^*.

The notation $\mathbf{1}(S)$, where S is some condition, means 1 if S is true and 0 if S is false.

The expectation operator is denoted by E. Its scope in a formula is as far as it can reach; thus, EX^2 means $E\left(X^2\right)$, not $(EX)^2$.

The vector space of polynomials of degree $<n$ with complex coefficients is denoted by P_n. If V is a vector space, then $x \cong y \pmod{V}$ means $x - y \in V$, and we say that x and y are equivalent modulo V.

Frequency (Hz) is denoted by ν. Often we use ω merely as an abbreviation for $2\pi\nu$ in formulas containing ν. A two-sided power-law spectral density of a process $x(t)$ is written as $S_x(\nu) = a_\beta |2\pi\nu|^\beta$. In the time and frequency literature, the corresponding one-sided spectral densities of $x(t)$ and $y(t) = dx/dt$ are written as $S_x^+(\nu) = g_\beta \nu^\beta$ and $S_y^+(\nu) = h_{\beta+2}\nu^{\beta+2}$. Therefore,

$$\left.\begin{aligned} g_\beta &= 2\,(2\pi)^\beta\, a_\beta \\ h_{\beta+2} &= 2\,(2\pi)^{\beta+2}\, a_\beta \end{aligned}\right\} \tag{1}$$

In Section IV, however, we set $a_\beta = 1$ for convenience.

An integral $\int_a^b f(t)\,x(t)\,dt$, where $f(t)$ is piecewise continuous and $x(t)$ is a mean-square continuous random process (see Section II.C), is interpreted as a Riemann quadratic-mean stochastic integral, the mean-square limit of Riemann sums $\sum f(t_j)\,x(t_j)\,\Delta t_j$. For a discussion of these integrals, see Cramér and Ledbetter [15, p. 85 ff.]; we shall avoid belaboring this point here.

The following notation will be used later in the generalized Fourier transform that gives the GACV [see Eq. (21)]. For complex z, define

$$\left.\begin{aligned} e_0(z) &= e^z \\ e_n(z) &= e^z - \sum_{k=0}^{n-1} \frac{z^k}{k!}, \quad n \geq 1 \end{aligned}\right\} \tag{2}$$

Then $e_n'(z) = e_{n-1}(z)$, and $e_n(z) = O(|z|^n)$ as $z \to 0$; thus, $e_n(i\omega t)$ will serve as a substitute for $e^{i\omega t}$ that cancels low-frequency spectral divergence. For proofs we shall require the estimate

$$|e_n(z)| \leq \frac{|z|^n}{n!} \max\left(1, e^{\operatorname{Re} z}\right) \tag{3}$$

which we now prove by induction on n. For $n = 0$ it is obvious. Assuming it for $n-1$, we have

$$e_n(z) = \int_0^z e_{n-1}(\zeta)\,d\zeta = z\int_0^1 e_{n-1}(zt)\,dt$$

$$|e_n(z)| \leq |z|^n \int_0^1 \frac{t^{n-1}}{(n-1)!} \max\left(1, e^{t\operatorname{Re} z}\right) dt \leq \frac{|z|^n}{n!} \max\left(1, e^{\operatorname{Re} z}\right)$$

B. Time-Limited Filters

The following class of linear time-invariant (LTI) filters covers most of our needs here.

Definition 1. *A time-limited (TL) filter is an LTI filter whose impulse response is a time-limited piecewise-continuous function plus a linear combination of a finite number of delta functions.*

Thus, the result of applying a TL filter H to a continuous[2] function $x(t)$ is the continuous function

$$Hf(t) = \int_{-\infty}^{\infty} a(u)\,x(t-u)\,du + \sum_j a_j x(t-t_j) \tag{4}$$

[2] This includes mean-continuous random processes.

Here the sum is finite, and $a(t)$ is a piecewise-continuous function that vanishes outside some unspecified bounded interval; we write $\int_{-\infty}^{\infty}$ to avoid having to specify this interval. The impulse response is $h(t) = H\delta(t) = a(t) + \sum_j a_j \delta(t - t_j)$, actually a bounded complex measure, and $\int_{-\infty}^{\infty} h(u)\, x(t-u)\, du$ is used as shorthand for the right side of Eq. (4). Thus, we use $h(t)\, dt$ in integrals instead of a measure-theoretic notation like $d\mu(t)$ or $\mu(dt)$. A discrete TL filter is one whose impulse response consists only of delta functions. A finite-impulse-response (FIR) filter is a discrete TL filter whose delta functions are concentrated at integer multiples of some sample period τ_0.

The transfer function is defined in the shorthand notation by

$$H(\nu) = \int_{-\infty}^{\infty} e^{-i2\pi\nu t} h(t)\, dt$$

The same symbol H is used both for the filter as an operator on functions and for the transfer function as a function of frequency ν.

Important FIR filters are the shift B_τ, defined by $B_\tau \delta(t) = \delta(t - \tau)$, and the backward τ-difference $\Delta_\tau = B_0 - B_\tau$. Then $B_\tau(\nu) = e^{-i2\pi\nu\tau}$ and $B_{\tau_1} B_{\tau_2} = B_{\tau_1+\tau_2}$. The nth τ-difference Δ_τ^n satisfies

$$\Delta_\tau^n = (B_0 - B_\tau)^n = \sum_{k=0}^{n} \binom{n}{k} (-1)^k B_{k\tau}$$

$$\Delta_\tau^n(\nu) = \left(1 - e^{-i2\pi\nu\tau}\right)^n$$

Two TL filters with the same transfer function are the same, because bounded complex measures with the same Fourier transform are equal. The set of TL filters is closed under the commutative operations of addition and operator composition, which corresponds to convolution of impulse responses and multiplication of transfer functions.

If H has impulse response $h(t)$, its adjoint H^* is defined to be the filter with impulse response $h^*(t) = \overline{h(-t)}$. Thus, $H^*(\nu) = \overline{H(\nu)}$. For example, $B_\tau^* = B_{-\tau}$, $(\Delta_\tau^n)^* = \Delta_{-\tau}^n$. The filter HH^* is self-adjoint; for example,

$$\begin{aligned} \Delta_\tau^n (\Delta_\tau^n)^* &= (\Delta_\tau \Delta_{-\tau})^n = [(1 - B_\tau)(B_\tau - 1) B_{-\tau}]^n = (-1)^n (1 - B_\tau)^{2n} B_{-n\tau} \\ &= (-1)^n \Delta_\tau^{2n} B_{-n\tau} = (-1)^n \delta_\tau^{2n} \end{aligned} \tag{5}$$

where δ_τ^{2n} is the $2n$th central τ-difference operator. Expanding $(1 - B_\tau)^{2n}$ gives

$$\begin{aligned} \Delta_\tau^n (\Delta_\tau^n)^* &= (-1)^n \sum_{k=0}^{2n} \binom{2n}{k} (-1)^k B_{(k-n)\tau} \\ &= \sum_{j=-n}^{n} \binom{2n}{n+j} (-1)^j B_{j\tau} \end{aligned} \tag{6}$$

1. Order. The transfer function of a TL filter H is an entire analytic function of the complex variable ν. It is convenient to express its Taylor expansion about $\nu = 0$ in the form

$$H(\nu) = \sum_{k=0}^{\infty} \lambda_k(H) (i2\pi\nu)^k$$

where

$$\lambda_k(H) = \frac{H^{(k)}(0)}{k!(i2\pi)^k} = \int_{-\infty}^{\infty} h(t) \frac{(-t)^k}{k!} dt$$

The numbers $\lambda_k(H)$ will be called the Taylor coefficients of the filter H. The adjoint H^* has Taylor coefficients $\lambda_k(H^*) = (-1)^k \overline{\lambda_k(H)}$.

Definition 2. *A TL filter H has* order n *($n \geq 0$) if $\lambda_n(H)$ is the first nonzero $\lambda_k(H)$.*

A TL filter of order n annihilates P_n. It also reduces the degree of polynomials of degree $\geq n$ by n: for if $m \geq n$, then

$$\int_{-\infty}^{\infty} h(u) \frac{(t-u)^m}{m!} du = \int_{-\infty}^{\infty} h(u) \left[\frac{t^{m-n}}{(m-n)!} \frac{(-u)^n}{n!} + \cdots \right] du$$

$$= \lambda_n(H) \frac{t^{m-n}}{(m-n)!} + \text{lower terms}$$

If the order of H is $\geq k$, then

$$\lambda_k(H) = \lim_{\nu \to 0} \frac{H(\nu)}{(i2\pi\nu)^k} = \int_{-\infty}^{\infty} h(u) \frac{(t-u)^k}{k!} du \quad \text{for all } t \tag{7}$$

The filter Δ_τ^n has order n, and $\lambda_n(\Delta_\tau^n) = \tau^n$.

If G and H have orders m and n, then GH has order mn, and $\lambda_{mn}(GH) = \lambda_m(G)\lambda_n(H)$. If H is a TL filter with all $\lambda_k(H) = 0$, then $H(\nu) = 0$, and so H is the zero filter. (There do exist nonzero LTI filters, not time-limited, that annihilate all polynomials.)

2. Integration and Moving Average. Define the operator J_0 by

$$J_0 f(t) = \int_0^t f(u)\, du \tag{8}$$

The choice of 0 to start the integration is arbitrary. Although J_0 is not an LTI filter, it can be manipulated somewhat as one according to the following rules.

Lemma 1.

(1) If H is a TL filter of order $n \geq 1$ with impulse response $h(t)$, then the operator HJ_0 is a TL filter with piecewise-continuous impulse response $\int_{-\infty}^{t} h(u)\,du$, transfer function $H(\nu)/(i2\pi\nu)$, order $n-1$, and adjoint $-H^*J_0$.

(2) Let H be a TL filter, and let $f(t)$ be a continuous function. Then $J_0Hf(t) = HJ_0f(t) - HJ_0f(0)$. Consequently, if G is a TL filter of order ≥ 1, then $GJ_0H = GHJ_0$.

Proof.

(1) Since $\int_{-\infty}^{\infty} h(t)\,dt = 0$, both $h(t)$ and $h_1(t) = \int_{-\infty}^{t} h(u)\,du$ are supported in some interval $]-T, T[$. Integration by parts gives

$$HJ_0f(t) = \int_{-T}^{T} h(u)\,J_0f(t-u)\,du = \int_{-T}^{T} h_1(u)\,f(t-u)\,du$$

We can also use integration by parts to derive the transfer-function formula, from which the order reduction can be seen. If $-t$ is a point of continuity of h_1, then

$$h_1^*(t) = \int_{-\infty}^{-t} \overline{h(u)}du = -\int_{-t}^{\infty} \overline{h(u)}du \quad \left(\text{since} \int_{-\infty}^{\infty} h(u)\,du = 0\right)$$

$$= -\int_{-\infty}^{t} \overline{h(-v)}dv = -\int_{-\infty}^{t} h^*(v)\,dv$$

which, by the above argument, is the impulse response of $-H^*J_0$.

(2) By Fubini's theorem for the product measure $du \times h(v)\,dv$,

$$J_0Hf(t) = \int_0^t du \int_{-T}^{T} dv\,h(v)\,f(u-v) = \int_{-T}^{T} dv\,h(v) \int_0^t du\,f(u-v)$$

$$= \int_{-T}^{T} dv\,h(v)\,[J_0f(t-v) - J_0f(-v)]$$

❐

For example, the continuous moving-average filter A_τ defined by

$$A_\tau x(t) = \frac{1}{\tau}\int_{t-\tau}^{t} x(u)\,du \tag{9}$$

satisfies

$$A_\tau = \frac{1}{\tau}\Delta_\tau J_0 \tag{10}$$

$$A_\tau \delta(t) = \frac{1}{\tau}\mathbf{1}(0 < t < \tau) \tag{11}$$

$$A_\tau(\nu) = \frac{\Delta_\tau(\nu)}{i2\pi\nu\tau} = e^{-i\pi\nu\tau}\frac{\sin(\pi\nu\tau)}{\pi\nu\tau} \tag{12}$$

$$A_\tau^* = B_{-\tau}A_\tau \tag{13}$$

C. Stationary Processes

We review some elementary facts about covariance-stationary processes [15,16]. A complex-valued random process $x(t)$ is called mean-square continuous if, for all t, $E|x(t)|^2 < \infty$ and $E|x(s) - x(t)|^2 \to 0$ as $s \to t$. Such a process is called covariance-stationary if there is a constant c and function $R_x(t)$ such that $Ex(s) = c$ and $Ex(s+t)\overline{x(s)} = R_x(t)$ for all s and t. Even though the means are not subtracted off in this product, it is customary to call $R_x(t)$ the autocovariance (ACV) function of $x(t)$. There is a positive measure (distribution of power) $S_x(d\nu)$ on the two-sided frequency axis, the spectrum of $x(t)$, such that $\int_{-\infty}^{\infty} S_x(d\nu) < \infty$ and

$$R_x(t) = \int_{-\infty}^{\infty} e^{i2\pi\nu t} S_x(d\nu)$$

For applications, we assume that the spectrum is the sum of a non-negative integrable function $w(\nu)$, the spectral density, and a linear combination $\sum_{j=1}^{\infty} w_j \delta(\nu - \nu_j)$ of delta functions (bright lines), where $w_j > 0$, $\sum_j w_j < \infty$. From here on we write $S_x(\nu)\,d\nu$ instead of $S_x(d\nu)$, just as we write $h(t)\,dt$ in integrals involving an impulse response that might contain delta functions.

It is also convenient to make a harmless ergodic assumption. It can be shown that $(1/T)\int_0^T x(t)\,dt$ tends in mean square as $T \to \infty$ to a random variable Z such that $EZ = c$ and $E|Z|^2$ is the power of the spectral bright line at $\nu = 0$ (DC). Because no observer of one instance of the process over all t can tell the difference between the infinite-time average Z and the ensemble average c, we might as well assume that Z is the deterministic constant c. Then the spectral bright line at DC is $|c|^2\delta(\nu)$. The process $x_0(t) = x(t) - c$, which has mean zero and ACV $R_x(t) - |c|^2$, has the same spectrum as $x(t)$ except with no DC bright line.

If $x(t)$ is real-valued, then $R_x(t) = R_x(-t)$, $S_x(\nu) = S_x(-\nu)$, and $2S_x(\nu)$, $\nu > 0$, is called the one-sided spectrum; the bright line at DC (if any) is not doubled.

From here on, "stationary" means "covariance-stationary."

The following theorem shows how to compute means and covariances of random variables that are linear functionals of $x(t)$ via TL filters.

Theorem 1. *Let $x(t)$ be a complex-valued stationary process with mean c, spectrum $S_x(\nu)$, and ACV $R_x(t)$. If H is a TL filter, then $Hx(t)$ is stationary and $EHx(s) = cH(0)$. For any TL filters G and H,*

$$EGx(s+t)\overline{Hx(s)} = \int_{-\infty}^{\infty} e^{i2\pi\nu t} G(\nu)\overline{H(\nu)} S_x(\nu)\,d\nu \tag{14}$$

$$= GH^* R_x(t) \tag{15}$$

When written out, the right side of Eq. (15) becomes

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(u)\,\overline{h(v)}R_x(t-u+v)\,dudv$$

Except for some argument about interchanging the E operator with quadratic-mean integrals, the proof of Theorem 1 consists of routine manipulations. The purpose here is to extend this theorem, especially Eq. (15), to a class of processes used for models of phase noise.

III. Processes With Stationary *n*th Increments

The definition that follows is a simplification of the definition given by A. M. Yaglom [1], who developed the correlation theory of these processes as a generalization of Kolmogorov's theory of processes with stationary first increments.

Definition 3. *A complex-valued mean-square continuous random process $x(t)$* has *stationary nth increments if for each real τ the process $\Delta_\tau^n x(t)$ is stationary.*

It is convenient to regard a stationary process as having stationary 0th increments.

Yaglom's first main result is a structure theorem for the expectations and correlations of the nth increments in terms of a spectrum that can have infinite power at low frequencies. The notation used here differs somewhat from his.

Theorem 2. *Let $x(t)$ have stationary nth increments. There are unique numbers c_n, w and a unique positive locally bounded measure $S_x(d\nu)$ on the punctured real axis $\nu \neq 0$ such that $|c_n|^2 \leq w$,*

$$E\Delta_\tau^n x(t) = c_n\tau^n \tag{16}$$

$$\int_{-1}^{1}\nu^{2n}S_x(d\nu) + \int_{|\nu|>1} S_x(d\nu) < \infty \tag{17}$$

$$E\Delta_{\tau_1}^n x(s+t)\,\overline{\Delta_{\tau_2}^n x(s)} = w\tau_1^n\tau_2^n + \int_{-\infty}^{\infty} e^{i2\pi\nu t}\Delta_{\tau_1}^n(\nu)\,\overline{\Delta_{\tau_2}^n(\nu)}S_x(d\nu) \tag{18}$$

Conversely, if $|c_n|^2 \leq w$, and the positive measure $S_x(d\nu)$ satisfies Eq. (17), then there exists a process $x(t)$ with stationary nth increments such that Eqs. (16) and (18) hold.

The cutoff $\nu = 1$ for the integrals in Eq. (17) is arbitrary: the reader can think of it as 1 Hz or replace it by any positive frequency. The requirement in Eq. (17) of finite power at high frequencies is the condition that restricts us to ordinary processes instead of generalized processes.

The measure $S_x(d\nu)$ is called the spectrum of $x(t)$. As with stationary processes, we write $S_x(\nu)\,d\nu$ in place of $S_x(d\nu)$ and assume for convenience that $S_x(\nu)$ is the sum of a locally integrable function and a linear combination of delta functions. Likewise, since it can be shown that $(1/T)\int_0^T \Delta_1^n x(t)\,dt$ tends in mean-square to a random variable Z such that $EZ = c_n$, $E|Z|^2 = w$, we assume that Z is deterministic, so that $Z = E\Delta_1^n x(t) = c_n$ and $w = |c_n|^2$. This will be called the "deterministic drift assumption." We call $c_n = c_n(x)$ the drift rate of degree n. Then the drift rate of the process $x(t) - c_n t^n/n!$ vanishes.

Note, however, that "deterministic" is not the same as "known": often we can only estimate c_n from the values of $x(t)$ on some interval (see Sections V.C and V.D). If $x(t)$ is real-valued, then $S_x(\nu)$ is even, and the one-sided spectrum is given by $S_x^+(\nu) = 2S_x(\nu)$ for $\nu > 0$.

Yaglom gives an analog of the first part of Theorem 1; with our present restrictions, assumptions, and notation, it reads as follows.

Theorem 3. *Let $x(t)$ have stationary nth increments with drift rate c_n and spectrum $S_x(\nu)$. If H is a TL filter of order $\geq n$, then $Hx(t)$ is stationary and $EHx(t) = c_n\lambda_n(H)$. For any TL filters G and H of orders $\geq n$,*

$$EGx(s+t)\,\overline{Hx(s)} = \lambda_n(G)\,\overline{\lambda_n(H)}\,|c_n|^2 + \int_{-\infty}^{\infty} e^{i2\pi\nu t} G(\nu)\,\overline{H(\nu)} S_x(\nu)\,d\nu \tag{19}$$

In particular, the spectrum of the stationary process $Hx(t)$ is $|\lambda_n(H)\,c_n|^2\,\delta(\nu) + |H(\nu)|^2\,S_x(\nu)$.

A. Structure Theorem, Time Domain

Theorem 3 is the point of departure of this study. Missing from it is an analog of Eq. (15) for stationary processes. The following theorem says that there is always a function that plays the role of the ACV, but only for filters of order $\geq n$.

Theorem 4. *Let $x(t)$ have stationary nth increments ($n \geq 1$) with drift rate c_n and spectrum $S_x(\nu)$. There exists a (non-unique) continuous function $R_x(t)$ such that*

$$EGx(s+t)\,\overline{Hx(s)} = GH^*R_x(t) \tag{20}$$

for all TL filters G and H of orders $\geq n$. Such a function is given by

$$R_x(t) = |c_n|^2(-1)^n\frac{t^{2n}}{(2n)!} + \int_{-1}^{1} e_{2n}(i2\pi\nu t)\,S_x(\nu)\,d\nu + \int_{|\nu|>1} e^{i2\pi\nu t} S_x(\nu)\,d\nu \tag{21}$$

(See Eq. (2) for the definition of $e_{2n}(z)$.) In particular, the ordinary ACV of the stationary process $Hx(t)$ is $HH^*R_x(t)$; in this connection, see the remark after Theorem 3.

Before proving this theorem, we establish some terminology and make some remarks.

Definition 4. *Let $x(t)$ have stationary nth increments. A continuous function $R_x(t)$ that satisfies*

$$E\Delta_{\tau_1}^n x(s+t)\,\overline{\Delta_{\tau_2}^n x(s)} = \Delta_{\tau_1}^n\Delta_{-\tau_2}^n R_x(t) \tag{22}$$

for all s, t, τ_1, and τ_2 is called a generalized autocovariance (GACV) function for $x(t)$.

According to Theorem 4, the function $R_x(t)$ given by Eq. (21) is a GACV because it satisfies Eq. (20), a generalization of Eq. (22). To what extent is the GACV unique? If we add a polynomial of degree $<2n$ to a GACV, then Eq. (20) still holds for all G and H with orders $\geq n$ because GH^* has order $\geq 2n$. Conversely, any function $f(t)$ that is the difference of two GACVs is a continuous function that satisfies $\Delta_\tau^n\Delta_{-\tau}^n f(t) = 0$ for all t and τ, which is the same as saying $\Delta_\tau^{2n} f(t) = 0$ for all t and τ. By Lemma A-1 in Section I of the Appendix, $f(t) \in P_{2n}$. Consequently, *any two GACVs are equivalent modulo P_{2n}*. In

particular, changing the cutoff in the integrals in Eq. (21) from 1 to some other frequency only adds a member of P_{2n} to $R_x(t)$.

If $n = 1$ and $S_x(-\nu) = S_x(\nu)$, then

$$-|c_1|^2 \frac{t^2}{2} + \int_{-\infty}^{\infty} [\cos(2\pi\nu t) - 1] S_x(\nu)\, d\nu \tag{23}$$

is a GACV because it differs by only a constant from Eq. (21).

An additional ambiguity occurs because a process with stationary nth increments also has stationary increments of any order $m > n$. If such an m is gratuitously substituted for n in Eq. (21), then a polynomial $p(t) \in P_{2m}$ is added to $R_x(t)$. In this situation, however, it is understood that we are only to apply filters G, H of order $\geq m$ to $x(t)$; thus, GH^* annihilates the offending polynomial $p(t)$, along with the original drift portion of $R_x(t)$.

Proof of Theorem 4.

(1) We assert that $R_x(t)$ as given by Eq. (21) is continuous. Let $R_1(t)$ and $R_2(t)$ be the two integrals in Eq. (21). If $|t| \leq 1$, then $|e_{2n}(i\omega t)| \leq \omega^{2n}/(2n)!$ by Eq. (3). Let $t_j \to t$. By Eq. (17) and Lebesgue's theorem on dominated convergence, $R_1(t_j) \to R_1(t)$ and $R_2(t_j) \to R_2(t)$.

(2) According to Theorem 3, we want to prove that $GH^* R_x(t)$ equals the right side of Eq. (19). Because $GH^*\left(t^{2n}/(2n)!\right) = \lambda_n(G)\lambda_n(H^*) = (-1)^n \lambda_n(G)\overline{\lambda_n(H)}$, we can assume that $c_n = 0$. Let $F = GH^*$, a TL filter with impulse response $f(t)$ and transfer function $F(\nu) = G(\nu)\overline{H(\nu)}$. Then,

$$FR_1(t) = \int_{-\infty}^{\infty} du\, f(u) \int_{-1}^{1} d\nu\ S_x(\nu)\, e_{2n}(i\omega(t-u)) \tag{24}$$

By Eq. (3), $|e_{2n}(i\omega(t-u))| \leq (\omega(t-u))^{2n}/(2n)!$. Since

$$\int du\ |f(u)|\,(t-u)^{2n} \int_{-1}^{1} d\nu\ S_x(\nu)\,\nu^{2n} < \infty$$

we may interchange orders of integration in Eq. (24). For fixed ω, the filter F, because it has order $\geq 2n$, annihilates the polynomial part of $e_{2n}(i\omega t)$, so that

$$FR_1(t) = \int_{-1}^{1} d\nu\ S_x(\nu) \int_{-\infty}^{\infty} du\, f(u)\, e^{i\omega(t-u)} = \int_{-1}^{1} e^{i\omega t} F(\nu)\, S_x(\nu)\, d\nu$$

Similarly, $FR_2(t) = \int_{|\nu|>1} e^{i\omega t} F(\nu)\, S_x(\nu)\, d\nu$. Thus,

$$FR_x(t) = FR_1(t) + FR_2(t) = \int_{-\infty}^{\infty} e^{i\omega t} F(\nu)\, S_x(\nu)\, d\nu$$

❒

If we know $|c_n|^2$ and $S_x(\nu)$, then we know $R_x(\nu)$ from Eq. (21). Conversely, if we know a GACV $R_x(t)$, then we know the left side of Eq. (22) for all s, t, τ_1, and τ_2. By Theorem 2 and the ergodic drift assumption, $|c_n|^2$ and $S_x(\nu)$ are uniquely determined. We record this situation as a theorem.

Theorem 5. *Let $x(t)$ have stationary nth increments. Knowledge of the spectrum $S_x(\nu)$ and squared drift rate $|c_n|^2$ is equivalent to knowledge of the GACV, modulo P_{2n}.*

From here on, $R_x(t)$ denotes any GACV of $x(t)$. There is always a GACV, namely, the version given by Eq. (21), that satisfies $R_x(-t) = \overline{R_x(t)}$ and is real if $x(t)$ is real. Unless $n = 0$, we can subtract a constant to make $R_x(0) = 0$.

Equation (21) for a GACV, although useful for theory, is somewhat artificial and might not be the easiest way to derive a closed formula for $R_x(t)$ in particular situations. A version that is often more useful, especially for power-law spectra, is given in the following theorem, which obtains a GACV by means of partial Fourier transforms in the complex time domain.

Theorem 6. *Let $x(t)$ have stationary nth increments ($n \geq 1$) with zero drift rate and spectrum $S_x(\nu)$. Fix a positive integer $m \leq 2n$ such that $\int_{-1}^{1} |\nu|^m S_x(\nu)\, d\nu < \infty$. Let*

$$B^+(z) = \int_0^\infty e^{i2\pi\nu z} (i2\pi\nu)^m S_x(\nu)\, d\nu, \quad \text{Im } z > 0 \tag{25}$$

$$B^-(z) = -\int_{-\infty}^0 e^{i2\pi\nu z} (i2\pi\nu)^m S_x(\nu)\, d\nu, \quad \text{Im } z < 0 \tag{26}$$

and let $R^+(z)$ and $R^-(z)$ be any mth indefinite integrals of $B^+(z)$ and $B^-(z)$ on these half-planes, i.e., any functions whose mth derivatives are $B^+(z)$ and $B^-(z)$, respectively. Then $R^+(z)$ and $R^-(z)$ extend continuously to the real line, and $R^+(t+i0) - R^-(t-i0)$ is a GACV for $x(t)$.

If $x(t)$ is real-valued, then both $R^+(t+i0) + R^+(-t+i0)$ and $2 \text{ Re } R^+(t+i0)$ are GACVs *for $x(t)$.*

By Theorem 2, we can always take $m = 2n$; often a smaller m is possible and desirable.

Before proving this theorem, we give an important and familiar example. Let $S_x(\nu) = (2\pi\nu)^{-2}$, the spectrum of one-dimensional unit Brownian motion, a real-valued process $x(t)$ with stationary first increments that satisfies $E[\Delta_\tau x(t)]^2 = |\tau|$. Setting $m = 2$, we have

$$B^+(z) = \frac{-1}{2\pi} \int_0^\infty e^{i\omega z} d\omega = \frac{1}{(i2\pi z)}, \quad \text{Im } z > 0$$

Integrating twice gives $R^+(z) = (z \ln z - z)/(i2\pi)$, and we can throw out the z term since its second derivative is zero. We shall use the form $R_x(t) = R^+(t+i0) + R^+(-t+i0)$, which is an even function. Let $t > 0$. Choosing arbitrarily the branch of $\ln z$ that is real on the positive real axis and cut on the negative real axis, we have

$$R_x(t) = \frac{1}{i2\pi} [t \ln t + (-t)(\ln t + i\pi)] = -\frac{t}{2}, \quad t > 0$$

Thus, the GACV of unit Brownian motion is given by

$$R_x(t) = -\frac{1}{2}|t| \tag{27}$$

for all t.

The form 2 Re $R^+(t+i0)$ leads to another solution, namely, $R_x(t) = 0$ for $t \geq 0$ and $R_x(t) = t$ if $t < 0$, which is equivalent to Eq. (27) modulo P_2. Although the minus sign in $-(1/2)|t|$ might seem strange, it is correct, as we see from the computation

$$E\left[\Delta_\tau x(t)\right]^2 = \Delta_\tau \Delta_{-\tau} R_x(0) = 2R_x(0) - R_x(\tau) - R_x(-\tau) = |\tau|$$

We can also understand how the negative GACV comes about by approximating the nonstationary process $x(t)$ (in a sense to be explained) by stationary processes with greater and greater power at low frequencies. To avoid messy algebra, we venture outside the class of time-limited filters without proving that it is all right to do so. For $a > 0$, let H_a be the order-1 filter with impulse response $\delta(t) - ae^{-at}\mathbf{1}(t > 0)$. Then $H_a(\nu) = i\omega/(a + i\omega)$. The process $x(t;a) = H_a x(t)$ is a stationary process with Lorentzian spectrum $S_x(\nu;a) = |H_a(\nu)|^2 S_x(\nu) = \left(a^2 + \omega^2\right)^{-1}$ and ACV $R_x(t;a) = (1/2a)e^{-a|t|}$, which has a cusp at $t = 0$. As $a \to 0$, $S_x(\nu;a)$ converges to $\omega^{-2} = S_x(\nu)$ and $R_x(t;a)$ diverges, but $R_x(t;a) - R_x(0;a) = (1/2a)\left(e^{-a|t|} - 1\right)$ does converge to $-(1/2)|t| = R_x(t)$. When $|t| \ll 1/a$, $R_x(t;a)$ looks like $-(1/2)|t|$ relative to its value at $t = 0$.

In a similar way, as $a \to 0$, the random variable $x(t;a)$, whose variance is $(2a)^{-1}$, does not converge to anything, but, with the help of Theorem 3 or Theorem 4, one can show that the "calibrated" process $x(t;a) - x(0;a)$ does converge in the mean-square sense to $x(t) - x(0)$.

Proof of Theorem 6. We prove that $R^+(t+i0) - R^-(t-i0)$ is equivalent to $R_x(t)$ of Eq. (21) modulo P_{2n} and is, therefore, a GACV for $x(t)$. Let

$$R_1^+(z) = \int_0^1 e_m(i\omega z) S_x(\nu)\, d\nu$$

$$R_1^-(z) = -\int_{-1}^0 e_m(i\omega z) S_x(\nu)\, d\nu$$

$$R_2^+(z) = \int_1^\infty e^{i\omega z} S_x(\nu)\, d\nu, \quad \text{Im } z \geq 0$$

$$R_2^-(z) = -\int_{-\infty}^{-1} e^{i\omega z} S_x(\nu)\, d\nu, \quad \text{Im } z \leq 0$$

Since $m \leq 2n$,

$$R_x(t) \cong R_1^+(t) - R_1^-(t) + R_2^+(t) - R_2^-(t) \quad (\text{mod} P_{2n})$$

Assume the following statements, the proofs of which are routine exercises (Section II of the Appendix).

(1) $R_1^+(z)$ and $R_1^-(z)$ are entire.

(2) $R_2^+(z)$ and $R_2^-(z)$ are continuous on Im $z \geq 0$ and Im $z \leq 0$, respectively.

(3)
$$\frac{d^m}{dz^m}\left[R_1^+(z) + R_2^+(z)\right] = B^+(z), \quad \text{Im } z > 0$$

$$\frac{d^m}{dz^m}\left[R_1^-(z) + R_2^-(z)\right] = B^-(z), \quad \text{Im } z < 0$$

By Statement (3) and the assumptions of the theorem, $R^+(z)$ and $R^-(z)$ have the same mth derivatives as $R_1^+(z) + R_2^+(z)$ and $R_1^-(z) + R_2^-(z)$ do on their respective half-planes. Therefore, they are equivalent modulo P_m, i.e., there are polynomials $p^+, p^- \in P_m$ such that

$$R^+(z) = R_1^+(z) + R_2^+(z) + p^+(z), \quad \text{Im } z > 0$$

$$R^-(z) = R_1^-(z) + R_2^-(z) + p^-(z), \quad \text{Im } z < 0$$

By Statements (1) and (2), the right sides of these equations extend to the real line by continuity; therefore, so do the left sides, and

$$R^+(t+i0) - R^-(t-i0) = R_1^+(t) + R_2^+(t) - R_1^-(t) - R_2^-(t) + p^+(t) - p^-(t)$$

$$\cong R_x(t) \pmod{P_{2n}}$$

To reduce $R^+(t+i0) - R^-(t-i0)$ to the GACV forms for real $x(t)$, observe that for even $S_x(\nu)$ we have

$$-B^-(z) = (-1)^m B^+(-z) = \overline{B^+(\overline{z})}, \quad \text{Im } z < 0$$

(Only the first of these equations requires $S_x(\nu)$ to be even.) These three expressions are the mth derivatives of $-R^-(z)$, $R^+(-z)$, and $\overline{R^+(\overline{z})}$; accordingly, the latter functions are equivalent modulo P_m, and so are their limits $-R^-(t-i0)$, $R^+(-t+i0)$, and $\overline{R^+(t+i0)}$. ❒

B. Operations on Processes

The operations in question are integration, summation, filtering, and change of time scale, applied to processes with stationary nth increments. The following theorems give properties of the new processes from those of the old.

Theorem 7. *Let $x(t)$ have stationary nth increments, and $w(t) = \int x(t)\,dt$. Then $w(t)$ has stationary increments of order $n+1$. The drift rates, spectra, and GACVs are related by*

$$c_{n+1}(w) = c_n(x) \tag{28}$$

$$S_w(\nu) = \frac{S_x(\nu)}{(2\pi\nu)^2} \tag{29}$$

$$R_w(t) = -\left(\int dt\right)^2 R_x(t) \tag{30}$$

i.e., $R_w(t)$ is any function such that $-R_w''(t) = R_x(t)$.

Proof. We can assume $w(t) = J_0 x(t)$ [see Eq. (8)]. By Lemma 1, the operator $H_\tau = \Delta_\tau^{n+1} J_0$ is a TL filter of order n and $\lambda_n(H_\tau) = \tau^{n+1}$. Then $E\Delta_\tau^{n+1} w(t) = EH_\tau x(t) = c_n(x)\tau^{n+1}$ by Theorem 3, and Eq. (28) follows from Theorem 2.

By Theorem 3 and Lemma 1,

$$\begin{aligned} E\Delta_{\tau_1}^{n+1} w(s+t)\,\overline{\Delta_{\tau_2}^{n+1} w(s)} &= EH_{\tau_1} x(s+t)\,\overline{H_{\tau_2} x(s)} \\ &= (\tau_1\tau_2)^{n+1} |c_n(x)|^2 + \int_{-\infty}^{\infty} e^{i\omega t} H_{\tau_1}(\nu)\,\overline{H_{\tau_2}(\nu)} S_x(\nu)\,d\nu \\ &= (\tau_1\tau_2)^{n+1} |c_{n+1}(w)|^2 + \int_{-\infty}^{\infty} e^{i\omega t} \Delta_{\tau_1}^{n+1}(\nu)\,\overline{\Delta_{\tau_2}^{n+1}(\nu)} \frac{S_x(\nu)}{\omega^2} d\nu \end{aligned} \tag{31}$$

By Theorem 2, this proves Eq. (29) (and also shows that the w-drift is deterministic).

By Theorem 4, the right side of Eq. (31) equals $H_{\tau_1} H_{\tau_2}^* R_x(t)$. But by Statement (2) of Lemma 1,

$$H_{\tau_1} H_{\tau_2}^* R_x(t) = -\Delta_{\tau_1}^{n+1} J_0 \Delta_{-\tau_2}^{n+1} J_0 R_x(t) = -\Delta_{\tau_1}^{n+1} \Delta_{-\tau_2}^{n+1} J_0^2 R_x(t)$$

This shows that $-J_0^2 R_x(t)$ is a GACV for $w(t)$. ❒

Another version of Eq. (30) is given by

$$R_w(t) = -\int_0^t (t-u) R_x(u)\,du \tag{32}$$

For some applications, we need properties of discrete-time summations of $x(t)$. Given a fixed sample period τ_0, let

$$\left.\begin{aligned} w_n &= \sum_{j=1}^{n} x(j\tau_0), \quad n \geq 1 \\ w_0 &= 0 \\ w_n &= -\sum_{j=n+1}^{0} x(j\tau_0), \quad n \leq -1 \end{aligned}\right\} \tag{33}$$

where $x(t)$ has stationary pth increments. Then $\Delta_1 w_n = x(n\tau_0)$. This discrete-time process has stationary increments of order $p+1$. Without going through the theory here, we assert simply that w_n has a discrete-time GACV sequence

$$R_{w,n} = -\frac{|n|}{2} R_x(0) - \sum_{j=1}^{|n|-1} (|n| - j) R_x(\tau_0 j \text{ sgn } n) \tag{34}$$

which is τ_0^{-2} times a trapezoidal approximation to Eq. (32) with $t = n\tau_0$. Just as we had $-R_w''(t) = R_x(t)$ for $w(t) = \int x(t)\,dt$, here we have $-\delta_1^2 R_{w,n} = R_x(n\tau_0)$, where δ_1^2 is the second central difference operator with step 1. This discrete-time GACV behaves like a continuous-time GACV, allowing us to calculate covariances of the outputs of FIR filters of order $\geq p+1$ acting on w_n; see the discussion of modified Allan variance (Section V.A.4).

Theorem 8. *Let $x(t)$ have stationary nth increments, and let H be a TL filter of order $m \leq n$. The process $y(t) = Hx(t)$ has stationary $(n-m)$th increments, and*

$$c_{n-m}(y) = \lambda_m(H)\, c_n(x) \tag{35}$$

$$S_y(\nu) = |H(\nu)|^2 S_x(\nu) \tag{36}$$

$$R_y(t) = HH^* R_x(t) \tag{37}$$

Proof. The proof is like that of Theorem 7, but simpler. Let $r = n - m$; then $\Delta_\tau^r H$ is a TL filter of order n. We show the equations without further comment.

$$E\Delta_\tau^r y(t) = E\Delta_\tau^r Hx(t) = \tau^r \lambda_m(H)\, c_n(x)$$

$$\begin{aligned} E\Delta_{\tau_1}^r y(s+t)\,\overline{\Delta_{\tau_2}^r y(s)} &= E\Delta_{\tau_1}^r Hx(s+t)\,\overline{\Delta_{\tau_2}^r Hx(t)} \\ &= (\tau_1\tau_2)^r |\lambda_m(H)\, c_n(x)|^2 + \int_{-\infty}^{\infty} e^{i\omega t} \Delta_{\tau_1}^r(\nu)\,\overline{\Delta_{\tau_2}^r(\nu)}\,|H(\nu)|^2 S_x(\nu)\,d\nu \\ &= \Delta_{\tau_1}^r H \left(\Delta_{\tau_2}^r H\right)^* R_x(t) = \Delta_{\tau_1}^r \Delta_{-\tau_2}^r HH^* R_x(t) \end{aligned}$$

❐

Theorem 9. *Let $x_\rho(t) = x(\rho t)$, where $x(t)$ has stationary nth increments, and ρ is a nonzero real number. Then $c_n(x_\rho) = \rho^n c_n(x)$, $S_{x_\rho}(\nu) = |\rho|^{-1} S_x(\nu/\rho)$, and $R_{x_\rho}(t) = R_x(\rho t)$.*

The proof is left to the reader. We also remark that the spectrum of the sum $x_1(t) + x_2(t)$ of two processes with stationary nth increments is the sum of the spectra, provided that the nth increments of $x_1(t)$ are uncorrelated with those of $x_2(t)$; it is likewise for the GACV of the sum, modulo P_{2n}.

IV. GACVs for Power-Law Spectra

A. Pure Power Laws

For the function $|2\pi\nu|^\beta$ to be the spectrum of an ordinary process $x(t)$, it is necessary that $\beta < -1$; otherwise, there is infinite power at high frequencies. A more complete and elegant theory that included generalized processes would not need this restriction. (The case $\beta \geq -1$ is treated in the next section

with the help of high-frequency roll-offs.) Consider, then, the two-sided spectrum $S_x(\nu) = |2\pi\nu|^\beta$, $\nu \neq 0$, where $\beta < -1$. According to Theorem 2, this is the spectrum of a process with stationary increments of order $\lfloor (1/2)(1-\beta) \rfloor$. The corresponding GACV is denoted by $R(t;\beta)$.

1. Even Powers. We treated the case $\beta = -2$, finding that $R(t;-2) = -(1/2)|t|$. Then $R(t;\beta)$ for $\beta = -4, -6, -8, \ldots$ can be derived from Theorem 7 by applying the operator $-J_0^2$ repeatedly to $-(1/2)|t|$. We find that

$$R(t;-2n) = \frac{(-1)^n}{2} \frac{|t|^{2n-1}}{(2n-1)!}, \quad n = 1, 2, \ldots \tag{38}$$

where n is also the order of the stationary increments.

2. Odd Powers. If we do $\beta = -3$, the others will follow from Theorem 7. Setting $m = 3$ in Theorem 6 gives

$$B^+(z) = \frac{1}{i2\pi} \int_0^\infty e^{i\omega z} d\omega = \frac{1}{2\pi z}$$

Integrating three times and dropping the irrelevant quadratic polynomial gives

$$R^+(z) = \frac{1}{4\pi} z^2 \ln z$$

$$R(t;-3) = 2 \text{ Re } R^+(t + i0) = \frac{1}{2\pi} t^2 \ln |t| \tag{39}$$

with the understanding that $R(0;-3) = 0$.

It is all right that the argument of the logarithm is not dimensionless: replacing $|t|$ by $|t/\tau|$, say, changes $R(t;-3)$ only by a quadratic, which is irrelevant to the GACV of a process with stationary second increments. In contrast to the case of stationary first increments, the author cannot interpret the shape of this GACV, which is negative for small t, positive for large t.

Applying $-J_0^2$ repeatedly to Eq. (39) and dropping irrelevant polynomials gives the GACV for odd powers, namely

$$R(t;-2n+1) = \frac{(-1)^n}{\pi} \frac{t^{2n-2}}{(2n-2)!} \ln |t|, \quad n = 2, 3, \ldots \tag{40}$$

where n is again the order of the stationary increments.

3. Non-Integral Powers. The case $-3 < \beta < -1$ corresponds to the fractional Brownian motions introduced by Mandelbrot and Van Ness [17], with $\beta = -1 - 2H$ in their notation. Let us compute their GACVs. Setting $m = 2$ in Theorem 6 and using a standard gamma-function integral formula [20, Eq. 6.1.1], we have

$$B^+(z) = \frac{-1}{2\pi}\int_0^\infty e^{i\omega z}\omega^{2+\beta}d\omega = \frac{-\Gamma(3+\beta)}{2\pi}(-iz)^{-3-\beta}$$

$$= \frac{\Gamma(3+\beta)}{2\pi}e^{i(\pi/2)(1+\beta)}z^{-3-\beta}, \quad \text{Im } z > 0$$

in which we define $a^\mu = |a|^\mu e^{i\mu \arg a}$ for $-\pi < \arg a < \pi$. For $\beta \neq -2$, two integrations give

$$R^+(z) = \frac{\Gamma(1+\beta)}{2\pi}e^{i(\pi/2)(1+\beta)}z^{-1-\beta}$$

If $t > 0$, then $(t+i0)^{-1-\beta} = |t|^{-1-\beta}$; if $t < 0$, then $(t+i0)^{-1-\beta} = \left(|t|e^{i\pi}\right)^{-1-\beta} = |t|^{-1-\beta}e^{i\pi(-1-\beta)}$. Thus, for all t,

$$R(t;\beta) = 2\ \text{Re}\ R^+(t+i0) = \frac{\Gamma(1+\beta)}{\pi}|t|^{-1-\beta}\cos\left(\frac{\pi}{2}(1+\beta)\right)$$

$$= \frac{-\Gamma(1+\beta)}{\pi}\sin\left(\frac{\pi}{2}\beta\right)|t|^{-1-\beta}$$

The formula $\Gamma(-\beta)\Gamma(1+\beta) = -\pi\csc(\pi\beta)$ transforms this equation to

$$R(t;\beta) = \frac{1}{2\cos\left(\frac{\pi}{2}\beta\right)}\frac{|t|^{-1-\beta}}{\Gamma(-\beta)} \tag{41}$$

which is valid even for $\beta = -2$. Applying $-J_0^2$ to this formula gives the same formula with β replaced by $\beta - 2$. It follows that Eq. (41) holds whenever $\beta < -1$ and is not an odd integer, that case being covered by Eq. (40).

Exercise 1. *Derive Eq. (40) from Eq. (41) by an appropriate limiting procedure.*

Exercise 2. *(Self-similarity of power-law processes.) Let $\beta < -1$ and $H = (1/2)(-1-\beta)$. If $x(t)$ has zero drift rate and spectrum proportional to $|\nu|^\beta$, then $\rho^{-H}x(\rho t)$ has the same spectrum and GACV as $x(t)$ does.*

B. Filtered Power Laws

If $-1 \leq \beta \leq 0$, then the spectrum $|2\pi\nu|^\beta$ has to be rolled off at high frequencies if it is to be the spectrum of an ordinary process. One choice of roll-off method is a sharp cutoff at some frequency ν_h; the corresponding GACV (if $\beta = -1$) or ACV (if $\beta > -1$) can be obtained by cutting off the integration in Eq. (23) at ν_h. Then $R_x(t)$ will come out in terms of non-elementary special functions. Unless given a reason to believe that a sharp cutoff is a close approximation to reality, the author prefers another roll-off method, moving averages, because it is easy to use. This method uses the filter $A_\epsilon = \epsilon^{-1}\Delta_\epsilon J_0$ (see Section II.B.2) as a lowpass filter with nominal roll-off frequency $\nu_h = (2\epsilon)^{-1}$. We want to have $x_\epsilon(t) = A_\epsilon x(t)$, where $x(t)$ is a generalized process with a pure $|2\pi\nu|^\beta$ spectrum for $0 < |\nu| < \infty$, the derivative of an ordinary process $w(t)$ with a pure $|2\pi\nu|^{\beta-2}$ spectrum. Then we would have $w(t) = J_0 x(t) + \text{const}$ and $A_\epsilon x(t) = \epsilon^{-1}\Delta_\epsilon J_0 x(t) = \epsilon^{-1}\Delta_\epsilon w(t)$. Thus, the rolled-off $|2\pi\nu|^\beta$ process is actually defined as the difference quotient

$$x_\epsilon(t) = \frac{1}{\epsilon}\Delta_\epsilon w(t) = \frac{1}{\epsilon}\left[w(t) - w(t-\epsilon)\right]$$

where $w(t)$ has stationary second increments, and $S_w(\nu) = |2\pi\nu|^{\beta-2}$. By Theorem 8,

$$S_{x_\epsilon}(\nu) = \frac{1}{\epsilon^2}\left|\Delta_\epsilon(\nu)\right|^2 S_w(\nu) = \frac{\sin^2(\pi\nu\epsilon)}{(\pi\nu\epsilon)^2}\left|2\pi\nu\right|^\beta \tag{42}$$

$$R_{x_\epsilon}(t) = \frac{1}{\epsilon^2}\Delta_\epsilon\Delta_{-\epsilon}R_w(t) = -\frac{1}{\epsilon^2}\delta_\epsilon^2 R(t;\beta-2)$$

$$= \frac{1}{\epsilon^2}\left[-R(t-\epsilon;\beta-2) + 2R(t;\beta-2) - R(t+\epsilon;\beta-2)\right] \tag{43}$$

where $R(t;\beta-2)$ is obtained from Eq. (40) or Eq. (41).

For $\beta < -1$, we could do the same thing in case we wished to calculate the effect of high-frequency roll-offs on functions of $|\nu|^\beta$ noise; usually, however, we just use pure $|\nu|^\beta$ noise without bandwidth restrictions.

The choice of unweighted moving averages as a high-frequency roll-off method is not entirely for mathematical convenience: the author designed and programmed the signal-processing routines for a real-time stability analyzer [19] that produces a sequence of adjacent unweighted ϵ-averages of amplitude and phase residuals of a sine-wave signal, with ϵ selectable by the user.

Although Eq. (43) can easily be used as is, it is interesting to examine it further for different cases of β. By expanding $R(u;\beta-2)$ in a Taylor series about $u = t$, we find approximations, valid for $\epsilon \ll |t|$, in which the second central difference in Eq. (43) is replaced by the second derivative.

Case 1: $\beta = 0$. The corresponding $x_\epsilon(t)$ is ϵ-averaged white noise. No approximations are needed. Substituting $R(t;-2) = -(1/2)|t|$ into Eq. (43), we obtain the familiar hat-shaped ACV

$$R_{x_\epsilon}(t) = \frac{1}{\epsilon}\left(1 - \frac{|t|}{\epsilon}\right), \quad |t| \le \epsilon$$

$$= 0, \quad |t| \ge \epsilon \tag{44}$$

Case 2: $-1 < \beta < 0$. The corresponding $x_\epsilon(t)$ is a stationary long-memory process, called fractional noise by Mandelbrot and Van Ness [17]. If the notation of Eq. (41) is extended to this range of β, then $R(t;\beta) = -R''(t;\beta-2)$ as before. The Taylor expansion then gives $R_{x_\epsilon}(t) = R(t;\beta)\left[1 + O\left(\epsilon^2/t^2\right)\right]$ for the ACV of this process. Explicitly, we have

$$\left.\begin{aligned} R_{x_\epsilon}(0) &= \frac{\epsilon^{-1-\beta}}{\cos\left(\frac{\pi}{2}\beta\right)\Gamma(2-\beta)} \\ R_{x_\epsilon}(t) &= \frac{|t|^{-1-\beta}}{2\cos\left(\frac{\pi}{2}\beta\right)\Gamma(-\beta)}\left[1 + O\left(\frac{\epsilon^2}{t^2}\right)\right] \end{aligned}\right\} \tag{45}$$

Case 3: $\beta = -1$. The corresponding $x_\epsilon(t)$, which has stationary first increments, is an example of flicker or "$1/f$" noise that has been rolled off at high frequencies. After calculating the GACV, we are allowed to add any constant to it; in this way, we obtain

$$\left.\begin{aligned} R_{x_\epsilon}(0) &= \frac{1}{\pi}\left(\frac{3}{2} - \ln\epsilon\right) \\ R_{x_\epsilon}(t) &= -\frac{1}{\pi}\ln|t| + O\left(\frac{\epsilon^2}{t^2}\right) \end{aligned}\right\} \tag{46}$$

This GACV can be compared to the GACV for flicker noise with a sharp cutoff at $\nu = \nu_h$ by equating ν_h to $(2\epsilon)^{-1}$. The sharp-cutoff asymptotic result is like Eq. (46) except that the $3/2$ in $R_{x_\epsilon}(0)$ is replaced by $\gamma + \ln\pi \approx 1.722$, where γ is Euler's constant, and the error term is $O(\epsilon/t)$ instead of $O\left(\epsilon^2/t^2\right)$—another reason for preferring the moving-average filter to the sharp cutoff.

These approximations are useful when the continuous-time process $x_\epsilon(t)$ is undersampled—that is, when we have samples $x_\epsilon(j\tau_0)$ whose Nyquist frequency $(2\tau_0)^{-1}$ is much less than the process bandwidth $\nu_h = (2\epsilon)^{-1}$, so that $\tau_0 \gg \epsilon$. In this case, when doing covariance computations related to the sampled data, we have to evaluate $R_{x_\epsilon}(t)$ only at integer multiples of τ_0; thus, $t = 0$ or $|t| \gg \epsilon$, as Eqs. (45) and (46) require. As Eq. (44) confirms, the τ_0 samples of ϵ-averaged white noise are uncorrelated if $\tau_0 \geq \epsilon$.

V. Applications to Clock Statistics

Suppose that the readings of two clocks, C_1 and C_2, at time t are represented by $t_1(t)$ and $t_2(t)$, where t is the time as given by a time-base clock (perhaps even C_1, in which case $t_1(t) = t$). Then $x(t) = t_2(t) - t_1(t)$ is called the time deviation. Actually, the time-base clock has only to be good enough to supply time tags for measurements of $x(t)$: we might be measuring $x(t)$ to the nearest nanosecond by means of a time interval counter, while recording t to the nearest second or even the nearest day, depending on the scale of the experiment. The derivative $y(t) = dx(t)/dt$ is called the normalized or fractional frequency deviation. In the following discussions, however, $y(t)$ is simply called "frequency" even though it is dimensionless.

A. Increment Variances

The nth increment variance of $x(t)$ is defined as the mean-square nth increment

$$D_n(\tau) = E\left[\Delta_\tau^n x(t)\right]^2 \tag{47}$$

provided that this quantity exists and is independent of t. If $x(t)$ has stationary nth increments, then $D_n(\tau)$ can be computed from $R_x(t)$ by

$$\begin{aligned} D_n(\tau) &= \Delta_\tau^n \Delta_{-\tau}^n R_x(0) = (-1)^n \delta_\tau^{2n} R_x(0) \\ &= \sum_{j=-n}^{n} \binom{2n}{n+j} (-1)^j R_x(j\tau) \end{aligned} \tag{48}$$

[see Eqs. (5) and (6)]. Several measures of clock stability are just scaled versions of $D_n(\tau)$, either for $x(t)$ or for related processes. The results given below can be shortened by assuming that $R_x(-\tau) = R_x(\tau)$, as we always can for real-valued $x(t)$.

1. Mean-Square Average Frequency. Define

$$y(t,\tau) = A_\tau y(t) = \frac{1}{\tau}\int_{t-\tau}^{t} y(u)\,du = \frac{1}{\tau}\Delta_\tau x(t) \tag{49}$$

If $x(t)$ has stationary first increments, in particular, if $S_x(\nu) \sim |\nu|^\beta$ with $\beta > -3$, then

$$Ey^2(t,\tau) = \frac{1}{\tau^2}D_1(\tau) = -\frac{1}{\tau^2}\delta_\tau^2 R_x(0)$$

$$= \frac{1}{\tau^2}\left[-R_x(-\tau) + 2R_x(0) - R_x(\tau)\right] \tag{50}$$

2. Allan Variance. If $x(t)$ has stationary second increments ($S_x(\nu) \sim |\nu|^\beta$ with $\beta > -5$), Allan variance is defined by

$$\sigma_{y\mathrm{A}}^2(\tau) = \frac{1}{2}E\left[y(t,\tau) - y(t-\tau,\tau)\right]^2 = \frac{1}{2}E\left[\Delta_\tau y(t,\tau)\right]^2 = \frac{1}{2\tau^2}E\left[\Delta_\tau^2 x(t)\right]^2$$

Then

$$\sigma_{y\mathrm{A}}^2(\tau) = \frac{1}{2\tau^2}D_2(\tau) = \frac{1}{2\tau^2}\delta_\tau^4 R_x(0)$$

$$= \frac{1}{2\tau^2}\left[R_x(-2\tau) - 4R_x(-\tau) + 6R_x(0) - 4R_x(\tau) + R_x(2\tau)\right] \tag{51}$$

3. Hadamard Variance. For $x(t)$ with stationary third increments ($S_x(\nu) \sim |\nu|^\beta$ with $\beta > -7$),

$$\sigma_{y\mathrm{H}}^2(\tau) = \frac{1}{6\tau^2}D_3(\tau) = -\frac{1}{6\tau^2}\delta_\tau^6 R_x(0)$$

$$= \frac{1}{6\tau^2}\left[-R_x(-3\tau) + 6R_x(-2\tau) - 15R_x(-\tau) + 20R_x(0) - 15R_x(\tau) + 6R_x(2\tau) - R_x(3\tau)\right] \tag{52}$$

4. Modified Allan Variance. Like Allan variance, the modified version $\sigma_{y\mathrm{M}}^2$ is applicable to $x(t)$ with stationary second increments. Although this stability measure can be applied only to sampled data $x(n\tau_0)$, we first treat a theoretical continuous-time analog $\sigma_{y\mathrm{Mc}}^2(\tau)$ studied by Bernier [18]. Consider the moving averages $x(t,\tau) = A_\tau x(t) = \tau^{-1}\Delta_\tau w(t)$, where $w(t) = J_0 x(t) = \int_0^t x(u)\,du$. If $x(t)$ has stationary second increments, then $w(t)$ has stationary third increments. By definition,

$$\sigma_{y\mathrm{Mc}}^2(\tau) = \frac{1}{2\tau^2}E\left[\Delta_\tau^2 x(t,\tau)\right]^2$$

Then

$$\sigma_{y\mathrm{Mc}}^2(\tau) = \frac{1}{2\tau^4} E\left[\Delta_\tau^3 w(t)\right]^2 = -\frac{1}{2\tau^4}\delta_\tau^6 R_w(0)$$

$$= \frac{1}{2\tau^4}\left[-R_w(-3\tau) + 6R_w(-2\tau) - 15R_w(-\tau) + 20R_w(0) - 15R_w(\tau) + 6R_w(2\tau) - R_w(3\tau)\right] \tag{53}$$

[6]. By Theorem 7, we can obtain $R_w(\tau)$ from two successive integrations of $R_x(t)$ or from Eq. (32). If $x(t)$ is a power-law process, then so is $w(t)$, and we can obtain its GACV directly from Section IV. More to the point, we might wish to ignore bandwidth limitations on $x(t)$, even when $\beta \geq -1$, by treating $w(t)$ as a *pure* power-law process with exponent $\beta - 2$, whose GACV is exactly $R(t; \beta - 2)$. When we do this, we find that $\sigma_{y\mathrm{Mc}}^2(\tau) \propto \tau^{-3-\beta}$ for $-5 < \beta < 1$ [18].

The actual modified Allan variance is denoted here by $\sigma_{y\mathrm{M}}^2(m, \tau_0)$, where τ_0 is the data sample period and m is a positive integer. To define it in terms of the sampled process $x(n\tau_0)$, we first define the discrete-time moving average:

$$x_n(m) = \frac{1}{m}\sum_{j=n-m+1}^{n} x(j\tau_0) = \frac{1}{m}\Delta_m w_n$$

where w_n is given by Eq. (33). By definition,

$$\sigma_{y\mathrm{M}}^2(m, \tau_0) = \frac{1}{2(m\tau_0)^2} E\left[\Delta_m^2 x_n(m)\right]^2$$

Then

$$\sigma_{y\mathrm{M}}^2(m, \tau_0) = \frac{1}{2m^4\tau_0^2} E\left[\Delta_m^3 w_n\right]^2 = -\frac{1}{2m^4\tau_0^2}\delta_m^6 R_{w,0}$$

$$= \frac{1}{2m^4\tau_0^2}\left(-R_{w,-3m} + 6R_{w,-2m} - 15R_{w,-m} + 20R_{w,0} - 15R_{w,m} + 6R_{w,2m} - R_{w,3m}\right) \tag{54}$$

where $R_{w,n}$ is given in terms of $R_x(j\tau_0)$ by Eq. (34). In turn, if $x(t)$ is modeled by ϵ-averaged ν^β noise, then $R_x(j\tau_0)$ is given by Eq. (43).

B. Frequency Transfer Variance

In this section and those to follow, we assume that $R_x(-t) = R_x(t)$.

Given four times t_1, t_2, t_3, t_4, let $\tau_A = t_2 - t_1$, $t_D = t_3 - t_2$, and $\tau_B = t_4 - t_3$. Assume only that $\tau_A > 0$ and $\tau_B > 0$. The frequency transfer variance of Boulanger and Douglas [22] is defined by

$$u_y^2(\tau_A, t_D, \tau_B) = E\left[\frac{x(t_4) - x(t_3)}{\tau_B} - \frac{x(t_2) - x(t_1)}{\tau_A}\right]^2 \tag{55}$$

which is the mean-square error of estimating the average frequency in an interval $B = [t_3, t_4]$ from the average frequency in an interval $A = [t_1, t_2]$. These intervals can be located arbitrarily with respect to each other. A special case is Allan variance $\sigma_{y\mathrm{A}}^2(\tau) = (1/2)u_y^2(\tau, 0, \tau)$.

Assuming that $x(t)$ has stationary second increments, we can regard the function u_y^2 as a complete structure function for $x(t)$ in that knowledge of Eq. (55) for all τ_A, t_D, and τ_B gives all the second-increment covariances and, hence, determines the GACV or spectrum. If the quadratic drift rate of $x(t)$ is zero, then u_y^2 is a variance.

To evaluate u_y^2 from the GACV of $x(t)$, we observe that the expression inside the brackets in Eq. (55) equals $Hx(0)$, where H is the filter with impulse response

$$h(t) = \frac{\delta(t+t_4) - \delta(t+t_3)}{\tau_B} - \frac{\delta(t+t_2) - \delta(t+t_1)}{\tau_A} = \sum_i h_i \delta(t+t_i) \tag{56}$$

Since this filter has order 2, it follows that $u_y^2 = HH^* R_x(0)$. The impulse response of $F = HH^*$ is given by

$$f(t) = \int_{-\infty}^{\infty} h(t+u)\, h(u)\, du = \sum_{i,j} h_i h_j \delta(t + t_i - t_j) \tag{57}$$

Then $FR_x(0) = \sum_{i,j} h_i h_j R_x(t_j - t_i)$. Working in a computer algebra system that knows how to handle Dirac δ-functions, such as Maple,[3] we can assign $h(t)$ by Eq. (56), evaluate the integral in Eq. (57), and evaluate $\int f(t) R_x(t)\, dt$. We arrive at the pleasantly symmetrical result of Boulanger and Douglas:

$$\begin{aligned} u_y^2(\tau_A, t_D, \tau_B) = {} & 2\left[\frac{I_x(\tau_A)}{\tau_A^2} + \frac{I_x(\tau_B)}{\tau_B^2}\right] \\ & + \frac{2}{\tau_A \tau_B}\left[-I_x(t_D) + I_x(t_D+\tau_A) + I_x(t_D+\tau_B) - I_x(t_D+\tau_A+\tau_B)\right] \end{aligned} \tag{58}$$

where $I_x(t) = R_x(0) - R_x(t)$.

For ν^β processes, $\beta = 0, -1, -2, -3, -4$, Boulanger and Douglas give a convenient means (A-B structure factors) for scaling the result, Eq. (58), to the corresponding Allan variance, $\sigma_{y\mathrm{A}}^2(\tau_A)$. In their formulas, one can see the GACVs that are derived here in Section IV. Their derivation of Eq. (58) uses the integral in Eq. (23) with $S_x(\nu)$ truncated below a frequency ν_1 to make the integral converge. After calculating u_y^2 as a linear combination of such integrals, they find that the result converges to a finite limit as $\nu_1 \to 0$; this happens because H has order 2. The advantage of the method given here is that the analytic problem, divergence of integrals over frequency, has already been taken care of by the GACV theory. To calculate the variance of the output of a discrete filter acting on $x(t)$, it remains only to substitute a specific form of $R_x(t)$ into sums $\sum_k a_k R_x(u_k)$ in which a_k and u_k depend only on the filter, not on the noise type.

1. Mean-Square Time Interval Error. Suppose that we observe the time difference $x(t)$ of a pair of clocks at times t_1 and $t_2 = t_1 + \tau_A$, and extrapolate these two observations linearly to a later time, $t_2 + \tau_B$. What is the mean-square error of the extrapolated time difference? As Boulanger and Douglas [22] point out, this is just a special case of the frequency transfer variance, u_y^2, discussed above. In fact, the mean-square extrapolation error is given by

[3] Trademark of Waterloo Maple Inc.

$$\text{MSTIE}\ (\tau_A, \tau_B) = E\left\{x\left(t_2+\tau_B\right) - x\left(t_2\right) - \frac{\tau_B}{\tau_A}\left[x\left(t_2\right) - x\left(t_1\right)\right]\right\}^2$$

$$= \tau_B^2 u_y^2\left(\tau_A, 0, \tau_B\right) \tag{59}$$

[see Eq. (58)]. An interesting special case is flicker FM noise, $S_x(\nu) = a_{-3}|2\pi\nu|^{-3}$, which dominates many quartz oscillators over a considerable range of frequencies. Using Eq. (39) for $R_x(t)$ in Eq. (58) and setting $r = \tau_B/\tau_A$, we find

$$\text{MSTIE}\ (\tau_A, \tau_B) = a_{-3}\frac{\tau_B^2}{\pi}\left[\left(1+\frac{1}{r}\right)\ln r + \left(1+\frac{1}{r}\right)^2 r \ln\left(1+\frac{1}{r}\right)\right]$$

$$= a_{-3}\frac{\tau_B^2}{\pi}\left(1+\ln r\right)\left(1+O\left(\frac{1}{r}\right)\right) \quad \text{as } r \to \infty$$

The factor of order $\ln(\tau_B/\tau_A)$ is a subtle long-memory property of this model of flicker FM, not shared by certain other models of flicker FM that do not have stationary second increments [23,24].

C. Drift Estimators

Assume that $x(t)$ has stationary second increments. Another special case of the frequency transfer variance u_y^2 is the variance of "4-point x" estimators of frequency drift rate $c = c_2(x)$. Observing $x(t)$ for $0 \le t \le T$, we define

$$\hat{c}(\tau) = \frac{y(T,\tau) - y(\tau,\tau)}{T-\tau} = \frac{x(T) - x(T-\tau) - x(\tau) + x(0)}{\tau(T-\tau)} = \frac{1}{\tau(T-\tau)}\Delta_\tau \Delta_{T-\tau} x(T) \tag{60}$$

[see Eq. (49)], where $0 < \tau \le T/2$. In other words, $\hat{c}(\tau)$ is the divided difference of the first and last available τ-average frequencies. We have $E\hat{c}(\tau) = c$, and because $\hat{c}(\tau) - c$ is invariant to c, we can assume that $c = 0$. Then

$$\text{var}\ \hat{c}(\tau) = E\hat{c}^2(\tau) = \frac{1}{(T-\tau)^2}u_y^2(\tau, T-2\tau, \tau)$$

When $\tau = T/2$, we have the "3-point x" drift estimator $\hat{c}(T/2) = 4T^{-2}\Delta_{T/2}^2 x(T)$, whose variance is $8T^{-2}\sigma_{y\text{A}}^2(T/2) = 4T^{-2}u_y^2(T/2, 0, T/2)$ [25].

An example of a different kind of unbiased frequency drift estimator is the "least-squares y" estimator defined by

$$\hat{c}_{\text{LS}y} = \frac{6}{T^3}\int_0^T (2t - T)\, y(t)\, dt \tag{61}$$

This is the continuous-time analog of the slope of the least-squares linear fit to sampled data $y_1, \cdots, y_N$. Since $y(t)$ might be a generalized process, we integrate Eq. (61) by parts to obtain

$$\hat{c}_{\mathrm{LS}y} = \frac{6}{T^2}\left[x(0) + x(T) - \frac{2}{T}\int_0^T x(t)\,dt\right]$$

$$= Hx(T)$$

where

$$H = \frac{6}{T^2}\left(B_0 + B_T - 2A_T\right)$$

a real TL filter of order 2 whose impulse response is supported on $[0, T]$ and is even about $T/2$. (See Section II.B for the definitions of the shift B_τ and moving average A_τ.) Using Eq. (13), we find that $H^* = B_{-T}H$ and $HH^* = B_{-T}H^2$. A straightforward calculation gives

$$HH^*\delta(t) = \frac{36}{T^4}\left[\delta(t+T) + 2\delta(t) + \delta(t-T) - \frac{4|t|}{T^2}\mathbf{1}(|t| < T)\right]$$

$$\operatorname{var}\hat{c}_{\mathrm{LS}y} = HH^*R_x(0) = \frac{72}{T^4}\left[R_x(0) + R_x(T) - \frac{4}{T^2}\int_0^T R_x(t)\,tdt\right]$$

where it is assumed that $R_x(-t) = R_x(t)$.

Formulas for the variance of several drift estimators in the presence of power-law noises, computed by this method, can be found in [7].

D. Allan Variance With Drift Removal

We conclude this set of applications with a moderately complex calculation, similar to others that the author has carried out [4,7]. Some motivation and numerical results are given. The problem is to find out how drift removal affects the mean and variance of an estimator of Allan variance. Let $N+1$ time-deviation data $x(n\tau_0)$, $0 \le n \le N$, be given. Write $T = N\tau_0$, $\tau = m\tau_0$, where $1 \le m \le N/2$. Let $z(t) = \Delta_\tau^2 x(t)$. The available second τ-increments of $x(t)$ are $z_n = z(n\tau_0)$, $2m \le n \le N$. The fully overlapped estimator of $\sigma_{y\mathrm{A}}^2(\tau)$ is given by

$$V = \frac{1}{2\tau^2 M}\sum_{n=2m}^{N} z_n^2 \tag{62}$$

where $M = N - 2m + 1$. This estimator is distinguished from the τ-overlap estimator, which uses the average of the subsequence $z_{2m}^2, z_{3m}^2, \ldots$. Both estimators are unbiased, but the fully overlapped estimator usually has the lesser variance [26,3,27].

If $x(t)$ includes a long-term drift component $(1/2)ct^2$, then, for large τ, the drift contribution $(1/2)c^2\tau^2$ to the Allan variance masks the part of $\sigma_{y\mathrm{A}}^2(\tau)$ that helps to characterize the random fluctuations. If we knew c, we could remove it; usually we have to make do with an estimate $\hat{c}$ derived from the current data. The estimator V applied to $x(t) - (1/2)\hat{c}t^2$ gives a different result than V applied to $x(t) - (1/2)ct^2$ does; here we show how to calculate the mean and variance for both situations.

Define a "drift-estimator τ" by $\tau_c = m_c\tau_0$, where $1 \le m_c \le N/2$, and let $\hat{c}$ be the 4-point (or possibly 3-point) drift estimator,

$$\hat{c} = \hat{c}(\tau_c) = H_c x(T) = \frac{1}{\tau_c (T - \tau_c)} \Delta_{\tau_c} \Delta_{T-\tau_c} x(T)$$

[see Eq. (60)]. The second τ-increments of $x(t) - (1/2)\hat{c}t^2$ being $z_n - \hat{c}\tau^2$, we define the estimated-drift-removed estimate of Allan variance to be

$$V_0 = \frac{1}{2\tau^2 M} \sum_{n=2m}^{N} \left(z_n - \hat{c}\tau^2\right)^2$$

Expanding the square and letting $\hat{z} = (1/M)\sum_{n=2m}^{N} z_n$, we have

$$V_0 = V - \hat{c}\hat{z} + \frac{\tau^2}{2}\hat{c}^2 \tag{63}$$

Because V_0 is invariant to the true drift rate c, we can assume that $c = 0$; in this case, V becomes an unbiased estimator of Allan variance with the true drift removed. Now our job is to compare the mean and variance of V and V_0 under the condition that $x(t)$ is a process with stationary zero-mean second increments. To compute variances of V and V_0, we also assume that the second increments are a Gaussian process, so that Isserlis's theorem (Section III of the Appendix) applies.

1. Means. Under the given assumptions, $z(t)$ is a stationary zero-mean process with ACV

$$R_z(t) = \delta_\tau^4 R_x(t)$$

and we have

$$E\hat{c}z_n = EH_c x(T)\,\Delta_\tau^2 x(n\tau_0) = H_c \Delta_{-\tau}^2 R_x((N-n)\tau_0)$$

$$= f((N-n)\tau_0)$$

where $f(t) = H_c\Delta_{-\tau}^2 R_x(t)$, the result of another fourth-order filter acting on $R_x(t)$. Then

$$EV = \frac{1}{2\tau^2} Ez^2(t) = \frac{1}{2\tau^2} R_z(0)$$

$$EV_0 = EV - E\hat{c}\hat{z} + \frac{\tau^2}{2} E\hat{c}^2$$

$$E\hat{c}\hat{z} = \frac{1}{M}\sum_{n=2m}^{N} E\hat{c}z_n = \frac{1}{M}\sum_{k=0}^{M-1} f(k\tau_0)$$

$$E\hat{c}^2 = E\left[H_c x(T)\right]^2 = H_c H_c^* R_x(0)$$

These formulas give EV and EV_0 as finite linear combinations of $R_x(n\tau_0)$.

2. Variances.

(1) From Eq. (62) and Isserlis's theorem, we have

$$\operatorname{var} V = \frac{1}{4\tau^4 M^2} \sum_{i,j=2m}^{N-1} \operatorname{cov}\left(z_i^2, z_j^2\right) = \frac{1}{2\tau^4 M^2} \sum_{i,j} \left(E z_i z_j\right)^2$$

$$= \frac{1}{2\tau^4 M^2} \sum_{i,j} R_z^2\left((i-j)\tau_0\right)$$

$$= \frac{1}{2\tau^4 M^2}\left[M R_z^2(0) + 2\sum_{k=1}^{M-1}(M-k) R_z^2(k\tau_0)\right]$$

where it is assumed that $R_z(-t) = R_z(t)$. This part of the calculation was carried out by Yoshimura [27].

From Eq. (63), the variance of V_0 is a sum of six terms:

$$\operatorname{var} V_0 = \operatorname{var} V + \operatorname{var}(\hat{c}\hat{z}) + \frac{\tau^4}{4}\operatorname{var}\left(\hat{c}^2\right) - 2\operatorname{cov}(V, \hat{c}\hat{z}) + \tau^2 \operatorname{cov}\left(V, \hat{c}^2\right) - \tau^2 \operatorname{cov}\left(\hat{c}\hat{z}, \hat{c}^2\right)$$

Let us take up the last five in turn, using Isserlis's theorem to evaluate the covariances.

(2)

$$\operatorname{var}(\hat{c}\hat{z}) = \operatorname{cov}(\hat{c}\hat{z}, \hat{c}\hat{z}) = \left(E\hat{c}^2\right)\left(E\hat{z}^2\right) + \left(E\hat{c}\hat{z}\right)^2$$

$$E\hat{z}^2 = \frac{1}{M^2}\sum_{i,j} E z_i z_j$$

$$= \frac{1}{M^2}\left[M R_z(0) + 2\sum_{k=1}^{M-1}(M-k) R_z(k\tau_0)\right]$$

(3)

$$\operatorname{var}\left(\hat{c}^2\right) = \operatorname{cov}(\hat{c}\hat{c}, \hat{c}\hat{c}) = 2\left(E\hat{c}^2\right)^2$$

(4)

$$\mathrm{cov}\left(V,\hat{c}\hat{z}\right)=\frac{1}{2\tau^2M^2}\sum_{i,j=2m}^{N}\mathrm{cov}\left(z_i^2,\hat{c}z_j\right)=\frac{1}{\tau^2M^2}\sum_{i,j=2m}^{N}\left(E\hat{c}z_i\right)\left(Ez_iz_j\right)$$

$$=\frac{1}{\tau^2M^2}\sum_{i,j=2m}^{N}f\left(\left(N-i\right)\tau_0\right)R_z\left(\left(i-j\right)\tau_0\right)$$

$$=\frac{1}{\tau^2M^2}\sum_{k=0}^{M-1}f\left(k\tau_0\right)\sum_{i=-k}^{M-1-k}R_z\left(i\tau_0\right) \tag{64}$$

For each new m, we can precompute the array $s_n=\sum_{i=-M}^{n}R_z\left(i\tau_0\right)$; then the inner sum in Eq. (64) is $s_{M-1-k}-s_{-1-k}$.

(5)

$$\mathrm{cov}\left(V,\hat{c}^2\right)=\frac{1}{2\tau^2M}\sum_{n=2m}^{N}\mathrm{cov}\left(z_n^2,\hat{c}^2\right)=\frac{1}{\tau^2M}\sum_{n=2m}^{N}\left(E\hat{c}z_n\right)^2$$

$$=\frac{1}{\tau^2M}\sum_{k=0}^{M-1}f^2\left(k\tau_0\right)$$

(6)

$$\mathrm{cov}\left(\hat{c}\hat{z},\hat{c}^2\right)=2\left(E\hat{c}^2\right)\left(E\hat{c}\hat{z}\right)$$

3. Numerical Results. The author programmed these calculations in Maple and carried them out numerically for $\tau_0=1$, $N=100$, $\tau_c=15$, and $m=\tau/\tau_0=1,2,4,8,16,32,50$, using the GACVs for power-law spectra $\left|2\pi\nu\right|^\beta$, $\beta=0,-1,-2,-3$, and -4 (white PM, flicker PM, white FM, flicker FM, and random-walk FM), as given in Section IV. For $\beta=0$ and -1, the author used Eq. (43) with $\epsilon=\tau_0$; this choice gives discrete-time white noise for the white-PM case and realizes flicker-PM noise as the first τ_0-increment of flicker-FM noise. Figure 1 shows the ratios EV_0/EV and $\sqrt{\mathrm{var}V_0/\mathrm{var}V}$ as functions of m and β. Also shown are results of simulation with 10,000 trials, carried out with Matlab;[4] the theoretical and simulated values confirm each other's correctness.

For the "redder" noises, $\beta=-4,-3$, drift removal causes a negative bias in the Allan variance estimate, as expected, and a decrease of its standard deviation. For white FM, $\beta=-2$, drift removal has only a small effect. For the PM noises, $\beta=-1,0$, drift removal causes a positive bias and an increase of the standard deviation for the largest values of m. The author was surprised by this result. One might say that it does not matter since clock data dominated by the PM noises rarely show drift; nevertheless, the act of drift removal produces these effects whether or not any drift is present.

[4] Copyright of The MathWorks, Inc.

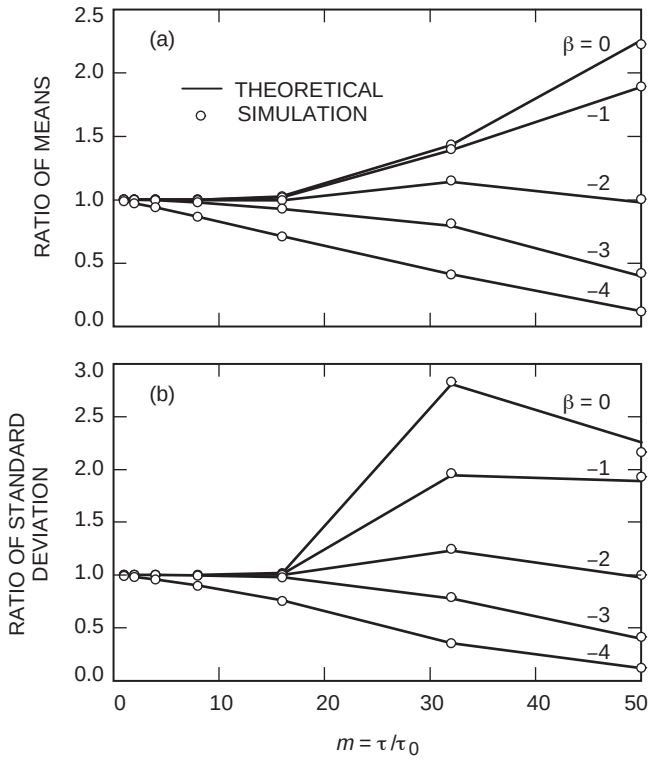

Fig. 1. The effect of drift removal on the fully overlapped Allan variance estimator. Shown are ratios (drift removed to drift not removed) of the (a) mean and (b) standard deviation of the estimator for 101 time samples and noise spectrum $S_x(\nu) = |2\pi\nu|^\beta$ (white PM to random walk FM).

VI. Conclusion

This article does not pretend to disclose new results in clock statistics; in fact, many of the results given here were previously derived for power-law processes by taking limits of frequency-domain integrals involving low-frequency cutoffs, or by expressing the desired quantities as functions of known quantities such as Allan variance with dead time. These ad hoc methods can be replaced by manipulations of the generalized autocovariance, a function of one time variable that extends the autocovariance (autocorrelation) function in a natural way from stationary noise models to noise models whose first or higher increments are stationary. Once the GACV for a particular noise model is obtained, results for a variety of statistics problems can be calculated from it by turning a crank. The GACV method, being based on sound mathematics, can be depended upon to give correct results for these noise models.

We close with a suggestion for future work. The extension from ACV to GACV was carried out only by generalizing the Fourier transform relationship between ACV and spectrum. On the other hand, since the ACV of a stationary process $x(t)$ is defined in the time domain as a mean lagged product, $Ex(t+\tau)\overline{x(t)}$, it would make sense to seek a time-domain expression of the GACV for an $x(t)$ with stationary nth increments. Having such, one might be able to estimate the GACV directly from a time series without going through an estimate of the spectrum. Perhaps the practice of computing mean lagged products to estimate an "ACV" for data that do not look stationary will turn out to have meaning after all.

References

[1] A. M. Yaglom, "Correlation Theory of Processes With Random Stationary nth Increments," *American Mathematical Society Translations*, ser. 2, vol. 8, pp. 87–141, 1958.

[2] P. Lesage and C. Audoin, "Characterization and Measurement of Time and Frequency Stability," *Radio Science*, vol. 14, pp. 521–539, 1979.

[3] C. A. Greenhall, "A Structure Function Representation Theorem With Applications to Frequency Stability Estimation," *IEEE Trans. Instrum Meas.*, vol. IM-32, pp. 364–370, 1983.

[4] C. A. Greenhall, "The Fundamental Structure Function of Oscillator Noise Models," *Proc. 14th Ann. PTTI Applications and Planning Meeting*, pp. 281–294, 1982.

[5] C. A. Greenhall, "Recipes for Degrees of Freedom of Frequency Stability Estimators," *IEEE Trans. Instrum. Meas.*, vol. 40, pp. 994–999, 1991.

[6] C. A. Greenhall, "The Third-Difference Approach to Modified Allan Variance," *IEEE Trans. Instrum. Meas.*, vol. 46, pp. 696–703, 1997.

[7] C. A. Greenhall, "A Frequency Drift Estimator and Its Removal From Modified Allan Variance," *Proc. 1997 IEEE International Frequency Control Symposium*, pp. 428–432, 1997.

[8] D. A. Howe and C. A. Greenhall, "Total Variance: A Progress Report on a New Frequency Stability Characterization," *Proc. 29th Ann. PTTI Systems and Applications Meeting*, pp. 39–48, 1997.

[9] H. Bremermann, *Distributions, Complex Variables, and Fourier Transforms*, Reading, Massachusetts: Addison-Wesley, 1965.

[10] M. J. Lighthill, *Introduction to Fourier Analysis and Generalised Functions*, Cambridge: Cambridge University Press, 1958.

[11] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, New York: Academic Press, 1980.

[12] I. N. Sneddon, *Fourier Transforms*, New York: McGraw-Hill, 1951.

[13] W. C. Lindsey and C. M. Chie, "Theory of Oscillator Instability Based Upon Structure Functions," *Proc. IEEE*, vol. 64, pp. 1652–1666, 1976.

[14] C. A. Greenhall, "Does Allan Variance Determine the Spectrum?," *Proc. 1997 IEEE International Frequency Control Symposium*, pp. 358–365, 1997.

[15] H. Cramér and M. R. Ledbetter, *Stationary and Related Stochastic Processes*, New York: Wiley, 1967.

[16] A. M. Yaglom, *An Introduction to the Theory of Stationary Random Functions*, Englewood Cliffs, New Jersey: Prentice-Hall, 1962.

[17] B. B. Mandelbrot and J. W. Van Ness, "Fractional Brownian Motions, Fractional Noises and Applications," *SIAM Review*, vol. 10, pp. 422–437, 1968.

[18] L. G. Bernier, "Theoretical Analysis of the Modified Allan Variance," *Proc. 41st Ann. Symposium on Frequency Control*, pp. 116–121, 1987.

[19] C. A. Greenhall, "Digital Signal Processing in the Radio Science Stability Analyzer," *The Telecommunications and Data Acquisition Progress Report 42-121,*

January–March 1995, Jet Propulsion Laboratory, Pasadena, California, pp. 271–287, May 15, 1995.
http://tmo.jpl.nasa.gov/tmo/progress_report/42-121/121G.pdf

[20] M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions*, New York: Dover, 1965.

[21] D. B. Percival and A. T. Walden, *Spectral Analysis for Physical Applications*, Cambridge: Cambridge University Press, 1993.

[22] J.-S. Boulanger and R. J. Douglas, "Traceability of Frequency Control Systems," *Proc. 1997 International Frequency Control Symposium*, pp. 366–372, 1997.

[23] J. A. Barnes and D. W. Allan, "A Statistical Model of Flicker Noise," *Proc. IEEE*, pp. 176–178, 1966.

[24] C. A. Greenhall, "Initializing a Flicker-Noise Generator," *IEEE Trans. Instrum. Meas.*, vol. IM-35, pp. 222–224, 1986.

[25] M. Weiss and C. Hackman, "Confidence on the Three-Point Estimator of Frequency Drift," *Proc. 24th Ann. Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, pp.451–455, 1992.

[26] D. A. Howe, D. W. Allan, and J. A. Barnes, "Properties of Signal Sources and Measurement Methods," *Proc. 35th Annual Symposium on Frequency Control*, pp. 1–47, 1981.

[27] K. Yoshimura, "Degrees of Freedom of the Estimate of the Two-Sample Variance in the Continuous-Sampling Method," *IEEE Trans. Instrum. Meas.*, vol. 38, pp. 1044–1049, 1989.

Acknowledgments

The author is grateful to Lute Maleki for encouraging and supporting this work, and to John Dick for serving as referee.

Appendix
Miscellaneous Arguments

I. Functions With Vanishing *n*th Increments

We want to prove that a continuous function whose nth increments are all zero must be a polynomial of degree $<n$. Here we prove the result in an unnecessarily general setting, both to purify the argument and to establish the result for vector-valued functions such as a mean-continuous random process, viewed as a mapping $t \to x(t)$ from the real line to a space of random variables with finite second moment. A real topological vector space is a vector space V over the reals R such that addition and scalar multiplication are continuous functions from $V \times V$ and $R \times V$, respectively, to V. In particular, a V-valued polynomial $p(t) = \sum_{i=0}^{m} t^i a_i$, where $a_i \in V$, is a continuous function from R to V.

Lemma A-1. *Let n be a positive integer and $f(t)$ a continuous function from the real line to a real topological vector space V. If $\Delta_\tau^n f(t) = 0$ for all rational τ* and *t, then $f(t)$ is a polynomial of degree $<n$ with coefficients in V.*

Proof. Let k be a positive integer. By means of the Lagrange interpolation formula, we can construct a polynomial $p_k(t)$ of degree $<n$, with coefficients in V, such that $p_k(t) = f(t)$ for $t = 1/k, 2/k, \cdots, n/k$. Then $r_j = f(j/k) - p_k(j/k)$ satisfies the difference equation $\Delta^n r_j = 0$ for all integers j, with the initial conditions $r_1 = \cdots = r_n = 0$. Solving this difference equation in both directions, we find that $r_j = 0$ for all integers j. This says that $f(t) = p_k(t)$ for all t of form j/k, in particular, for all integral t. Therefore, $p_k = p_1$ on the integers. We assert that $p_k = p_1$ everywhere; for if that is not so, then $q(t) = p_k(t) - p_1(t)$ is a nonzero polynomial $\sum_{i=0}^{m} t^i a_i$, where $m < n$ and $a_m \neq 0$. Since $q = 0$ on the integers, we have $\Delta_1^m q(0) = m! a_m = 0$, a contradiction.

We now see that f and p_1 agree on the rationals. Since both functions are continuous, they agree everywhere. ❒

II. Details of the Proof of Theorem 6

The purpose of this appendix is to prove statements (1) through (3) of the proof.

(1) Fix z and a sequence $z_j \to z$ $(z_j \neq z)$, and let $M \geq |z|, |z_j|$. Set

$$\phi_j(\nu) = \frac{e_m(i\omega z_j) - e_m(i\omega z)}{z_j - z} = \frac{i\omega}{z_j - z} \int_z^{z_j} e_{m-1}(i\omega\zeta)\, d\zeta$$

Then $\phi_j(\nu) \to i\omega e_{m-1}(i\omega z)$. By Eq. (3), the continuous function $\phi_j(\nu)$ and its limit are majorized in absolute value by $|\omega|^m M^{m-1} e^{|\omega M|}/(m-1)!$, which by the assumptions of the theorem is integrable with respect to $S_x(\nu)$ for $|\nu| \leq 1$. By Lebesgue's theorem on dominated convergence,

$$\frac{R_1^+(z_j) - R_1^+(z)}{z_j - z} = \int_0^1 \phi_j(\nu)\, S_x(\nu)\, d\nu \to \int_0^1 i\omega e_{m-1}(i\omega z)\, S_x(\nu)\, d\nu$$

and similarly for R_1^-. Thus, both are differentiable.

(2) Let $\operatorname{Im} z, \operatorname{Im} z_j \geq 0$, $z_j \to z$. Then $\left|e^{i\omega z}\right|, \left|e^{i\omega z_j}\right| \leq 1$ for $\nu \geq 0$, and $\int_1^\infty S_x(\nu)\, d\nu < \infty$. By dominated convergence, $R_2^+(z_j) \to R_2^+(z)$, and similarly for R_2^-.

(3) The exchange of differentiation with respect to z and the integration with respect to ν carried out in the proof of Item (1) can be performed m times, giving

$$\frac{d^m}{dz^m} R_1^\pm(z) = \int_0^{\pm 1} e^{i\omega z} (i\omega)^m S_x(\nu)\, d\nu$$

To deal with $R_2^\pm$, let $\operatorname{Im} z > 0$, $\operatorname{Im} z_j > 0$, $z_j \to z$, and $z_j \neq z$. Then there is a $b > 0$ such that $\operatorname{Im} z$ and $\operatorname{Im} z_j \geq b$. Set

$$\psi_j(\nu) = \frac{e^{i\omega z_j} - e^{i\omega z}}{z_j - z} = \frac{i\omega}{z_j - z} \int_z^{z_j} e^{i\omega\zeta} d\zeta$$

If $\nu > 0$, then $\psi_j(\nu)$ and its limit, $i\omega e^{i\omega z}$, are majorized in absolute value by $\omega e^{-b\omega}$, which is integrable with respect to $S_x(\nu)$ for $\nu \geq 1$. Thus, $dR_2^+(z)/dz = \int_1^\infty e^{i\omega z} i\omega S_x(\nu)\, d\nu$. Since $\omega^k e^{-b\omega}$ is S_x-integrable on $\nu \geq 1$ for any k, these derivatives can be repeated indefinitely under the integral sign, giving

$$\frac{d^m}{dz^m} R_2^+(z) = \int_1^\infty e^{i\omega z} (i\omega)^m S_x(\nu)\, d\nu, \quad \operatorname{Im} z > 0$$

and similarly for R_2^-.

III. Isserlis's Theorem

This is a general theorem about fourth moments of jointly Gaussian random variables with zero means. A set of N complex-valued random variables is said to be jointly Gaussian if the $2N$ real and imaginary parts have a joint Gaussian distribution.

Theorem A-1. *Let X_1, X_2, X_3, and X_4 be complex-valued jointly Gaussian random variables with expectations equal to zero. Then*

$$\operatorname{cov}(X_1 X_2, X_3 X_4) = \left(EX_1\overline{X_3}\right)\left(EX_2\overline{X_4}\right) + \left(EX_1\overline{X_4}\right)\left(EX_2\overline{X_3}\right) \tag{A-1}$$

See [21] for a reference. Here is a sketch of a proof. We can replace X_3 and X_4 by their conjugates; then Eq. (A-1) is equivalent to

$$EX_1X_2X_3X_4 = (EX_1X_2)(EX_3X_4) + (EX_1X_3)(EX_2X_4) + (EX_1X_4)(EX_2X_3) \tag{A-2}$$

To prove Eq. (A-2), we use the theorem that N real jointly Gaussian random variables can be expressed as linear combinations of at most N independent standard (mean 0, variance 1) Gaussians U_i. Applying this theorem to the 8 real and imaginary parts of the X_k, we see that the X_k are linear combinations, with complex coefficients, of a set of independent standard Gaussians U_i. Since both sides of Eq. (A-2) are linear in each X_k separately, it follows that Eq. (A-2) reduces to evaluation of $EU_iU_jU_kU_l$, of which the only nonzero cases are $EU_i^2U_j^2 = EU_i^2EU_j^2$ $(i \neq j)$ and $EU_i^4 = 3EU_i^2$.