

AN EXAMPLE OF AN INFINITE DIMENSIONAL FILTERING PROBLEM: FILTERING FOR GYROSCOPIC NOISE

by

L. Horowitz
Lincoln Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

S.K. Mitter
Electronic Systems Laboratory
Department of Electrical Engineering
and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

An infinite-dimensional model is given for the generation of gyroscopic noise, which exhibits power spectral density proportional to $(1/f)$ over a wide frequency range. The optimal filter is given for separating a statistically described signal from additive gyroscopic noise, using discrete-time observations. This filter is expressed as a discrete-time infinite-dimensional Kalman-Bucy filter, with an associated Riccati covariance operator equation. Sufficient conditions are specified such that this Kalman-Bucy filter will possess various desired properties.

1. Introduction

The gyroscope is an instrument used to detect angular movement. The problem of the removal from the gyro output signal of noise inherent to the gyroscope in a constant gravitational field is one which has received considerable attention in the literature. Sutherland and Gelb [1], for example, discuss an aided inertial guidance system, where periodic telescopic sightings are used along with gyro output to develop gyro error observations. The error observations are used as the input to a Kalman filter, which is used to estimate the gyro error at the observation times. An estimate of the true angular position is then obtained by subtracting the estimated gyro error from the gyro output samples. Mehra and Bryson [2] discuss smoothing of the gyro output to obtain estimates of the input signal.

Gyroscopic noise has often been modeled as either a first-order Gauss-Markov process [3], or as a Gaussian random walk (integral of Gaussian white noise) [4,5]. However, recent studies performed at The Charles Stark Draper Laboratory [6] of the power spectral characteristics of the random noise associated with various gyroscopes indicate that gyro noise is often characterized by a $(1/f)$ behavior in power spectral density over a wide frequency range. An explanation of the source of this noise in the magnetic materials of the gyroscope (e.g. the gyro float rebalance torquer) is

The research of the first author was supported by the National Science Foundation under a Graduate Fellowship. The research of the second author was supported by the Air Force Office of Scientific Research under grant AF-AFOSR-72-2273.

proposed by Harris and Koenigsberg [7]. In Section 2 we discuss their findings and add others. We present an infinite-dimensional state space model which generates noise with the power spectral properties of gyroscopic noise. We also discuss the possible relationship between magnetic disaccommodation and gyroscopic noise.

In Section 3 we introduce and solve the filtering problem to be treated in the paper. Using discrete-time observations, a statistically described gyro output signal (resulting from angular motion inputs to the gyroscope) is optimally separated from additive gyroscopic noise. Because observations are made at discrete times, we first determine a discrete-time infinite-dimensional linear system to generate samples of the gyroscopic noise, as modeled in continuous time in Section 2. The filtering problem can be solved as a conditional expectation filter in the case where the input signal is Gaussian (this solution being equivalent to the minimum variance linear estimator for non-Gaussian input signals). The resulting optimal filter is expressed as a discrete-time infinite-dimensional Kalman filter with an associated Riccati covariance operator equation. We note here that steady-state filtering of a random process with a $(f^{-1-2\epsilon})$ power spectrum has been discussed by Moran [8]. However, the performance of Moran's filter degrades as $\epsilon \rightarrow 0$.

We indicate how theorems concerning Hilbert space Kalman filters and Riccati operator equations can be applied to the gyro noise filtering problem. By specifying conditions on the system generating the signal to be recovered, we are able to guarantee a number of desirable properties for the Kalman filter.

The optimal filter derived in Section 3 involves integrations over a free time constant parameter. In applications, these integrations must be implemented discretely. This discretization can be achieved by making a finite-dimensional approximation to the infinite-dimensional gyroscopic noise model. The optimal filter becomes an ordinary finite-dimensional discrete-time Kalman filter, with an associated matrix Riccati equation. It can be shown [19] that the mean-squared estimation error incurred in using the Kalman filter of the finite-dimensional approximate model can be made, through the use of a sufficient number of dimensions in the approximation, to approach the mean-squared estimation error associated with optimal filtering of gyroscopic noise.

The results found here for gyroscopic noise are applicable to any random process characterized by a wideband (1/f) power spectral density, as given in Section 2. (1/f) spectra are found, for example in semiconductor flicker noise and in the noise characterizing the frequency fluctuations of quartz crystal oscillators.

2. An Infinite-Dimensional Model for Gyroscopic Noise

Gyroscopic noise has often been modeled as either a first-order Gauss-Markov process [3], or as a Gaussian random walk (integral of Gaussian white noise) [4,5]. The Gaussian nature of the noise is inferred from histogram plots of gyro output. A linearized version on log-log scales of the power spectral density of a first-order Gauss-Markov process is shown in Figure 1. The random walk has a variance proportional to time, hence is nonstationary. Thus in a strict sense the power spectral density of a random walk process does not exist. When discrete samples of bandlimited white noise are generated by computer and summed (to resemble the integration of white noise), the resulting noise is found to be characterized by a (1/f²) power spectral density over the bandwidth of the original bandlimited white noise. (The power spectral density is found through evaluation of the squared magnitudes of the Fourier coefficients of the output signal.) For the following reasons we intuitively expect this result. The power spectral density, S_{yy}(f), of the output of a time-invariant linear system (transfer functions H(f)) to an input signal of PSD (power spectral density) S_{xx}(f) is given by:

$$S_{yy}(f) = S_{xx}(f) \cdot |H(f)|^2 \quad (1)$$

The transfer function of an integrator is proportional to (1/s), hence we would have:

$$|H(f)|^2 = \frac{1}{(j2\pi f)(-j2\pi f)} = \frac{1}{4\pi^2 f^2} \quad (2)$$

Bandlimited white noise has a PSD constant with frequency (over its band limits), so we would intuitively expect our approximation to random walk to have behavior proportional to (1/f²). The PSD resulting from the computer simulation described above is shown in Figure 2. Notice that both random processes discussed here exhibit (1/f²) behavior in PSD (slopes of (-2) on log-log scales).

Recent studies performed at The Charles Stark Draper Laboratory [6] of the power spectral characteristics of the random noises associated with various gyroscopes indicate that gyro noise is often characterized by a (1/f) behavior in power spectral density. (The gyro is set up as an input rate integrator, with a binary torque rebalance loop. The units of PSD are (input rate)²/Hz.) A linearized graph of the observed form of gyro power spectral density is given in Figure 3. The (f²) portion of this graph is primarily attributed to quantization noise due to the binary torque loop. This effect of quantization is currently under investigation. Power spectral analyses of separate record lengths of gyro noise show the power spectral density to be constant in time, hence we will treat the gyro noise as stationary. An explanation

of the source of this noise in the magnetic materials of the gyroscope (e.g. the gyro float rebalance torquer) is proposed by Harris and Koenigsberg [7]. In this section we shall discuss their findings and add others. We first discuss a model for magnetic relaxation (disaccommodation). This model is then used to develop an infinite-dimensional state space model for the generation of gyroscopic noise.

Examination of the literature on magnetic relaxation (e.g. Ref. [9]) indicates that the response of iron to transients in applied magnetic field can be characterized as the impulse response of a continuum of first-order linear systems with a uniform volume density distribution of time constants. The term "uniform volume density distribution" is used here to mean a spatial distribution of systems such that each volume element contains many systems, and such that the systems in each volume element have time constants distributed according to the same probability density function. Each individual system is characterized by a transfer function of the form:

$$G_{\tau}(s) = \frac{\tau}{\tau s + 1} \quad (3)$$

The probability density function of time constants (τ) is given by (see insert in Figure 4):

$$P_d(\tau) = \begin{cases} (1/\ln(\tau_2/\tau_1))(1/\tau); & \tau_1 \leq \tau \leq \tau_2 \\ 0 & ; \text{otherwise} \end{cases} \quad (4)$$

We shall demonstrate that the above density function is effective in explaining the gyro noise PSD in addition to magnetic relaxation, which is observed when the gyro is operated in the presence of power supply transients. Incidentally, other possible density forms (for anelastic relaxation of strain in crystalline solids, a related phenomenon) are discussed by Nowick and Berry [10]. The impulse response of each linear system (Eq. (3)) is given by:

$$h_{\tau}(t) = e^{-t/\tau} \quad (5)$$

The magnetic relaxation of the material is then characterized (see Ref. [9]) by the weighted integral of the impulse responses of the linear systems, with the time constant density of (Eq. 4):

$$m(t) \stackrel{\Delta}{=} K \int_0^{\infty} P_d(\tau) h_{\tau}(t) d\tau \quad (6)$$

Substituting Eq. (4) and Eq. (5) into Eq. (6), we find that:

$$m(t) = \frac{K}{\ln(\tau_2/\tau_1)} \int_{\tau_1}^{\tau_2} (1/\tau) (e^{-t/\tau}) d\tau \quad (7)$$

changing variables, we obtain:

$$m(t) = \frac{K}{\ln(\tau_2/\tau_1)} \int_{t/\tau_2}^{t/\tau_1} (e^{-y}/y) dy \quad (8)$$

where we have made the substitution:

$$y = t/\tau \quad (9)$$

Finally, we obtain:

$$m(t) = \frac{K}{\ln(\tau_2/\tau_1)} [E_1(t/\tau_2) - E_1(t/\tau_1)] \quad (10)$$

Where $E_1(z)$ is the exponential integral, defined by:

$$E_1(z) = \int_z^\infty \left(\frac{e^{-u}}{u}\right) du \quad (11)$$

We choose K to normalize $m(t)$ to $\psi(t)$, where we require for normalization that:

$$\psi(0) = 1; \psi(\infty) = 0 \quad (12)$$

We find that:

$$K = 1 \quad (13)$$

Thus, the magnetic disaccommodation (relaxation) is normalized to:

$$\psi(t) = \left(\frac{1}{\ln(\tau_2/\tau_1)}\right) [E_1(t/\tau_2) - E_1(t/\tau_1)] \quad (14)$$

Graphs of $\psi(t)$, for $\tau_1 = 0.01$, $\tau_2 = 1.0$, on linear-linear, semilog, and log-log scales are found in Figures 4, 5, and 6, respectively. As discussed in Ref. (7), $\psi(t)$, with proper choice of τ_1 and τ_2 , often fits the time record of gyro output in the presence of power supply transients. Gyro output is the record of the torques applied by the magnetic gyro torquer in order to keep the gyro float angle close to zero. For t between τ_1 and τ_2 , $\psi(t)$ is proportional to $(-\ln(t))$, a familiar result in the study of magnetic relaxation (see Ref. (11)). (Incidentally, τ_1 and τ_2 may be estimated by observing the gyro output and using an analytic approximation (12) for $\psi(t)$, for t between τ_1 and τ_2 .) In summary, the time constant density given in Eq. (4) can be used to explain the deterministic gyro response to transients. The reader should be aware that we do not have empirical confirmation that the relaxation exhibited by gyro output in the presence of power supply transients is necessarily magnetic in origin. We can only suggest this as a possible source, and note that this mechanism is effective in explaining the observed power spectral characteristics of gyroscopic output noise, which we shall now discuss.

If a linear system (Eq. (3)) with time constant (T) is fed by an input function $w(\tau, t)$ then its response, $x(\tau, t)$, is characterized by:

$$\frac{\partial}{\partial t} x(\tau, t) = -\left(\frac{1}{T}\right)x(\tau, t) + w(\tau, t) \quad (15)$$

Let the input function of two variables, $w(\tau, t)$, to the systems be characterized by covariance: ($\delta(\cdot)$ is the Dirac delta function)

$$E\{w(\tau, t)w(\gamma, t-\alpha)\} \stackrel{\Delta}{=} W\delta(\tau-\gamma)\delta(\alpha) \quad (16)$$

$w(\tau, t)$ is formally a "two-dimensional white noise". The inputs to two systems with time constants τ and γ are independent if $\tau \neq \gamma$. Eq. (12) may be regarded as a state equation, where state $x(\tau, t)$ is a function of $\tau \in [\tau_1, \tau_2]$. Gyroscopic noise is now modeled as the weighted integral of the outputs of the filters (where $x(\tau, t)$ is the output (at time t) of a filter with time constant T), and is given by:

$$g(t) = \int_0^\infty x(\tau, t) p_d(\tau) d\tau \quad (17)$$

In more rigorous form, Eq. (15), (16) and (17) are shorthand for:

$$g(t) = \int_{\tau_1}^{\tau_2} p_d(\tau) e^{-t/\tau} d\mu(\tau, 0) + \int_{\tau_1}^{\tau_2} \int_0^t p_d(\tau) e^{-(t-s)/\tau} d\beta(\tau, s) \quad (18)$$

where the first integral, the initial condition propagation, is a Wiener integral and the second is a "two-dimensional Wiener integral", defined in Appendix A. In this appendix we also discuss the two-dimensional Wiener process $\beta(\tau, s)$ whose (formal) mixed double partial derivative is the two-dimensional white noise, $w(\tau, s)$, in Eq. (13). Further, as discussed in Appendix A, normalization of $g(t)$ so that the noise has unit variance requires:

$$W = 2\ln(\tau_2/\tau_1) \quad (19)$$

The power spectral density of the noise is then given by:

$$S_{gg}(f) = \left(\frac{2}{\ln(\tau_2/\tau_1)}\right) \left(\frac{1}{2\pi f}\right) \tan^{-1} \left[\frac{2\pi f(\tau_2 - \tau_1)}{1 + 4\pi^2 f^2 \tau_2 \tau_1} \right] \quad (20)$$

A graph of $S_{gg}(f)$ is plotted in Figure 7. Note that the $(1/f)$ characteristic of gyro noise observed experimentally is inherent in the linearized version of this plot. The (f^2) section of Figure 3, the experimentally observed gyro noise, due to quantization dominates over the $(1/f^2)$ line of Figure 7 at high frequencies, masking that portion of the gyro noise. Further, it is felt that the low frequency breakpoint of Figure 7 corresponds to times longer than the record lengths normally employed for observations of gyro output, accounting for its absence from Figure 3 (see caption of Figure 2). Ongoing experiments at The Charles Stark Draper Laboratory with long record lengths of gyro output indicate that the power spectral density is flat at very low frequencies (~ 1 cycle/month) for some of the gyroscopes being tested.

Henceforth, we shall use the term gyroscopic noise to refer to the stochastic process generated by our state space model. The gyroscopic noise is assumed to have started at $(t=-\infty)$, hence to be stationary at $(t=0)$.

Incidentally, alternative models, in terms of diffusion mechanisms, for stochastic processes with the power spectral characteristics of gyroscopic noise are discussed in Ref. [13] and [14]. Another model, with a $(f^{-1-2\epsilon})$ power spectrum, is discussed in Ref. [15]. Note however that because our gyro noise filter is a linear estimator only the second-order properties of the gyroscopic noise influence the filter mean-squared error

sequence. Thus all mathematical models which generate stochastic processes with the same PSD as gyro noise (hence the same second-order properties as gyro noise) will yield the same optimal (minimum variance) filter. In the next section we formulate the problem and indicate how it can be solved in an infinite dimensional context.

3. Filtering and Properties of the Filter

Take X, U real separable Hilbert spaces, (Ω, S, μ) a complete probability space.

3.1 Separable Hilbert space-valued random variables

The reader is referred to [20] for more detailed exposition of this material.

$x: \Omega \rightarrow X$ is called an X-valued random variable (r.v.) if it is a (weakly) measurable map. The linear space of X-valued r.v.'s is denoted $Mes(\Omega, \mu; X)$.

An X-valued stochastic process is a map $x(\cdot): \mathbb{R}^+ \rightarrow Mes(\Omega, \mu; X)$. $x(\cdot)$ is a measurable process if the map $(t, \omega) \rightarrow x(t, \omega)$ is measurable w.r.t. $\mu_L \times \mu$ (μ_L denotes Lebesgue measure on \mathbb{R}^+). $x \in Mes(\Omega, \mu; X)$ is first order if $x \in L^1[\Omega, \mu; X]$ and second order if $x \in L^2[\Omega, \mu; X]$. For a first order r.v. $x(\omega)$ we define the mean $E\{x(\omega)\}, (\bar{x})$

$$E\{x(\omega)\} = \int_{\Omega} x(\omega) d\mu \quad (\text{Bochner Integral})$$

For a second order r.v. $x(\omega)$, $(h, \bar{h}) \rightarrow E\{ \langle x(\omega) - E\{x(\omega)\}, h \rangle \langle x(\omega) - E\{x(\omega)\}, \bar{h} \rangle \}$ is a continuous, symmetric bilinear form which has unique representation through $Q \in L(X)$, $Q \geq 0$, $Q^* = Q$ as $(h, \bar{h}) \rightarrow \langle Qh, \bar{h} \rangle$. Q is the covariance of $x(\omega)$. The covariance of a second order random variable $x(\omega)$ is necessarily nuclear.

Given two X-valued second order r.v.'s $x(\omega)$, $y(\omega)$, $(h, \bar{h}) \rightarrow E\{ \langle x(\omega) - \bar{x}, h \rangle \langle y(\omega) - \bar{y}, \bar{h} \rangle \}$ has unique representation $(h, \bar{h}) \rightarrow \langle Rh, \bar{h} \rangle$, $R \in L(X)$. R is called the covariance of $x(\omega)$, $y(\omega)$ and is written $cov\{x(\omega), y(\omega)\}$.

$x(\omega)$, $y(\omega) \in Mes(\Omega, \mu; X)$ are independent if $\langle h, x(\omega) \rangle$, $\langle \bar{h}, y(\omega) \rangle$ are independent for all $h, \bar{h} \in X$. $x \in L^2[\Omega, \mu; X]$ is Gaussian if $\langle x(\omega), h \rangle$ is normally distributed for each $h \in X$.

3.2 Wiener Process

The U-valued stochastic process $W(t, \omega)$ is a Wiener process if (i) for finite collections $\{t_i\} \in \mathbb{R}^+, \{e_j\} \in U, (w(t_i, \omega), e_j)$ is a family of real-valued gaussian r.v.'s (ii) $W(t, \omega)$ is second order for each $t \geq 0$ and there exists some nuclear $Q \in L(X)$ s.t.

$$\begin{aligned} E\{ \langle w(t_1, \omega), h \rangle \langle w(t_2, \omega), \bar{h} \rangle \} \\ = \langle Qh, \bar{h} \rangle \min\{t_1, t_2\} \end{aligned}$$

each $t_1, t_2 \geq 0, h, \bar{h} \in U$, (iii) $E\{w(t, \omega)\} = 0$ for each $t \geq 0$. See ([20], p. 167 et seq.) for properties of $w(t, \omega)$.

Notice that since Q is nuclear, $Q \geq 0$, $Q^* = Q$

$$Q(\cdot) = \sum_i \lambda_i e_i \langle e_i, \cdot \rangle$$

for some $\{\lambda_i\}, \lambda_i \geq 0$ with $\sum_i \lambda_i < \infty$, some orthonormal sequence $\{e_i\}$ in U . We shall make use of the property that the Wiener process $W(t, \omega)$ has unique representation

$$w(t, \omega) = \lim_{N \rightarrow \infty} \sum_{i=1}^N \beta_i(t, \omega) e_i$$

(limit in $L^2[\Omega, \mu; X]$)

with the β_i 's independent real valued Wiener processes.

3.3 The Wiener Integral

Suppose $b: \mathbb{R}^+ \rightarrow X$ is locally essentially bounded, measurable and that $\beta(t, \omega)$ is a real-valued Wiener process. Then the Wiener Integral

$$\int_0^T b(t) d\beta(t, \omega)$$

is defined in the usual manner as a limit in $L^2[\Omega, \mu; X]$ through a sequence of simple functions approximating $b(t)$ in $L^2[0, T; X]$. Now suppose that $B(\cdot): \mathbb{R}^+ \rightarrow L(U, X)$ satisfies (i) $\|B(\cdot)\|$ is locally essentially bounded, measurable (ii) $t \rightarrow B(t)x$ is measurable for each $x \in X$. The Wiener Integral

$$\int_0^T B(t) dw(t, \omega)$$

is defined in this case as

$$\lim_{N \rightarrow \infty} \sum_{i=1}^N \int_0^T B(t) e_i d\beta(t, \omega)$$

(limit in $L^2[0, T; X]$)

where each element in the sequence is evaluated as above. ($e_i, (t, \omega) i = 1, w, \dots$ as in Sec. 3.2) For $B(\cdot)$ measurable w.r.t. the uniform operator topology this definition coincides essentially with that in ([20], p. 180 et seq.). Notice that the Wiener Integral is defined modulo null-functions in $L^2[\Omega, \mu; X]$.

3.4 Infinite Dimensional Formulation of the Filtering Problem

We first show how equations (15) and (17) can be represented in the infinite-dimensional stochastic setting just described. Let $X = L^2(\tau_1, \tau_2; R)$ be the space of square-integrable functions with values in R , and let $\langle \cdot, \cdot \rangle$ denote the natural scalar product on X . All random variables are considered with respect to some fixed complete probability space (Ω, A, P) .

Denote by $\beta(t) = w(\cdot, t)$ the X-valued Wiener process with covariance operator W obtained from the two dimensional Wiener process $w(\tau, t)$.

Consider $A: X \rightarrow X: x(\tau, t) \rightarrow \frac{1}{\tau} x(\tau, t)$. This is clearly a bounded linear operator.

Let $y(t) = x(\cdot, t)$ be an X-valued random variable, given as the solution of the integral equa-

tion

$$y(t) = e^{At} y_0 + \int_0^t e^{A(t-\sigma)} d\beta(\sigma) \quad (21)$$

where y_0 is an X -valued Gaussian random variable with zero mean and covariance operator Π . It is assumed that y_0 and $\beta(t)$ are independent. It can be shown that (i) $y(t)$ is a Gaussian random variable, (ii) $E(y(t)) = 0$ and (iii) writing $\Lambda(t,s) = \text{cov}[y(t), y(s)]$, $\langle \Lambda(t,s)h, \bar{h} \rangle = \langle e^{At} \Pi e^{A^*t} h, \bar{h} \rangle +$

$$+ \int_0^{\min(t,s)} \langle e^{A^*(s-\sigma)} h, \langle e^{A^*(t-\sigma)} \bar{h} \rangle \rangle d\sigma \quad \forall h, \bar{h} \in X.$$

A^* denotes the adjoint of A .

The output equation of the gyro noise model is

$$g(t) = \int_{\tau_1}^{\tau_2} p_d(\tau) x(\tau, t) d\tau \quad (22)$$

where $p_d(\cdot)$ is a bounded measurable function. The above defines a bounded linear operator $C: X \rightarrow R$. (21) and (22) constitute the infinite dimensional representation of the gyro model. In practice, the gyro output is sampled. Then, doing a "sampled-data" approximation to (21) and (22) we obtain the discrete-time representation

$$\begin{cases} y(n+1) = \Psi y(n) + \beta(n) \\ y(0) = y_0 \end{cases} \quad (23)$$

$$g(n) = \int_{\tau_1}^{\tau_2} p_d(\tau) x(\tau, n) d\tau, \quad (24)$$

$$n = 0, 1, 2, \dots$$

Here $\Psi \in L(X, X)$ is the mapping $f(\tau) \rightarrow e^{-\frac{T}{T}} f(\tau)$ where $T > 0$ is the sample-time increment, and $\alpha(n)$ is an independent sequence of X -valued Gaussian random variables with covariance operator \hat{W} (which can be calculated from W and the sampling data). y_0 and $\beta(n)$ are assumed to be independent.

Now let a finite-dimensional discrete-time linear stochastic system be given by

$$a(n+1) = a(n) + Bu(n) \quad (25)$$

$$p_1(n) = h'a(n) \quad (26)$$

Here $a(0)$ is a Gaussian R^n -valued random variable with mean 0 and covariance P , $u(n)$ is a "white" Gaussian sequence with mean zero and covariance Q_1 , Φ and B are matrices of appropriate size and h a vector.

The observation equation is

$$z(n) = p_1(n) + g(n) + v(n), \quad (27)$$

where $v(n)$ is a white gaussian scalar sequence with zero mean and covariance $\gamma > 0$.

$p_1(n)$ is the ideal gyro noise output which needs to be estimated. In order to estimate it we have to estimate $a(n)$ and $y(n)$. This filtering problem can now be solved using standard infinite-dimensional filtering methods (see, for example, [21]). By duality arguments, it can be shown that this problem is equivalent to the following optimal

control problem (in backward time)

$$\tilde{x}(t+1) = F\tilde{x}(t) + Gu(t), \quad t = 0, \dots, T-1 \quad (28)$$

with cost functional of the form

$$\begin{aligned} (S\tilde{x}(T), \tilde{x}(T)) + \sum_{t=0}^{T-1} \left[\int_{\tau_1}^{\tau_2} q(\tau) x(\tau, t)^2 d\tau \right. \\ \left. + a(t)' Q a(t) + \alpha u(t)^2 \right] \end{aligned} \quad (29)$$

(\cdot, \cdot) denotes the natural scalar produce on $X \times R^n$.

In the above,

$$\tilde{x}(t) = (y(t), z(t)) \in X \times R^n$$

$$u(t) \in R$$

$F: X \times R^n \rightarrow X \times R^n$ is the bounded linear mapping

$$(f(\tau), a) \rightarrow (e^{-\frac{T}{T}} f(\tau), \Phi a), \quad T > 0,$$

$G: R \rightarrow X \times R^n$ is the bounded linear mapping

$u \rightarrow (p(\cdot)u, hu)$ where $p(\cdot)$ is bounded measurable and $h \in R^n$, S is a positive operator, Q is a symmetric positive semidefinite matrix, $q(\tau) \geq c > 0$ is a bounded measurable function and $\alpha > 0$ is a scalar. What is of interest is the asymptotic behavior of this control problem (that is, the asymptotic behavior of the filter).

We say that (28) is reachable if there exists an integer $\gamma \geq 0$ and a constant $0 < a < \infty$ such that

$$(\tilde{x}, \sum_{i=0}^{\gamma} F^i G G^* F^{i*} \tilde{x}) \geq a \|x\|^2 \quad \forall \tilde{x} \in X \times R^n$$

Let us check whether this is possible for the X part of the system. We would require

$$\begin{aligned} \sum_{i=0}^{\gamma} \left\{ \int_{\tau_1}^{\tau_2} p(\tau) e^{-\frac{T}{T}} f(\tau) d\tau \right\}^2 \\ \geq a \int_{\tau_1}^{\tau_2} f(\tau)^2 d\tau \quad \text{for some } a > 0 \end{aligned}$$

and $\forall f \in L^2(\tau_1, \tau_2; R)$ which is clearly impossible. The X -part of the system is however stable. Now

$$\int_{\tau_1}^{\tau_2} q(\tau) x(\tau, t)^2 d\tau = \langle \sqrt{q(\cdot)} x(\cdot, t), \sqrt{q(\cdot)} x(\cdot, t) \rangle.$$

Hence the mapping $x(\cdot, t) \rightarrow \sqrt{q(\cdot)} x(\cdot, t) = \tilde{Q}x(\cdot, t)$ may be thought of as an observation equation for the X -part of the system. We say that the X -part of (28) with the above observation equation is observable if \exists integer $s \geq 0$ and a constant $0 < b < \infty$ such that

$$(x, \sum_{i=0}^s F^{*i} \tilde{Q} F^i x) \geq b \|x\|^2 \quad \forall x \in X.$$

Since $q(\tau) \geq c > 0$ the X -part of the system is observable. Now assume that $a(k+1) = \Phi a(k) +$

$hu(k)$ is stabilizable and $\sum_{k=0}^s \Phi^{*k} Q \Phi^k > 0$ for some s .

It then follows from the results of Hager and Horowitz [22], on the asymptotic behavior of Discrete Riccati Operator Equations, that the resulting filter is asymptotically stable. For the appropriate concepts of filter stability see the forthcoming paper by Vinter [23].

Appendix A
The Two-Dimensional Wiener Process

In this appendix, we discuss the two-dimensional Wiener process, with covariance:

$$E[\beta(\tau, t)\beta(s, \sigma)] = W \cdot \min(\tau, s) \cdot \min(t, \sigma) \quad (A-1)$$

Note that, formally, the mixed double partial derivative of this process will have the covariance of a two-dimensional white noise, because, formally:

$$E\left[\left(\frac{\partial^2}{\partial \tau \partial t} \beta(\tau, t)\right) \left(\frac{\partial^2}{\partial \sigma \partial s} \beta(s, \sigma)\right)\right] = \frac{\partial^4}{\partial \tau \partial t \partial \sigma \partial s} E[\beta(\tau, t)\beta(s, \sigma)] \quad (A-2)$$

and from Eq. (A-1) we have that:

$$\frac{\partial^4}{\partial \tau \partial t \partial \sigma \partial s} E[\beta(\tau, t)\beta(s, \sigma)] = W \cdot \delta(\tau-s) \cdot \delta(t-\sigma) \quad (A-3)$$

We shall first show the existence of a (Gaussian) random process with the covariance in Eq. (A-1). We then shall give meaning to a two-dimensional Wiener integral (where $f \in L^2([0, \infty) \times [0, \infty))$):

$$\int_0^\infty \int_0^\infty f(\tau, t) d\beta(\tau, t) \quad (A-4)$$

Finally, we shall show that the model we have given for gyroscopic noise in Eq. (18):

$$g(t) = \int_{\tau_1}^{\tau_2} p_d(\tau) e^{-t/\tau} d\mu(\tau, 0) + \int_{\tau_1}^{\tau_2} \int_0^t p_d(\tau) e^{-(t-s)/\tau} d\beta(\tau, s) \quad (A-5)$$

in fact yields the desired power spectral density (Eq. 20):

$$S_{gg}(f) = \left(\frac{2}{2\pi(\tau_2/\tau_1)}\right) \frac{1}{2\pi f} \cdot \tan^{-1} \left[\frac{2\pi f(\tau_2 - \tau_1)}{1 + 4\pi^2 f^2 \tau_2 \tau_1} \right] \quad (A-6)$$

Further details concerning the multiparameter Wiener process and Wiener integral are discussed by Park [1].

We first show the existence of the two-dimensional Wiener process. The argument closely parallels that of J.M.C. Clark [17]. Choose two sets of complete orthonormal functions in $L^2([0, \infty))$: (where $\langle \cdot, \cdot \rangle$ is scalar product notation)

$$\{\psi_i(t)\} \dots \langle \psi_i, \psi_j \rangle = \int_0^\infty \psi_i(t) \psi_j(t) dt =$$

$$= \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}; \quad i=1, 2, \dots \quad (A-7)$$

$$\{\phi_i(t)\} \dots \langle \phi_i, \phi_j \rangle = \int_0^\infty \phi_i(\tau) \phi_j(\tau) d\tau =$$

$$= \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}; \quad i=1, 2, \dots \quad (A-8)$$

(Note that $\{\psi_i(t)\}$ and $\{\psi_i(\tau)\}$ may be the same set of functions.) Next, define a sequence of doubly-indexed Gaussian random variables:

$$\{a_{ij}\} \dots E[a_{ij} a_{mq}] = W \cdot \delta_{im} \cdot \delta_{jq};$$

$$\begin{cases} i=1, 2, \dots \\ j=1, 2, \dots \end{cases} \quad (A-9)$$

where δ_{im} is the Kronecker delta function, defined by

$$\delta_{im} = \begin{cases} 0, & i \neq m \\ 1, & i = m \end{cases} \quad (A-10)$$

We now define a sequence of random processes $\{\beta^N(\tau, t)\}$. Each random process is a function of two variables, (τ) and (t) . The definition is given by:

$$\beta^N(\tau, t) = \sum_{i=0}^N \sum_{j=0}^N a_{ij} \int_0^\tau \int_0^t \psi_i(\gamma) \phi_j(\alpha) d\alpha d\gamma \quad (A-11)$$

Fix (τ) and (t) . We claim that $\{\beta^N(\tau, t)\}$ is a quadratic mean Cauchy convergent sequence of random variables. Observe that: (say $M > N$)

$$E[(\beta^M(\tau, t) - \beta^N(\tau, t))^2]$$

$$= E\left[\sum_{i=N+1}^M \sum_{j=N+1}^M a_{ij} \int_0^\tau \int_0^t \psi_i(\gamma) \phi_j(\alpha) d\alpha d\gamma \times \sum_{k=N+1}^M \sum_{q=N+1}^M a_{kq} \int_0^\tau \int_0^t \psi_k(\lambda) \phi_q(\eta) d\eta d\lambda\right] \quad (A-12)$$

By Eq. (A-9), we obtain:

$$E[(\beta^M(\tau, t) - \beta^N(\tau, t))^2] = W \cdot \left[\sum_{i=N+1}^M \left(\int_0^\tau \psi_i(\gamma) d\gamma\right)^2\right] \cdot \left[\sum_{j=N+1}^M \left(\int_0^t \phi_j(\lambda) d\lambda\right)^2\right] \quad (A-13)$$

Define the following function:

$$I_t(\gamma) = \begin{cases} 1, & 0 \leq \gamma \leq t \\ 0, & \gamma > t \end{cases} \quad (A-14)$$

We may now express Eq. (A-13) in dot product notation:

$$E[(\beta^M(\tau, t) - \beta^N(\tau, t))^2] = W \cdot \left[\sum_{i=N+1}^M \langle \psi_i, I_t \rangle^2\right] \cdot \left[\sum_{j=N+1}^M \langle \phi_j, I_t \rangle^2\right] \quad (A-15)$$

By the orthonormality of the sequences $\{\psi_i(\gamma)\}$ and $\{\phi_j(\lambda)\}$ we have that each factor on the right of Eq. (A-15) approaches zero as $N \rightarrow \infty$. Thus $\{\beta_N(\tau, t)\}$ is a quadratic mean Cauchy convergent sequence of random variables. Call the limit $\beta(\tau, t)$. We now demonstrate that $\beta(\tau, t)$ has covariance as in Eq. (A-1):

$$E[\beta(\tau, t)\beta(s, \sigma)] = W \cdot \left[\sum_{i=0}^{\infty} \langle \psi_i, I_t \rangle \langle \psi_i, I_\sigma \rangle \right] \cdot \left[\sum_{j=0}^{\infty} \langle \phi_j, I_\tau \rangle \langle \phi_j, I_s \rangle \right] \quad (A-16)$$

By Parseval's theorem, we have that:

$$E[\beta(\tau, t)\beta(s, \sigma)] = W \cdot \langle I_t, I_\sigma \rangle \langle I_\tau, I_s \rangle \\ = W \cdot \left(\int_0^\infty I_t(\gamma) I_\sigma(\gamma) d\gamma \right) \left(\int_0^\infty I_\tau(\gamma) I_s(\gamma) d\gamma \right)$$

$$E[\beta(\tau, t)\beta(s, \sigma)] = W \cdot \min(t, \sigma) \cdot \min(\tau, s) \quad (A-17)$$

We have thus established the existence of the two-dimensional Wiener process.

We now wish to define a two-dimensional Wiener integral. We first develop an analog of the "orthogonal increment" property of Brownian motion. For a Brownian motion, $\mu(t)$, it is well known (17) that:

($\lambda(\cdot)$ denotes length)

$$E[(\mu(t) - \mu(s))(\mu(q) - \mu(r))] = K \cdot \lambda([s, t] \cap [r, q]) \quad (A-18)$$

Define the following function, over a box ($[\tau_1, \tau_2] \times [t_1, t_2]$) in the first quadrant (no generality is lost by defining our process for $\tau \geq 0, t \geq 0$) of the R^2 plane:

$$F(\tau_1, \tau_2, t_1, t_2) = \beta(\tau_2, t_2) - \beta(\tau_2, t_1) - \beta(\tau_1, t_2) + \beta(\tau_1, t_1) \quad (A-19)$$

Note, incidentally, that if a deterministic function $q(\tau, t) \in C^2$ had its mixed double partial derivative integrated over this domain, we would obtain:

$$\int_{\tau_1}^{\tau_2} \int_{t_1}^{t_2} \frac{\partial^2 q(\tau, t)}{\partial \tau \partial t} dt d\tau = q(\tau_2, t_2) - q(\tau_2, t_1) - q(\tau_1, t_2) + q(\tau_1, t_1) \quad (A-20)$$

This can be taken as motivation for Eq. (A-19). It is easily seen, using only Eq. (A-17) that: ($A(\cdot)$ denotes area)

$$E[F(\tau_1, \tau_2, t_1, t_2)F(\tau_3, \tau_4, t_3, t_4)] \\ = W \cdot A([[\tau_1, \tau_2] \times [t_1, t_2]] \cap [[\tau_3, \tau_4] \times [t_3, t_4]]) \quad (A-21)$$

Equation (A-21) is the analog to Eq. (A-18). Using this property, we proceed, as Wong (17) does for a one-dimensional orthogonal increment process, to define the two-dimensional Wiener integral.

(1) If $f = I[a_1, a_2] \times [b_1, b_2]$, the indicator func-

tion of the rectangle $[a_1, a_2] \times [b_1, b_2]$, we set:

$$\int_0^\infty \int_0^\infty f(\tau, t) d\beta(\tau, t) = F(a_1, a_2, b_1, b_2) \quad (A-22)$$

(2) If $f = \sum_{v=1}^n a_v f_v$, with f_v functions as in (1),

we set:

$$\int_0^\infty \int_0^\infty f(\tau, t) d\beta(\tau, t) = \sum_{v=1}^n a_v \int_0^\infty \int_0^\infty f_v(\tau, t) d\beta(\tau, t)$$

(3) If $\int_0^\infty \int_0^\infty |f_n(\tau, t) - f(\tau, t)|^2 dt d\tau \rightarrow 0$, we set:

$$\int_0^\infty \int_0^\infty f(\tau, t) d\beta(\tau, t) = \lim_{n \rightarrow \infty} \text{in q.m.} \int_0^\infty \int_0^\infty f_n \cdot (\tau, t) d\beta(\tau, t) \quad (A-23)$$

The class of functions $f(\tau, t)$ for which this is possible is $L^2([0, \infty) \times [0, \infty))$. In addition, as in the one-dimensional case in Wong (18), we find that:

$$E \left\{ \left[\int_0^\infty f(\tau, t) d\beta(\tau, t) \right] \left[\int_0^\infty g(\tau, t) d\beta(\tau, t) \right] \right\} \\ = \int_0^\infty \int_0^\infty f(\tau, t) \overline{g(\tau, t)} dt d\tau \quad (A-24)$$

We shall make use of Eq. (A-24) in showing that our model for gyroscopic noise (Eq. (A-5) yields the desired power spectral density (Eq. (A-6)). From Eq. (A-5) we have that (for $\alpha \geq 0$):

$$E[g(t)g(t-\alpha)] = \left[\left(\int_{\tau_1}^{\tau_2} p_d(\tau) e^{-t/\tau} d\mu(\tau, 0) \right) \right. \\ \left. + \int_{\tau_1}^{\tau_2} \int_0^t p_d(\tau) e^{-(t-s)/\tau} d\beta(\tau, s) \right] \times \\ \times \left[\left(\int_{\tau_1}^{\tau_2} p_d(\lambda) e^{-(t-\alpha)/\lambda} d\mu(\lambda, 0) \right) \right. \\ \left. + \int_{\tau_1}^{\tau_2} \int_0^{t-\alpha} p_d(\lambda) e^{-[(t-\alpha)-\sigma]/\lambda} d\beta(\lambda, \sigma) \right] \quad (A-25)$$

The process $\mu(\tau, 0)$ is an initial " τ -axis-scaled" Brownian motion characterized by (analog of Eq. (A.24) in one dimension (18):

$$E \left[\left(\int_{\tau_1}^{\tau_2} f(\tau) d\mu(\tau, 0) \right) \left(\int_{\tau_1}^{\tau_2} g(\gamma) d\mu(\gamma, 0) \right) \right] \\ = \int_{\tau_1}^{\tau_2} f(\tau) \overline{g(\tau)} h(\tau) d\tau \quad (A-26) \\ \text{(for some } h(\tau))$$

Also, we have that $\beta(\tau, s)$ ($\tau \geq 0, s > 0$) is independent of $\mu(\tau, 0)$ ($\tau \geq 0$). Thus we obtain from Eq. (A-25) that:

$$\begin{aligned} E[g(t)g(t-\alpha)] &= e^{-\alpha/\tau} \int_{\tau_1}^{\tau_2} p_d^2(\tau) e^{-(t-\alpha)/\tau} h(\tau) d\tau \\ &+ e^{-\alpha/\tau} \int_{\tau_1}^{\tau_2} p_d^2(\tau) e^{-2(t-\alpha)/\tau} \int_0^{t-\alpha} e^{2s/\tau} w ds d\tau \end{aligned} \quad (A-27)$$

Integrate in Eq. (A-27) to obtain

$$\begin{aligned} E[g(t)g(t-\alpha)] &= e^{-\alpha/\tau} \int_{\tau_1}^{\tau_2} p_d^2(\tau) e^{-(t-\alpha)/\tau} h(\tau) d\tau \\ &+ e^{-\alpha/\tau} \int_{\tau_1}^{\tau_2} \left(\frac{W\tau}{2}\right) p_d^2(\tau) [1 - e^{-2(t-\alpha)/\tau}] d\tau \end{aligned} \quad (A-28)$$

Let $t \rightarrow \infty$ in Eq. (A-28), to obtain a stationary random process characterized by:

$$\begin{aligned} R_{gg}(\alpha) &= E[g(t)g(t-\alpha)] \\ &= \int_{\tau_1}^{\tau_2} \left(\frac{W\tau}{2}\right) e^{-|\alpha|/\tau} p_d^2(\tau) d\tau \end{aligned} \quad (A-29)$$

A Fourier transform of Eq. (A-29) yields:

$$S_{gg}(f) = \int_0^{\infty} \left(\frac{W}{1+4\pi^2 f^2 \tau^2}\right) [p_d(\tau)]^2 d\tau \quad (A-30)$$

Substituting Eq. (4) into Eq. (A-30) and using a normalization ($W = 2 \ln(\tau_2/\tau_1)$) to give $S_{gg}(f)$ unit variance, we obtain:

$$S_{gg}(f) = \left(\frac{2}{\ln(\tau_2/\tau_1)}\right) \left(\frac{1}{2\pi f}\right) \tan^{-1} \left[\frac{2\pi f(\tau_2 - \tau_1)}{1+4\pi^2 f^2 \tau_2 \tau_1} \right] \quad (A-31)$$

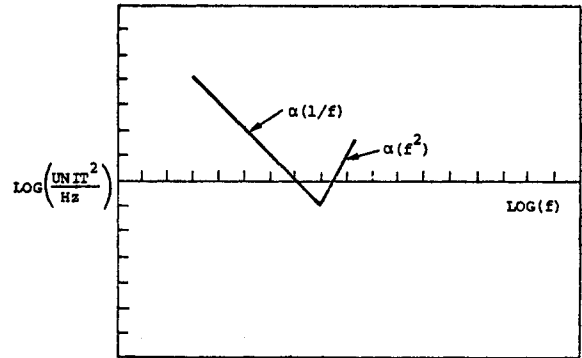
Equation (A-31) is the power spectral density for gyroscopic noise which we had in Eq. (20). Incidentally, note that our discussion of the two-dimensional Wiener process and two-dimensional Wiener integral can easily be extended to (n) dimensions.

References

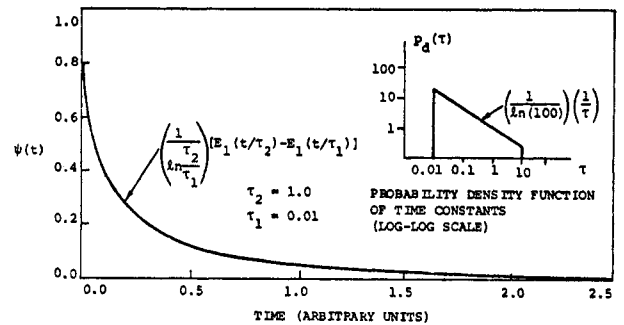
- [1] Sutherland, A.A., Jr., Gelb, A., "Application of the Kalman Filter to Aided Inertial Systems", The Analytic Sciences Corp., 1967.
- [2] Mehra, R.K., Bryson, A.E., Jr., "Linear Smoothing Using Measurements Containing Correlated Noise with an Application to Inertial Navigation", IEEE Transactions on Automatic Control, Vol. AC-13, No. 5, Oct. 1968, pp. 496-503.
- [3] McDonald, W.T., "A Gyrocompassing Spacecraft Navigator", Ph.D. Thesis, MIT Dept. of Aeronautics and Astronautics, 1968, pp. 51-62.
- [4] Wilkinson, R.H., "A Stochastic Model for the One-Degree-of-Freedom Inertial Gyro", MIT Instrumentation Laboratory Report E-2466, 1969.
- [5] Wilkinson, R.H., "Measurement of Gyro Stochastic Parameters", MIT Instrumentation Laboratory Report E-2467, 1970.
- [6] Truncala, A. Koenigsberg, W., Harris, R., "Spectral Density Measurements of Gyro Noise", MIT Charles Stark Draper Laboratory Report E-2641, 1972.
- [7] Harris, R.A., Koenigsberg, W.D., "Gyro Drift Noise, Measurement, Analysis, and Possible Sources", MIT Charles Stark Draper Laboratory, Sixth Biennial Guidance Test Symposium, Holloman Air Force Base, N.M., 1972.
- [8] Moran, K.W., "Linear Least-Square Prediction of Flicker Noise", Information Theory Symposium, 1974, p. 74.
- [9] Richter, G., "Uber die Magnetische Nachwirkung am Carbonyleisen", Ann. Physik, Vol. 29, 1937, pp. 605-635 (English trans. available through Dr. W. Koenigsberg, The Charles Stark Draper Laboratory, Inc., Cambridge, Mass., 02139).
- [10] Nowick, A.S., Berry, B.S., Anelastic Relaxation in Crystalline Solids, Academic Press, N.Y., 1972, pp. 40-114.
- [11] Chickazumi, S., Physics of Magnetism, John Wiley and Sons, 1964.
- [12] Stegun, I., Abramowitz, M., Handbook of Mathematical Functions, Dover Publications, Inc., N.Y., 1965, pp. 228-229.
- [13] Richardson, J.M., "The Linear Theory of Fluctuations Arising from Diffusional Mechanisms-An Attempt at a Theory of Contact Noise", Bell System Technical Journal, Vol. 29, 1950, pp. 117-140.
- [14] Van Vliet, K.M., Van der Ziel, A., "On the Noise Generated by Diffusion Mechanisms", Physica, Vol. XXIV, Zernike Issue, 1958, pp. 415-421.
- [15] Rutman, J. Vebersfeld, J., "A Model for Flicker Frequency Noise of Oscillators", Proc. IEEE, Vol. 60, No. 2, Feb. 1972, pp. 233-235.

- [16] Park, W.J., "A Multi-Parameter Gaussian Process," Ann. of Math. Stat., Vol. 41, No. 5, 1970, pp. 1582-1595.
- [17] Clark, J.M.C., "An Introduction to Stochastic Differential Equations on Manifolds", in Geometric Methods in System Theory, R.W. Brockett and D.Q. Mayne, eds., Reidel Publishing Co., The Netherlands, 1973.
- [18] Wong, E., Stochastic Processes in Information and Dynamical Systems, McGraw-Hill Book Co., N.Y., 1971, pp. 51, 98, 99.
- [19] Horowitz, L., "Finite Dimensional Approximate Models for Gyroscopic Noise," to appear.
- [20] Bensoussan, A., Filtrage Optimal des Systèmes Linéaires, Dunod, Paris, 1971.
- [21] Mitter, S.K. and Vinter, R.B., "Filtering for Linear Stochastic Hereditary Differential Systems", in Control Theory, Numerical Methods and Computer Systems Modelling, Int. Symposium, Rocquencourt, June 17-21, 1974 (Bensoussan, A., Lions, J.L., eds.), pp.1-21.
- [22] Hager, W.W., and Horowitz, L.L., "Convergence and Stability Properties of the Discrete Riccati Operator Equation and the Associated Optimal Control and Filtering Problems", SIAM J. Control and Optimization, Vol. 14, No. 2, Feb. 1976, pp. 295-312.
- [23] Vinter, R.B., "Stability of Filters for Infinite Dimensional Stochastic Systems, to appear.

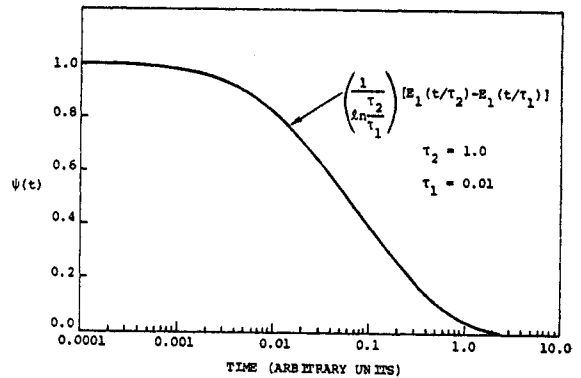
simulated random walk. The lowest frequency sample (aside from the sample at 0 corresponding to the mean) of the power spectral density is at $(f = 1/T)$, where T is the record length used for analysis. The highest frequency sample is at $(f = 1/2T_s)$, where T_s is the sample time.



3. The observed form of gyro noise power spectral density.

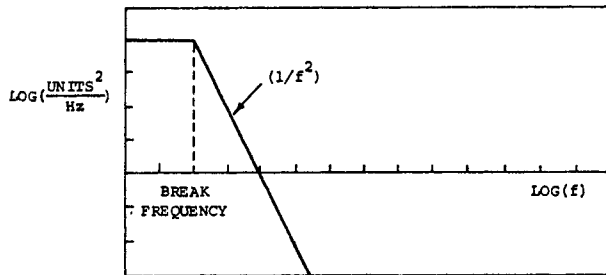


4. $\psi(t)$ vs. t ; linear-linear scale.

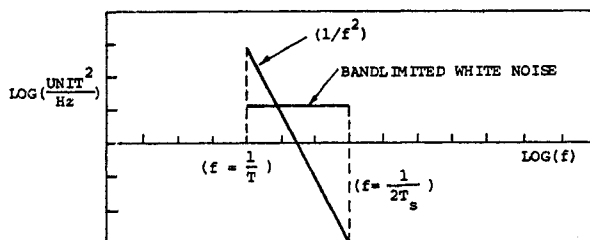


5. $\psi(t)$ vs. t ; semilog scale.

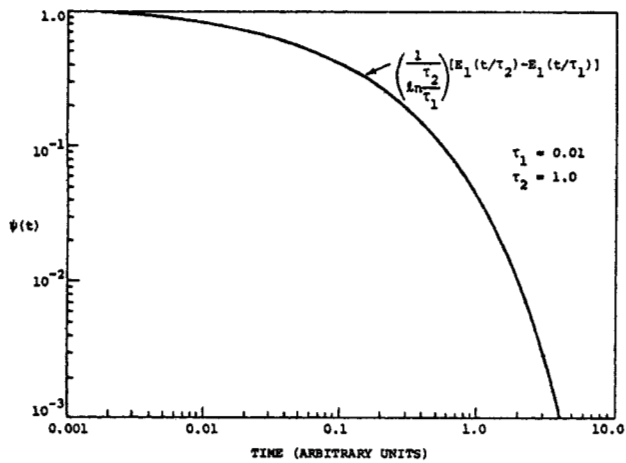
Figures



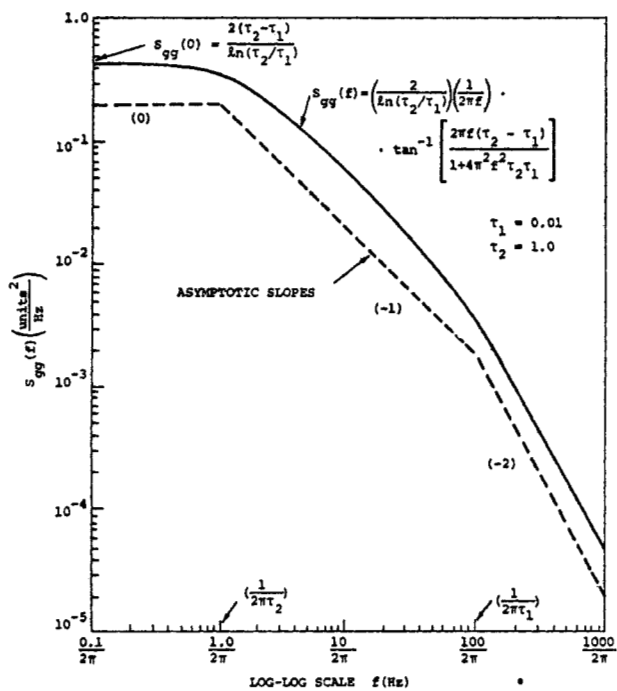
1. Linearized power spectral density of first-order Markov process



2. Linearized power spectral density of computer-



6. $\psi(t)$ vs. t ; log-log scale.



7. Power spectral density of gyroscopic noise model.