Nonstationary Velocity Estimation

By T. M. BURFORD

(Manuscript received January 3, 1958)

A nonstationary noise may frequently be approximated by the product of a stationary noise and a deterministic function of time. From observations of the sum of such a nonstationary noise and a linear signal, an estimate of the rate of change of the signal is found. More exactly, a random function, x(t), *is assumed to be one of the following:*

$$\mathrm{a} + \mathrm{bt} + \mathrm{g(t)n(t)},$$

$$\mathrm{a} + \mathrm{bt} + \mathrm{g(t)} \int_{-\infty}^{\mathrm{t}} \mathrm{h}(\mathrm{t} - \tau)\mathrm{n}(\tau)\, \mathrm{d}\tau, \quad \textit{or}$$

$$\mathrm{a} + \mathrm{bt} + \int_{-\infty}^{\mathrm{t}} \mathrm{h}(\mathrm{t} - \tau)\mathrm{g}(\tau)\mathrm{n}(\tau)\, \mathrm{d}\tau,$$

where a *and* b *are constants,* h(t) *is the impulse response of a lumped parameter filter,* n(t) *is white noise and* g(t) *is a nonzero deterministic function. A least squares estimate of* b *is found as a linear operation on a finite sample of* x(t).

I. WHITE NOISE CASE

A well-known estimation problem is that of forming a mean square estimate of the constant, b, in the random function $x(t) = a + bt + n(t)$, where $n(t)$ is white noise and a and b are unknown constants. The estimate of b is required to be the linear operation

$$\hat{b} = \int_{t-T}^{t} K(t - \tau)x(\tau)\, d\tau,$$

where $K(z)$ vanishes outside $(0, T)$ and $\hat{b}$ is to equal b in the absence of noise. The solution is known[1,2] to be

$$\begin{aligned} K(z) &= \frac{6}{T^3}(T - 2z); \quad 0 \leqq z \leqq T \\ &= 0; \qquad\qquad\qquad \text{elsewhere.} \end{aligned} \tag{1}$$

The above problem will be generalized here to include one type of nonstationary noise.

Let

$$x(t) = a + bt + g(t)n(t), \tag{2}$$

where a and b are unknown constants,* $n(t)$ is white noise of unit spectral density and $g(t)$ is a nonzero, deterministic function. Formally, it follows that

$$E[g(t)n(t)g(t')n(t')] = g(t)g(t')\delta(t - t'). \tag{3}$$

We wish an estimate of b in the form

$$\hat{b} = \int_{t-T}^{t} K(t, t - \tau)x(\tau)\, d\tau, \tag{4}$$

where $K(t, z)$ vanishes for z outside $(0, T)$. A constraint is that, in the absence of noise, (4) should give b exactly. Therefore

$$b = \int_{t-T}^{t} K(t, t - \tau)(a + b\tau)\, d\tau,$$

which implies, with a change of variable, that

$$\int_{0}^{T} K(t, z)\, dz = 0, \tag{5}$$

and

$$\int_{0}^{T} zK(t, z)\, dz = -1. \tag{6}$$

Using (3), (5) and (6) and again changing variables we find that the expected error is now

$$E(\hat{b} - b)^2 = \int_{0}^{T} K^2(t, z)g^2(t - z)\, dz. \tag{7}$$

The minimization of (7) is easily done by the usual variational technique, using (5) and (6) as isoperimetric constraints. The variation gives

$$2g^2(t - z)K(t, z) + \lambda + \mu z = 0.$$

Therefore

$$K(t, z) = \frac{-\lambda - \mu z}{2g^2(t - z)}, \tag{8}$$

* The linear $a + bt$ is used here for simplicity. The coefficients of a general polynomial may be estimated in a similar way.

where λ and μ are undetermined multipliers to be fixed by (5) and (6). From (5),

$$\int_0^T K(t, z)\, dz = -\frac{\lambda}{2}\int_0^T \frac{dz}{g^2(t - z)} - \frac{\mu}{2}\int_0^T \frac{z\, dz}{g^2(t - z)} = 0,$$

and, from (6),

$$\int_0^T zK(t, z)\, dz = -\frac{\lambda}{2}\int_0^T \frac{z\, dz}{g^2(t, z)} - \frac{\mu}{2}\int_0^T \frac{z^2\, dz}{g^2(t - z)} = -1.$$

Solving the preceding equations for λ and μ and substituting into (8) gives

$$\begin{aligned} K(t, z) &= \frac{A_1 - zA_0}{(A_0A_2 - A_1^2)g^2(t - z)}; \quad 0 \leqq z \leqq T \\ &= 0; \qquad\qquad \text{elsewhere}, \end{aligned} \tag{9}$$

with

$$A_j(t) = \int_0^T \frac{z^j\, dz}{g^2(t - z)},$$

which gives the minimized error

$$E(b - \hat{b})^2 = \frac{A_0}{A_0A_2 - A_1^2}.$$

The denominator of $K(t, z)$ does not vanish, since $g(t - z)$ is assumed nonzero, and

$$\int_0^T \frac{\left(z - \dfrac{A_1}{A_0}\right)^2}{g^2(t - z)}\, dz > 0,$$

$$A_2 - \frac{A_1^2}{A_0} > 0,$$

$$A_0A_2 - A_1^2 > 0.$$

In general, the $K(t, z)$ found here defines a time variable filter over a finite time interval. However, if g is a constant, (9) reduces to (1). Another interesting special case is that in which g is assumed to be of exponential form, which implies that

$$g(t - z) = g(t)g(-z),$$

and therefore

$$A_j = \frac{1}{g^2(t)} \int_0^T \frac{z^j\,dz}{g^2(-z)} = \frac{B_j}{g^2(t)}\,.$$

Substitution into (9) gives

$$\begin{aligned} K(z) &= \frac{B_1 - zB_0}{(B_0B_2 - B_1^2)g^2(-z)}\,; & 0 \leqq z \leqq T \\ &= 0; & \text{elsewhere.} \end{aligned}$$

Now, however, the B_j's are constants instead of functions of t, as the A_j's were. For an exponential g, therefore, the filter given by (9) is time-invariant.

The function $g(t)$ is not necessarily continuous. For example, let

$$\begin{aligned} g(x) &= g_1; & x < \beta \\ &= g_2\,; & x > \beta. \end{aligned}$$

Then

$$\begin{aligned} A_j &= \frac{1}{g_2^2}\frac{T^{j+1}}{j+1}\,; & \beta < t - T, \\ &= \frac{1}{g_1^2}\frac{T^{j+1}}{j+1} + \left(\frac{1}{g_2^2} - \frac{1}{g_1^2}\right)\frac{\beta^{j+1}}{j+1}\,; & t - T < \beta < t, \\ &= \frac{1}{g_1^2}\frac{T^{j+1}}{j+1}\,; & \beta > t, \end{aligned}$$

which implies that $K(t, z)$ is linear but has a discontinuity in value and slope at $z = \beta$.

Many applications involve a linear $g(t)$. For this case, it is possible to plot a one-parameter family of curves which completely describe $K(t, z)$. If $g(t)$ equals $\alpha + \beta t$, we define $Q = -(\alpha + \beta t)/\beta t$. Then, by a change of variable,

$$A_j = \frac{T^{j-1}}{\beta^2} \int_0^1 \frac{s^j\,ds}{(Q + s)^2} = \frac{T^{j-1}}{\beta^2}\,C_j\,,$$

and substitution into (9) gives

$$T^2K(t, z) = \frac{C_1 - C_0\left(\dfrac{z}{T}\right)}{(C_0C_2 - C_1^2)\left[Q + \left(\dfrac{z}{T}\right)\right]^2}\,.$$

All of the time dependence is now included in Q. Fig. 1 shows $T^2K(t, z)$ plotted against z/T for several values of Q. A positive value of Q implies that $| g(t) |$ is decreasing with time and therefore that recent data are superior. Values of Q in $(-1, 0)$ are not permitted, since such a Q would imply that $g(t)$ vanished somewhere in $(t - T, t)$. If $g(t)$ is constant, Q is infinite and (1) applies.

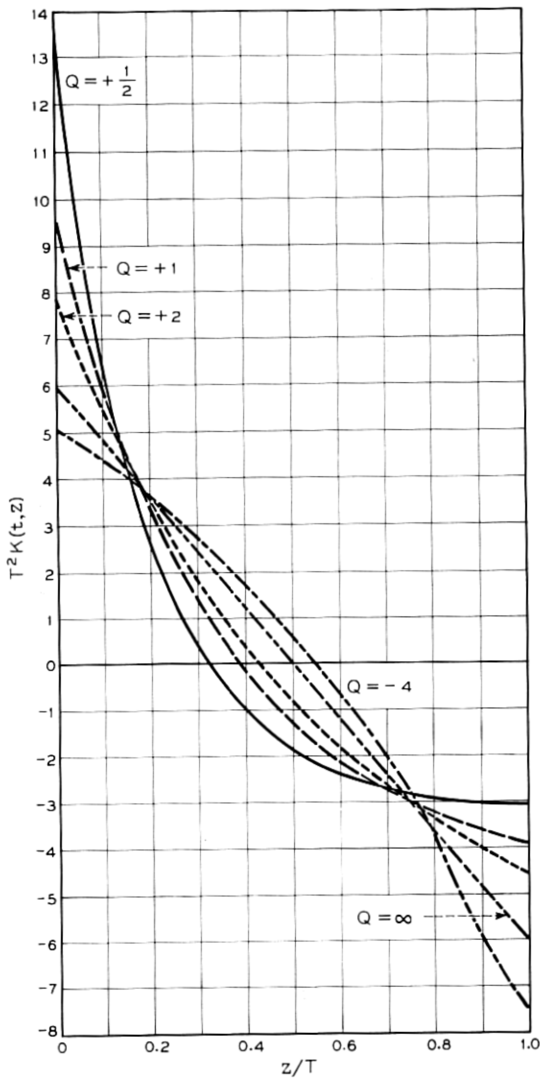

Fig. 1 — Weighting functions, $K(t, z)$, for linear $g(t)$.

The $K(t, z)$ just found for the linear $g(t)$ case may be substituted into (7) to give the variance of $\hat{b}$ as a function of Q:

$$\frac{T}{\beta^2} E(\hat{b} - b)^2 = \left[1 - \left(\log \frac{Q+1}{Q}\right)^2 Q(Q+1)\right]^{-1}. \tag{10}$$

An interesting comparison may be made with the variance resulting from using $K(z)$ of (1) in (7), which gives

$$\frac{T}{\beta^2} E(\hat{b} - b)^2 = 3(2Q + 1)^2 + \frac{9}{5}. \tag{11}$$

Equation (11) shows the error to be expected from using a $K(z)$, which is optimum in the stationary case, on this type of nonstationary data.

Equations (10) and (11) are plotted in Fig. 2. The difference between the curves is a measure of the improvement possible* with time-variable smoothing.

As Q approaches ∞, (10) may be written

$$\frac{T}{\beta^2} E(\hat{b} - b)^2 = 3(2Q + 1)^2 - \frac{7}{5} + 0\left(\frac{1}{Q}\right);$$

therefore (10) and (11) will asymptotically differ by 16/5.

Fig. 2 is plotted for $Q > 0$; however, it may be used for $Q < -1$ since both (10) and (11) are even on either side of $Q = -0.50$.

Fig. 1 also suggests a simple approximate realization for $K(t, z)$. The several curves have approximately common intersections near $z/T = 0.2$ and 0.8. Therefore, $K(t, z)$ might be represented as

$$K(t, z) = K_1(z) + f(Q)K_2(z),$$

where $K_1(z)$ is one of the $K(t, z)$ curves near the middle of the range of Q's to be considered. The function $K_2(z)$ vanishes near the points $z/T = 0.2$ and 0.8 and takes on relatively small values elsewhere. This approximate realization thus involves two time-invariant filters and a multiplier which depends on Q and thereby on t.

II. PRE-FILTERING OF NOISE

Filtering and multiplication by $g(t)$ are not, in general, commutative operations, so a choice must be made between

$$g(t) \int_{-\infty}^{t} h(t - \tau)n(\tau)\, d\tau \tag{12}$$

* It has been shown by E. N. Gilbert that, even if a linear $g(t)$ vanishes in the smoothing interval, the best possible estimate has variance β^2/T and, therefore, perfect estimation is not possible.

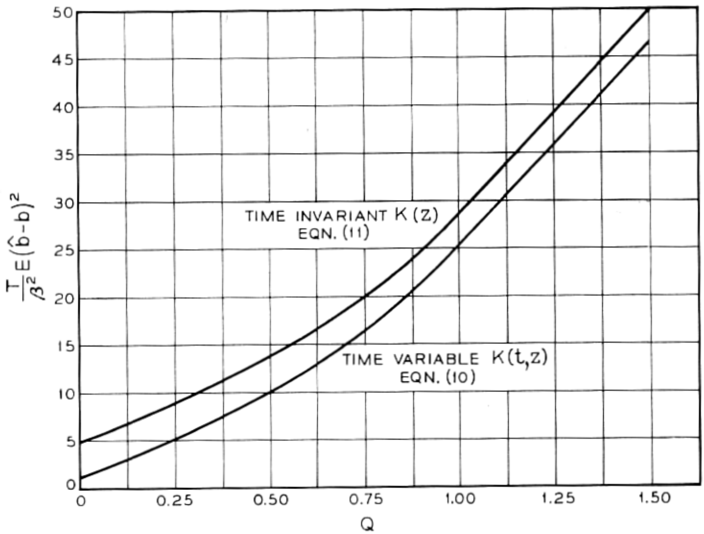

Fig. 2 — Variance for linear $g(t)$.

and

$$\int_{-\infty}^{t} h(t - \tau)g(\tau)n(\tau)\, d\tau \tag{13}$$

as noise functions to be added to $a + bt$. In certain cases, the representations (12) and (13) are virtually equivalent in the sense that, given $h(t - \tau)$, there exists an $\tilde{h}(t - \tau)$ such that

$$g(t) \int_{-\infty}^{t} h(t - \tau)n(\tau)\, d\tau = \int_{-\infty}^{t} \tilde{h}(t - \tau)g(\tau)n(\tau)\, d\tau.$$

From the point-wise uncorrelation of $n(\tau)$, the preceding equation, if it is true at all, must be true for all values of τ. Therefore

$$\tilde{h}(t - \tau) = \frac{g(t)}{g(\tau)}\, h(t - \tau),$$

which implies that $g(t)/g(\tau)$ must be a function of $(t - \tau)$. The only nonconstant $g(t)$ satisfying this condition is the exponential. If $g(t)$ is an exponential, the choice between (12) and (13) is arbitrary. In general, however, the choice is not arbitrary and depends on physical considerations. The forms (12) and (13) will now be treated as Cases 1 and 2.

Case 1

The observed quantity, $x(t)$, is

$$x(t) = a + bt + g(t) \int_{-\infty}^{t} h(t - \tau)n(\tau)\, d\tau,$$

where $n(\tau)$ is white noise of spectral density unity. The transform of $h(t)$ is assumed to be

$$T[h(t)] = \frac{\sum_{0}^{m} \beta_j (i\omega)^j}{\sum_{0}^{n} \alpha_j (i\omega)^j}; \quad m < n.$$

It is also assumed that the numerator and denominator of $T[h(t)]$ have neither common factors nor multiple roots. The covariance of

$$\int_{-\infty}^{t} h(t - \tau)n(\tau)\, d\tau$$

is defined to be $\rho(\tau)$. A particular $K(t, z)$ is sought such that

$$\hat{b} = \int_{t-T}^{t} K(t, t - \tau)x(\tau)\, d\tau$$

minimizes

$$E(\hat{b} - b)^2 = \int_0^T dz\, g(t - z)K(t, z) \int_0^T dz'\, g(t - z')K(t, z')\rho(z - z'). \quad (14)$$

In obtaining (14), it was assumed that, in the absence of noise, $\hat{b}$ equals b, or

$$\int_0^T K(t, z)\, dz = 0 \qquad \int_0^T zK(t, z)\, dz = -1. \quad (15)$$

Using the constraints of (15), the variation of (14) gives the integral equation*

$$\int_0^T dz' K(t, z')g(t - z')\rho(z - z') = \frac{\lambda + \mu z}{g(t - z)}. \quad (16)$$

From the form of $T[h(t)]$, it is evident that

$$\rho(z - z') = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{M(\omega^2)}{N(\omega^2)} \epsilon^{i\omega(z-z')}\, d\omega,$$

* A similar equation, of somewhat different origin, has been considered by Ule.[3]

where $M(\omega^2)$ and $N(\omega^2)$ are polynomials in ω^2 of degrees m and n. Now form[4] the operation $N(-d^2/dz^2)$ and apply it to (16), giving

$$\frac{1}{2\pi}\int_0^T dz' K(t, z')g(t - z') \int_{-\infty}^{\infty} M(\omega^2)\epsilon^{i\omega(z-z')}\, d\omega = N\left(-\frac{d^2}{dz^2}\right)\frac{\lambda + \mu z}{g(t - z)},$$

which is, formally, the same as

$$\frac{1}{2\pi} M\left(-\frac{d^2}{dz^2}\right)\int_0^T dz' K(t, z')g(t - z') \int_{-\infty}^{\infty} \epsilon^{i\omega(z-z')}\, d\omega = N\left(-\frac{d^2}{dz^2}\right)\frac{\lambda + \mu z}{g(t - z)},$$

or, the differential equation

$$M\left(-\frac{d^2}{dz^2}\right)K(t, z)g(t - z) = N\left(-\frac{d^2}{dz^2}\right)\frac{\lambda + \mu z}{g(t - z)}. \tag{17}$$

The general bounded solution of (17) will not necessarily satisfy (16), however. As is well known,[2] singularity functions up to order $n - m - 1$ should be added at the end points of the interval in order to satisfy the boundary conditions of the integral equation (16).

Therefore, let

$$K(t, z)g(t - z) = \sum_0^{2m-1} a_j\epsilon^{m_j z} + \int_0^z k(z - x)N\left(-\frac{d^2}{dx^2}\right)\left(\frac{\lambda + \mu x}{g(t - x)}\right)dx + \sum_0^{n-m-1} b_j\delta^{(j)}(z) + \sum_0^{n-m-1} c_j\delta^{(j)}(z - T), \tag{18}$$

where $\delta^{(0)}$ is the Dirac delta function, $\delta^{(1)}$ its derivative, etc. The function $k(z - x)$ is the Green's function associated with $M(-(d^2/dz^2))$ and the $2m$ constants m_j are the roots of $M(-X^2)$. The a_j, b_j and c_j comprise $2n$ undetermined constants. They may be determined by substituting (18) into the integral equation (16) and requiring an identity in the $2n$ exponentials of ρ. For example, assume $h(x) = \epsilon^{-ax}$, which implies that

$$\frac{M(\omega^2)}{N(\omega^2)} = \frac{1}{a^2 + \omega^2}.$$

Then

$$K(t, z)g(t - z) = \left(a^2 - \frac{d^2}{dz^2}\right)\frac{\lambda + \mu z}{g(t - z)} + b_0\delta(z) + c_0\delta(z - T),$$

and substitution into (16) gives the following equations for b_0 and c_0:

$$b_0 = \left(a - \frac{d}{dz}\right)\frac{\lambda + \mu z}{g(t - z)}\bigg|_{z = 0} \qquad c_0 = \left(a + \frac{d}{dz}\right)\frac{\lambda + \mu z}{g(t - z)}\bigg|_{z = T}.$$

Except for evaluating λ and μ from (15), the solution is

$$K(t, z) = \frac{1}{g(t-z)}\left[a^2 - \frac{d^2}{dz^2} + \delta(z)\left(a - \frac{d}{dz}\right) + \delta(z - T)\left(a + \frac{d}{dz}\right)\right]\frac{\lambda + \mu z}{g(t-z)}.$$

Case 2

The observed quantity $x(t)$ now is

$$x(t) = a + bt + \int_{-\infty}^{t} h(t-\tau)g(\tau)n(\tau)\,d\tau.$$

We define $\hat{b}$, $K(t, z)$ and $h(t - \tau)$ as in Case 1 and again require that, in the absence of noise, $\hat{b}$ equal b, or that the constraints (15) hold. These assumptions lead to the following integral equation:

$$\int_0^T dz' K(t, z')\Phi(t - z, t - z') = \lambda + \mu z,$$

where

$$\Phi(t - z, t - z') = \int_0^\infty h(y)h(y + z' - z)g^2(t - z' - y)\,dy.$$

Instead of dealing directly with the integral equation, we first rewrite $x(t)$ as

$$x(t) = \int_{-\infty}^{t} h(t-\tau)[\beta + \gamma\tau + g(\tau)n(\tau)]\,d\tau, \tag{19}$$

where

$$\gamma = \frac{\alpha_0 b}{\beta_0},$$

and

$$\beta = \frac{\alpha_0 a}{\beta_0} + \frac{b}{\beta_0}\left(\alpha_1 - \frac{\beta_1\alpha_0}{\beta_0}\right),$$

since it is easily shown that

$$\int_0^\infty h(x)\,dx = \frac{\beta_0}{\alpha_0},$$

and

$$\int_0^\infty xh(x)\,dx = \frac{\beta_0\alpha_1}{\alpha_0^2} - \frac{\beta_1}{\alpha_0},$$

from the assumed form of $T[h(t)]$, and that

$$\int_{-\infty}^{t} h(t - \tau)(\beta + \gamma\tau)\, d\tau =$$

$$\beta \int_0^\infty h(x)\, dx + \gamma t \int_0^\infty h(x)\, dx - \gamma \int_0^\infty xh(x)\, dx = a + bt$$

when the values of β and γ are substituted from (19).

A function $y(t)$ is now defined as

$$y(t) = \beta + \gamma t + g(t)n(t)$$

and a function $\tilde{K}(t, z)$ is defined as the convolution of $K(t, z)$ and $h(z)$, so that (14) is replaced by

$$\hat{b} = \int_{-\infty}^{t} \tilde{K}(t, t - \tau) y(\tau)\, d\tau.$$

The function $y(t)$ is similar to $x(t)$ as defined in (2). Therefore, $\tilde{K}(t, z)$ may be found in a manner analogous to that used in the case in which the noise was not pre-filtered. Recalling that $K(t, z)$ is required to vanish outside $(0, T)$, we may define $\tilde{K}(t, z)$ more exactly by

$$\begin{aligned} \tilde{K}(t, z) &= \int_0^z dv K(t, v) h(z - v); \quad 0 < z < T \\ &= \int_0^T dv K(t, v) h(z - v); \quad z > T \\ &= 0; \quad z < 0, \end{aligned} \tag{20}$$

or the equivalent:

$$\begin{aligned} \sum_0^n \alpha_j \frac{d^j}{dz^j} \tilde{K}(t, z) = \sum_0^m \beta_j \frac{d^j}{dz^j} K(t, z); \quad 0 < z < T \\ = 0; \quad z > T, \quad z < 0. \end{aligned} \tag{21}$$

Since $\tilde{K}(t, z)$ may be discontinuous at $z = 0$ and T, these points have been excluded. While $K(t, z)$ is constrained by (15), it follows that $\tilde{K}(t, z)$ must satisfy

$$\begin{aligned} \int_0^\infty \tilde{K}(t, z)\, dz &= 0 \\ \int_0^\infty z\tilde{K}(t, z)\, dz &= -\frac{\beta_0}{\alpha_0}. \end{aligned} \tag{22}$$

The variance to be minimized is now

$$E(b - \hat{b})^2 = \int_0^\infty \tilde{K}^2(t, z)g^2(t - z)\, dz.$$

From (21),

$$\tilde{K}(t, z) = \sum_0^{n-1} s_j \epsilon^{\zeta_j z}; \quad z > T, \tag{23}$$

where the s_j are presently unknown but will be selected to minimize $E(\hat{b} - b)^2$ and the ζ_j are known in terms of the α_j. A variation of

$$\int_0^\infty [\tilde{K}^2(t, z)g^2(t - z) + \lambda \tilde{K}(t, z) + \mu z \tilde{K}(t, z)]\, dz$$

gives

$$2\tilde{K}(t, z)g^2(t - z) + \lambda + \mu z = 0; \quad 0 < z < T \tag{24}$$

and a set of n equations defining the s_j:

$$2 \sum_{j=0}^{n-1} s_j \int_T^\infty \epsilon^{(\zeta_j + \zeta_k)z} g^2(t - z)\, dz + \int_T^\infty (\lambda + \mu z)\epsilon^{\zeta} k^z\, dz = 0, \tag{25}$$

where

$$k = 0, 1, \cdots, n - 1.$$

Equations (23) and (24), together with the constraints (22), define $\tilde{K}(t, z)$ completely. The function $K(t, z)$ may now be found from (21).

From the discontinuities in $\tilde{K}(t, z)$ at $z = 0$ and T, delta functions of order as high as $n - m - 1$ may be expected in $K(t, z)$ at $z = 0$ and T. For example, if $h(x)$ equals ϵ^{-ax}, (21) becomes

$$\begin{aligned} \frac{d\tilde{K}(t, z)}{dz} + a\tilde{K}(t, z) &= K(t, z); \quad 0 < z < T \\ &= 0; \quad z < 0, \qquad z > T \end{aligned}$$

and

$$\tilde{K}(t, z) = s\epsilon^{-az}; \quad z > T.$$

From (24),

$$\tilde{K}(t, z) = \frac{-\lambda - \mu z}{2g^2(t - z)}; \quad 0 < z < T$$

and, from (25),

$$2s \int_T^\infty \epsilon^{-2az} g^2(t - z)\, dz + \int_T^\infty (\lambda + \mu z)\epsilon^{-az}\, dz = 0,$$

which defines s and therefore $\tilde{K}(t, z)$ for $z > T$. Except for evaluating λ and μ from (15), the solution is

$$K(t, z) = \left[\delta(z) - \delta(z - T) + a + \frac{d}{dz}\right]\left(\frac{\lambda + \mu z}{g^2(t - z)}\right) + \frac{\lambda + \mu\left(T + \frac{1}{a}\right)}{aH(t, T)}\,\delta(z - T),$$

where

$$H(t, T) = \epsilon^{2aT} \int_T^\infty \epsilon^{-2az} g^2(t - z)\, dz.$$

III. ACKNOWLEDGMENTS

The author appreciates the helpful suggestions of S. Darlington, E. N. Gilbert, and D. Slepian.

REFERENCES

1. Blackman, R. B., Bode, H. W. and Shannon, C. E., *Data Smoothing and Prediction in Fire-Control Systems*, NDRC Report MGC 12/1, 1948 (Confidential).
2. Zadeh, L. A. and Ragazzini, J. R., An Extension of Wiener's Theory of Prediction, J. Appl. Phys., **21,** July 1950, p. 645.
3. Ule, L. A., A Theory of Weighted Smoothing, Trans. I.R.E., **IT-3**, June 1957, p. 131.
4. Youla, D. C., The Use of the Method of Maximum Likelihood in Estimating Continuous-Modulated Intelligence Which Has Been Corrupted by Noise, Trans. I.R.E., **PGIT-3,** March 1954, p. 90.