The Annals of Mathematical Statistics
1972, Vol. 43, No. 6, 1935-1943

ERROR FORMULAE FOR OPTIMAL LINEAR FILTERING, PREDICTION AND INTERPOLATION OF STATIONARY TIME SERIES

BY JAKOV SNYDERS[1]

Technion—Israel Institute of Technology

Several explicit expressions are presented for the minimum mean square error in linear causal filtering, prediction and interpolation of weakly stationary discrete-time processes corrupted by additive noise. A general procedure for deriving error expressions of this kind is established.

1. Introduction. Several explicit formulae expressing the minimum mean square error of linear filtering and prediction of wide-sense stationary processes are known. The basic and earliest of these results is due to Szegö-Kolmogoroff-Krein ([1]; [2], page 49); it states that

$$\lim_{n\to\infty} \min_{a_1,\cdots,a_n} \frac{1}{2\pi} \int |1 + a_1 e^{i\theta} + \cdots + a_n e^{in\theta}|^2 \, d\mu(\theta) = \exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(e^{i\theta}) \, d\theta\right]$$

where μ is a finite positive Baire measure on C, the unit circle of a complex plane, and f is the derivative of the absolutely continuous part of μ with respect to Lebesgue measure (the right-hand side is interpreted as zero if $\log f$ is not integrable). Equivalently,

$$\min_{h\in H_+^2(d\mu)} \frac{1}{2\pi} \int |h(e^{i\theta}) - e^{-i\theta}|^2 \, d\mu(\theta) = \exp\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(e^{i\theta}) \, d\theta\right] \tag{1}$$

where $H_+^2(d\mu)$ is the (Hardy) subspace of $L^2(d\mu)$ spanned by $\{e^{in\theta}; n = 0, 1, 2, \cdots\}$. If μ stands for the spectral measure of a stationary discrete-time process, then (1) expresses the minimum mean-square error of linear prediction one time-unit ahead. Our concern in this paper is to derive formulae for cases when the process is (possibly) disturbed by noise and when other types of operations, such as prediction several units of time ahead, is desired. This is accomplished by following simple procedures developed in the sequel. However, only absolutely continuous spectral densities will be considered.

2. Statement of the problem. A stochastic sequence $\{q_t\}$ is said to be derived by a linear operation on a wide-sense stationary sequence $\{P_t\}$ if, for each t, q_t is either of the form $\sum_{k=0}^{K} a_k P_{t+t(k)}$ or is the quadratic mean limit of such finite sums. A linear operation is called causal if only non-positive $t(k)$ appear in its defining sums of the above mentioned form.

Let $\{m_t\}$ and $\{n_t\}$ be wide-sense stationary sequences, considered to be the

Received October 26, 1971; revised April 7, 1972.

[1] At present with the Electrical Engineering Dept., University of Toronto.

message and the noise, respectively. Assume that these processes are uncorrelated with each other and that they have spectral densities denoted, respectively, by f and g. Further assume that the mean of $\{n_t\}$ is zero. Let h be a function on C which represents a linear causal operation on $\{m_t + n_t\}$ and which belongs, therefore, to $H_+^2(f\,d\theta + g\,d\theta)$; and let $r \in L^2(f\,d\theta)$ represent a linear "desired" operation on $\{m_t\}$. The mean-square difference between the "desired" and the best obtainable sequences, shortly termed minimum (mean-square filtering) error and denoted by E, is given by

$$(2) \qquad E = \min_{h \in H_+^2(f d\theta + g d\theta)} \frac{1}{2\pi} \int_{-\pi}^{\pi} [|h(e^{i\theta}) - r(e^{i\theta})|^2 f(e^{i\theta}) + |h(e^{i\theta})|^2 g(e^{i\theta})]\, d\theta \,.$$

It is easily verifiable that the minimum in (2) is indeed attained by some $h_0 \in H_+^2(f\,d\theta + g\,d\theta)$, which is uniquely defined a.e. with respect to $(f + g)\theta$, and is called the optimal transfer function.

Some further properties, especially the rationality of g and r, are required for the manipulations performed in the sequel. Stated more precisely, we assume the following:

A_1: f, g and r are integrable functions on C;

A_2: f and g are nonnegative on C;

A_3: g and r are rational (thus defined over the whole plane);

A_4: at least one of the conditions $g \neq 0$, $\log f \in L$ holds;

A_5: the structures of g and r are known, i.e. the number and multiplicity of zeros and poles of each function are known;

A_6: in case $g = 0$, the number and multiplicity of the poles of r located inside the unit disc are known;

and our purpose is to express E, defined by (2), in terms of f, g and coefficients of g and r. Observe that f is *not* necessarily *rational*. A_4 is assumed since if both $g = 0$ and $\log f \notin L$ then $E = 0$ (this is obvious for $r \in H_+^2(f\,d\theta)$, for $r(e^{i\theta}) = e^{-i\theta}$ it follows from (1), which in turn implies the minimum's vanishing for all $r \in L^2(f\,d\theta)$). The information included in A_5 and A_6 is important: in general different structures of g and r induce different formulae.

Some not too complicated error expressions, arising from simple structures of g and r, are given in the next section. They include examples of filtering, prediction and interpolation in white noise, colored noise and without noise. These expressions were derived by a method which is summarized in Propositions 1 and 2 of the last section. The Poisson transform extension (to the unit disc) of the optimal transfer function corresponding to each case is also obtainable by these propositions. A detailed proof of one error formula is presented at the end of Section 4. It helps to clarify the routine but sometimes tedious nature of derivation.

Previous work includes a matricial extension ([8]) of (1), and several expressions derived by using Toeplitz forms ([6], [9]). Choosing the imaginary axis version of (2) as starting point, Yovits and Jackson [11] arrived at an explicit error expression for continuous-time filtering in white noise. This and other results related to continuous-time were usually derived under more restrictive conditions involving either rationality of the signal spectral density ([3], [4], [7]) or minimum-phase property of the optimal transfer function in [10]. In [5] the particular case $r = 1$ of (2) and its imaginary axis version were considered. The method applied here is similar to that of [5]. However, the results obtained, in particular Propositions 1 and 2, are not an immediate generalization of the results in [5]. In a recent report ([12]) it was demonstrated that four error formulae, among them the fourth assertion of Theorem 1 and the second assertion of Theorem 2, stand also for the non-absolutely continuous case in the same sense as (1) does, i.e. the minimal error is unaffected by the singular part of the signal spectral measure. This property of the minimal error is possibly quite general.

3. Some error formulae. Throughout this section we assume that A_1—A_4 hold and E is defined by (2).

THEOREM 1. *Let g be a nonzero constant and $r(e^{i\theta}) = e^{-in\theta}$ with n specified below. Then*

$$\begin{aligned}
n &= -3: && E = g\{1 - [1 + |a_1|^2 + |a_2 - \tfrac{1}{2}a_1{}^2|^2 + |a_3 - a_1a_2 + \tfrac{1}{6}a_1{}^3|^2]P^{-2}(0)\}\,,\\
n &= -2: && E = g\{1 - [1 + |a_1|^2 + |a_2 - \tfrac{1}{2}a_1{}^2|^2]P^{-2}(0)\}\,,\\
n &= -1: && E = g\{1 - [1 + |a_1|^2]P^{-2}(0)\}\,,\\
n &= 0: && E = g\{1 - P^{-2}(0)\}\,,\\
n &= 1: && E = g\{P^2(0) - 1\}\,,\\
n &= 2: && E = g\{[1 + |a_1|^2]P^2(0) - 1\}\,,\\
n &= 3: && E = g\{[1 + |a_1|^2 + |a_2 + \tfrac{1}{2}a_1{}^2|^2]P^2(0) - 1\}\,,\\
n &= 4: && E = g\{[1 + |a_1|^2 + |a_2 + \tfrac{1}{2}a_1{}^2|^2 + |a_3 + a_1a_2 + \tfrac{1}{6}a_1{}^3|^2]P^2(0) - 1\}
\end{aligned}$$

where

$$P^2(0) = \exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi} \log\left(1 + \frac{f(e^{i\theta})}{g}\right) d\theta\right].$$

$$a_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{ij\theta} \log\,[g + f(e^{i\theta})]\, d\theta\,, \qquad j = 1, 2, 3\,.$$

These results were obtained by direct application of Proposition 1. Corresponding results for $g = 0$ are derivable by Proposition 2 or, more simply, by observing that $P^{-2}(0) \to 0$, $gP^2(0) \to \exp\,[(2\pi)^{-1}\int_{-\pi}^{\pi} \log f(e^{i\theta})\, d\theta]$ and $a_j \to (2\pi)^{-1}\int_{-\pi}^{\pi} e^{ij\theta} \log f(e^{i\theta})\, d\theta$ as $g \to 0$. For $n \leqq 0$, $E = 0$ and for $n = 1$ we obtain (the absolutely continuous case of) (1), as expected. The manipulations leading to the first assertion of Theorem 1 are presented in the next section.

THEOREM 2. *Let $g(e^{i\theta}) = |a|^2/|e^{i\theta} - b|^2$ where $|b| > 1$ and $r(e^{i\theta}) = e^{-in\theta}$ with n*

specified below. Then

$$n=-1:\quad E=\frac{|a|^2}{|b|^2-1}[1-|P(b)|^2]-|a|^2|P^{-1}(0)-P(b)|^2\,,$$

$$n=0:\quad E=\frac{|a|^2}{|b|^2-1}[1-|P(b)|^2]\,,$$

$$n=1:\quad E=\frac{|a|^2}{|b|^2-1}[1-|P(b)|^2]+\left|\frac{a}{b}\right|^2|P(0)-P(b)|^2$$

where

$$P(b)=\exp\left[\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{e^{i\theta}+b}{e^{i\theta}-b}\log\left(1+\frac{f(e^{i\theta})}{g(e^{i\theta})}\right)d\theta\right],$$

$$P(0)=\exp\left[\frac{1}{4\pi}\int_{-\pi}^{\pi}\log\left(1+\frac{f(e^{i\theta})}{g(e^{i\theta})}\right)d\theta\right].$$

THEOREM 3. *Let* $r(e^{i\theta})=a/(e^{i\theta}-b)$ *where* $|b|\neq 1$ *and let* g *be constant. If* $g\neq 0$ *then*

$$E=\frac{|a|^2g}{|b|^2-1}\left\{1-\exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1-|b|^2}{|e^{i\theta}-b|^2}\log\left(1+\frac{f(e^{i\theta})}{g}\right)d\theta\right]\right\},$$

whereas if $g=0$ *then*

$$E=\frac{|a|^2}{1-|b|^2}\exp\left[\frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1-|b|^2}{|e^{i\theta}-b|^2}\log f(e^{i\theta})\,d\theta\right],\qquad |b|<1\,,$$
$$=0\,,\qquad |b|>1.$$

4. The derivation of error formulae. $H_+{}^2$ will be shorthand for $H_+{}^2(d\mu)$ when μ is the Lebesgue measure; $H_-{}^2$ denotes the subspace of L^2 spanned by $\{e^{-in\theta};\ n=0,1,2,\cdots\}$. For a nonnegative measurable function on C satisfying $\log f\in L$ we define the outer functions ([2], page 62) f^+, f^- by

$$f^+(z)=\exp\left[\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{e^{i\theta}+z}{e^{i\theta}-z}\log f(e^{i\theta})\,d\theta\right],\qquad |z|<1\,,$$

$$f^-(z)=\exp\left[-\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{e^{i\theta}+z}{e^{i\theta}-z}\log f(e^{i\theta})\,d\theta\right],\qquad |z|>1\,.$$

On C, f^+ and f^- are defined, respectively, as the nontangential limits of the above expressions. Obviously $f(e^{i\theta})=f^+(e^{i\theta})f^-(e^{i\theta})$, and it is well known ([2], page 53) that $f\in L$ implies $f^+\in H_+{}^2$, $f^-\in H_-{}^2$. For a rational f with known structure this factorization is performable by inspection and produces rational functions.

The next lemma and its proof were communicated to the author by M. Zakai. Since it is apparently unpublished except for a similar statement in [5], we shall give also the proof. Dash above a function indicates complex conjugation.

LEMMA 1 (Zakai). *Let* k *be a nonnegative integrable function on* C *satisfying* $\log k\in L$. *Then* $H_+{}^2(k\,d\theta)=\{h:\ hk^+\in H_+{}^2\}$.

PROOF. Denote $W = \{h: hk^+ \in H_+^2\}$ and equip this set with the $L^2(k\,d\theta)$ norm. By writing $k = |k^+|^2$ it is easily seen that $H_+^2(k\,d\theta) \subset W$. Assuming therefore, that $H_+^2(k\,d\theta)$ is a proper subspace of W, there exists $h \in W$ such that

$$\int_{-\pi}^{\pi} \overline{h(e^{i\theta})}k(e^{i\theta})e^{in\theta}\,d\theta = 0\,, \qquad n \geqq 0\,. \tag{3}$$

Thus $\bar{h}k \in H_+^1$, which implies, using the inner-outer functions factorization ([2], page 62) that $\bar{h}k^- \in H_+^2$. But obviously $\bar{h}k^- \in H_-^2$, hence $\bar{h}k^-$ is a constant. Considering (3) with $n = 0$ and since $\int_{-\pi}^{\pi} k^+(e^{i\theta})\,d\theta \neq 0$, this constant is zero. Therefore $h = 0$ a.e. with respect to $k\,d\theta$.

Notation of the kind $\sum \operatorname{Res}[p(s, z); q(s), s = \infty]$ means the sum of the residues of $p(s, z)$ over the poles of $q(s)$ and at $s = \infty$. If the summation is only over those poles of $q(s)$ which are located outside the closed unit disc or inside the open unit disc, then $q(s)$ will be replaced by $q_+(s)$ or $q_-(s)$, respectively. It should be emphasized that $q_+(s)$ *differs* from $q^+(s)$ since q_+ is introduced above as a *notation* related to any rational function q, while q^+ is a *function* which is defined only if q is *nonnegative* on C. Nevertheless, if q is rational without poles on C and if q^+ exists then obviously $\sum \operatorname{Res}[p(s, z); q^+(s)] = \sum \operatorname{Res}[p(s, z); q_+(s)]$. Similar remarks apply to the difference between $q_-(s)$ and $q^-(s)$ (it is customary to associate the notations q_+ and q_- with a decomposition of the form $q = q_+ + q_-$, but we do not need these functions).

LEMMA 2. *Let $f \in H_+^2$ and $g \in H_-^2$. If g is rational then*

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{i\theta})f(e^{i\theta})e^{i\theta}\,d\theta = \sum \operatorname{Res}[g(z)f(z); g(z)]\,. \tag{4}$$

Similarly, if f is rational then

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} g(e^{i\theta})f(e^{i\theta})e^{i\theta}\,d\theta = -\sum \operatorname{Res}[g(z)f(z); f(z), z = \infty]\,.$$

PROOF. Since the two cases are provable by the same method, we shall consider the first one only. Choose $\rho < 1$ such that the disc $|z| < \rho$ contains all the poles of g, and define $f_\rho(e^{i\theta}) = f(\rho e^{i\theta})$ (f inside the unit disc is obtained by Poisson transform, as usual). An equation like (4) certainly holds for f replaced by f_ρ, and this implies (4) since $f_\rho \to f$ as $\rho \to 1$ both pointwise a.e. and in L^2 [2].

PROPOSITION 1. *Assume* A_1—A_3 *and* $g \neq 0$, *and let E be defined by* (2). *Then*

$$E = \sum \operatorname{Res}[T(z)P^{-1}(z)g^-(z); T(z), g^-(z), z = \infty]$$

where

$$T(z) = \frac{\overline{r(\bar{z}^{-1})}}{z} \sum \operatorname{Res}\left[\frac{r(s)P(s)g^+(s)}{s - z}; r(s), g^+(s), s = \infty\right] \tag{5}$$

and

$$P(z) = \exp\left[\frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \log\left(1 + \frac{f(e^{i\theta})}{g(e^{i\theta})}\right) d\theta\right], \qquad |z| \neq 1\,.$$

Furthermore, the Poisson transform extension of the optimal transfer function h_0 (to the unit disc) is given by

$$h_0(z) = r(z) + \frac{zT(z)P^{-1}(z)}{r(\bar{z}^{-1})g^+(z)}, \qquad |z| < 1 . \tag{6}$$

PROOF. First note that by the assumptions, $\log g$ and, consequently, $\log(f+g)$ are integrable. Set $f + g = k$. Perturbation of the equation

$$E = \frac{1}{2\pi} \int_{-\pi}^{\pi} [|h_0 - r|^2 f + |h_0|^2 g]\, d\theta \tag{7}$$

yields

$$\int_{-\pi}^{\pi} \overline{\delta h}[(h_0 - r)f + h_0 g]\, d\theta = 0 \tag{8}$$

or

$$\int_{-\pi}^{\pi} \overline{\delta h}[(h_0 - r)k + rg]\, d\theta = 0 \tag{9}$$

for every $\delta h \in H_+^2(k\, d\theta)$. By Lemma 1 we may choose $\delta h(e^{i\theta}) = [(1 - \bar{z}e^{i\theta})k^+(e^{i\theta})]^{-1}$ with $|z| < 1$ and obtain

$$\int_{-\pi}^{\pi} \left[\frac{h_0(e^{i\theta})k^+(e^{i\theta})}{e^{i\theta} - z} - \frac{r(e^{i\theta})k^+(e^{i\theta})}{e^{i\theta} - z} + \frac{r(e^{i\theta})g(^{i\theta})}{(e^{i\theta} - z)k^-(e^{i\theta})}\right] e^{i\theta}\, d\theta, \qquad |z| < 1 .$$

Each term in the brackets is factorable into a kind of product appearing in Lemma 2, since $h_0 k^+ \in H_+^2$ (by Lemma 1) and $|g^-(z)/k^-(z)| \leqq 1$ for $|z| \geqq 1$, therefore,

$$\begin{aligned} h_0(z)k^+(z) - r(z)k^+(z) - \sum \operatorname{Res}\left[\frac{r(s)k^+(s)}{s - z}; r_-(s)\right] \\ = \sum \operatorname{Res}\left[\frac{r(s)g(s)}{(s - z)k^-(s)}; r_+(s), g^+(s), s = \infty\right]. \end{aligned} \tag{10}$$

Observing that

$$\begin{aligned} \frac{k^+(s)}{g^+(s)} &= P(s) \qquad \text{if } |s| < 1 , \\ \frac{g^-(s)}{k^-(s)} &= P(s) \qquad \text{if } |s| > 1 , \end{aligned} \tag{11}$$

it is possible to rewrite (10) as follows:

$$[h_0(z) - r(z)]k^+(z) = \sum \operatorname{Res}\left[\frac{r(s)g^+(s)P(s)}{s - z}; r(s), g^+(s), s = \infty\right].$$

Define T by (5), then (6) obviously holds and

$$T(e^{i\theta}) = e^{-i\theta}[h_0(e^{i\theta}) - r(e^{i\theta})]k^+(e^{i\theta})\overline{r(e^{i\theta})} .$$

Setting $\delta h = h_0$ in (8) and comparing with (7)

$$\begin{aligned} E &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{r(e^{i\theta})} - [r(e^{i\theta}) - h_0(e^{i\theta})]f(e^{i\theta})\, d\theta \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{r(e^{i\theta})}[r(e^{i\theta}) - h_0(e^{i\theta})]k^+(e^{i\theta})\left[k^-(e^{i\theta}) - \frac{g(e^{i\theta})}{k^+(e^{i\theta})}\right] d\theta , \end{aligned} \tag{12}$$

thus

$$E = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{T(e^{i\theta})g(e^{i\theta})}{k^+(e^{i\theta})} e^{i\theta}\, d\theta - \frac{1}{2\pi} \int_{-\pi}^{\pi} T(e^{i\theta})k^-(e^{i\theta})e^{i\theta}\, d\theta .$$

Lemma 2 is again applicable since T is rational without poles on C, $|g^+(z)/k^+(z)| \leqq 1$ for $|z| \leqq 1$ and $k^- \in H_-^2$, hence

$$E = \sum \operatorname{Res}\left[\frac{T(z)g(z)}{k^+(z)};\, T_-(z),\, g^-(z)\right] + \sum \operatorname{Res}\,[T(z)k^-(z);\, T_+(z),\, z = \infty]$$

and by (11)

$$E = \sum \operatorname{Res}\,[T(z)P^{-1}(z)g^-(z);\, T(z),\, g^-(z),\, z = \infty] .$$

Note. It is possible to replace $\overline{r(\bar z^{-1})}$ by $r(z^{-1})$ in (5), (6) and in the next proposition, if and only if the coefficients of r are real. This is the natural case as long as the stationary sequences involved (the message and the noise) are real.

PROPOSITION 2. *Let E be defined by* (2) *with* $g = 0$, *and assume* A_1—A_4. *Then*

$$E = \sum \operatorname{Res}\,[T(z)Q^{-1}(z);\, T_+(z),\, z = \infty]$$

and

$$h_0(z) = r(z) + \frac{zT(z)Q^{-1}(z)}{\overline{r(\bar z^{-1})}}\,, \qquad |z| < 1$$

where

$$T(z) = \frac{\overline{r(\bar z^{-1})}}{z} \sum \operatorname{Res}\left[\frac{r(s)Q(s)}{s - z};\, r_-(s)\right],$$

$$Q(z) = \exp\left[\frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} \log f(e^{i\theta})\, d\theta\right], \qquad z \neq 1 .$$

PROOF. It is sufficient to note that (10) and (12) reduce, respectively, to

$$[h_0(z) - r(z)]f^+(z) = \sum \operatorname{Res}\left[\frac{r(s)f^+(s)}{s - z};\, r_-(s)\right],$$

$$E = \frac{1}{2\pi} \int_{-\pi}^{\pi} \overline{r(e^{i\theta})}[r(e^{i\theta}) - h_0(e^{i\theta})]f^+(e^{i\theta})f^-(e^{i\theta})\, d\theta$$

$$= \sum \operatorname{Res}\,[T(z)f^-(z);\, T_+(z),\, z = \infty]$$

and $f^+(z) = Q(z)$ for $|z| < 1$, $f^-(z) = Q^{-1}(z)$ for $|z| > 1$.

PROOF OF THEOREM 1 for $n = -3$. Obviously $r(z) = z^3$, $\overline{r(\bar z^{-1})} = r(z^{-1}) = z^{-3}$ and $g^+(z) = g^-(z) = g^{\frac{1}{2}}$. Therefore by (5)

$$T(z) = z^{-4} \operatorname{Res}\left[\frac{s^3P(s)g^{\frac{1}{2}}}{s - z};\, s = \infty\right]$$

$$= -g^{\frac{1}{2}}z^{-4} \operatorname{Res}\left[\frac{P(s^{-1})}{s^4(1 - sz)};\, s = 0\right]$$

$$= -\frac{g^{\frac{1}{2}}z^{-4}}{3!} \lim_{s\to 0} \frac{d^3}{ds^3}\left[\frac{P(s^{-1})}{1 - sz}\right]$$

$$= -g^{\frac{1}{2}}z^{-4}\lim_{s\to 0}\left[\frac{z^3P(s^{-1})}{(1-sz)^4}+\frac{z^2}{(1-sz)^3}\frac{dP(s^{-1})}{ds}\right.$$
$$\left.+\frac{z}{2(1-sz)^2}\frac{d^2P(s^{-1})}{ds^2}+\frac{1}{6(1-sz)}\frac{d^3P(s^{-1})}{ds^3}\right]$$

where

$$P(s^{-1}) = \exp\left[\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{se^{i\theta}+1}{se^{i\theta}-1}\log\left(1+\frac{f(e^{i\theta})}{g}\right)d\theta\right], \qquad |s| \neq 1 .$$

Denote for $|s| \neq 1$

$$K_j(s) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{ij\theta}}{(1-se^{i\theta})^{1+j}}\log\left(1+\frac{f(e^{i\theta})}{g}\right)d\theta , \qquad j = 1, 2, 3 .$$

Then $K_j(s) \to a_j = (2\pi)^{-1}\int_{-\pi}^{\pi} e^{ij\theta}\log[g + f(e^{i\theta})]\, d\theta$ as $s \to 0$ and

$$\frac{dP(s^{-1})}{ds} = -K_1(s)P(s^{-1}) ,$$
$$\frac{d^2P(s^{-1})}{ds^2} = -[2K_2(s) - K_1^2(s)]P(s^{-1}) ,$$
$$\frac{d^3P(s^{-1})}{ds^3} = -[6K_3(s) - 6K_1(s)K_2(s) + K_1^3(s)]P(s^{-1}) .$$

Consequently

$$(13)\quad T(z) = g^{\frac{1}{2}}z^{-4}\left[-z^3 + a_1z^2 + \frac{2a_2 - a_1^2}{2}z + \frac{6a_3 - 6a_1a_2 + a_1^3}{6}\right]P(\infty)$$
$$= g^{\frac{1}{2}}\left[-z^{-1} + a_1z^{-2} + \frac{2a_2 - a_1^2}{2}z^{-3} + \frac{6a_3 - 6a_1a_2 + a_1^3}{6}z^{-4}\right]P^{-1}(0) .$$

Using this result we obtain

$$E = \sum \mathrm{Res}\,[T(z)P^{-1}(z)g^{\frac{1}{2}};\ z = 0,\ z = \infty]$$
$$= \mathrm{Res}\,[gz^{-1}P^{-1}(z^{-1})P^{-1}(0);\ z = 0] + \mathrm{Res}\,[T(z)P^{-1}(z)g^{\frac{1}{2}};\ z = 0]$$
$$= g + g\lim_{z\to 0}\left[-P^{-1}(z) + a_1\frac{dP^{-1}(z)}{dz} + \frac{2a_2 - a_1^2}{4}\frac{d^2P^{-1}(z)}{dz^2}\right.$$
$$\left.+\frac{6a_3 - 6a_1a_2 + a_1^3}{36}\frac{d^3P^{-1}(z)}{dz^3}\right]P^{-1}(0) .$$

Now the assertion follows since

$$P^{-1}(z) = \exp\left[\frac{1}{4\pi}\int_{-\pi}^{\pi}\frac{z+e^{i\theta}}{z-e^{i\theta}}\log\left(1+\frac{f(e^{i\theta})}{g}\right)d\theta\right],$$
$$\frac{dP^{-1}(z)}{dz^2} = -M_1(z)P^{-1}(z) ,$$
$$\frac{d^2P^{-1}(z)}{dz^2} = -[2M_2(z) - M_1(z)]P^{-1}(z) ,$$
$$\frac{d^3P^{-1}(z)}{dz^3} = -[6M_3(z) - 6M_1(z)M_2(z) + M_1^3(z)]P^{-1}(z)$$

where for $|z| \neq 1$

$$M_j(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{e^{i\theta}}{(e^{i\theta}-z)^{1+j}}\log\left(1+\frac{f(e^{i\theta})}{g}\right)d\theta , \qquad j = 1, 2, 3 ,$$

and $M_j(z) \to \overline{a_j}$ as $z \to 0$. Note also that insertion of (13) into (6) yields

$$\begin{aligned} h_0(z) &= z^3 + z^3\left[-1 + a_1 z^{-1} + \frac{2a_2 - a_1^2}{2} z^{-2} + \frac{6a_3 - 6a_1 a_2 + a_1^3}{6} z^{-3}\right] \\ &\qquad \times P^{-1}(0)P^{-1}(z) \\ &= z^3[1 - P^{-1}(0)P^{-1}(z)] \\ &\qquad + \left[a_1 z^2 + \frac{2a_2 - a_1^2}{2} z + \frac{6a_3 - 6a_1 a_2 + a_1^3}{6}\right] P^{-1}(0)P^{-1}(z)\,. \end{aligned}$$

Propositions 1 and 2 were applied in order to obtain all the results of Section 3. Naturally, they can be used for deriving other error formulae and the optimal transfer function corresponding to each case. It is also remarkable that the method used in proving Propositions 1 and 2 is applicable to certain filtering problems involving noise correlated with the signal.

Acknowledgment. The author is grateful to Professor Moshe Zakai for his guidance and useful suggestions. Thanks are also due to Professor Jacob Ziv and to the referee.

REFERENCES

[1] GRENANDER, U. and SZEGÖ, G. (1958). *Toeplitz Forms and their Applications.* Univ. of California Press, 44.

[2] HOFFMAN, K. (1962). *Banach Spaces of Analytic Functions.* Prentice Hall, Englewood Cliffs.

[3] SNYDER, D. L. (1965). Some useful expression for optimum linear filtering in white noise. *Proc. IEEE* **53** 629–630.

[4] SNYDER, D. L. (1966). Optimum linear filtering of an integrated signal in white noise. *IEEE Trans. Aerospace and Electro. Systems* AES-2 231–232.

[5] SNYDERS, J. (1972). Error expressions for optimal linear filtering of stationary processes. *IEEE Trans. Information Theory* **IT-18** 574–582.

[6] VITERBI, A. J. (1965). On the minimum mean square error resulting from linear filtering of stationary signals in white noise. *IEEE Trans. Information Theory* **IT-11** 594–595.

[7] VITERBI, A. J. and CAHN, C. R. (1964). Optimum coherent phase and frequency demodulation of a class of modulating spectra. *IEEE Trans. Space Electronics and Telemetry* SET-**10** 95–102.

[8] WIENER, N. and MASANI, P. (1957). The prediction theory of multivariate stochastic processes, I. *Acta Math.* **98** 111–150.

[9] YAO, K. (1971a). On the direct calculations of MMSE of linear realizable estimator by Toeplitz form method. *IEEE Trans. Information Theory* **IT-17** 95–97.

[10] YAO, K. (1971b). An alternative approach to the linear casual least square filtering theory. *IEEE Trans. Information Theory* **IT-17** 232–240.

[11] YOVITS, M. C. and JACKSON, J. L. (1955). Linear filter optimization with game theory considerations. *1955 IRE Nat. Conv. Rec.* Part 4 193–199.

[12] ZAKAI, M. and SNYDERS, J. (1971). The Yovits-Jackson and related formulae for general spectral measures. EE Report No. 156, Department of Electrical Engineering, Technion, Haifa.

FACULTY OF ELECTRICAL ENGINEERING
TECHNION
HAIFA, ISRAEL