ARKIV FÖR MATEMATIK Band 4 nr 28

Communicated 22 February 1961 by H. CRAMÉR and H. WOLD

Extremal representation of stationary stochastic processes

By ENDERS A. ROBINSON

1. We let the real variables t and ω represent time and angular frequency respectively. A time-function $b(t)$ may be considered to be the *impulse response* function of a time-invariant linear system. If $b(t)$ is one-sided (i.e. if $b(t)=0$ for $t<0$ so that no response occurs prior to its stimulus) then the system is said to be *realizable*. The *transfer function* of a system with impulse response $b(t)$ is defined to be the formal Laplace transform

$$B(p)=\int_{-\infty}^{\infty} b(t)\,e^{-pt}\,dt, \quad p=\sigma+i\omega \quad (\sigma, \omega \text{ real}).$$

Letting $\sigma=0$ we obtain the formal Fourier transform

$$B(i\omega)=|B(i\omega)|\,e^{iP(i\omega)}=\int_{-\infty}^{\infty} b(t)\,e^{-i\omega t}dt,$$

where $|B(i\omega)|$ is called the *gain*, and $P(i\omega)$ the *phase-shift*, of the system. The *group-delay* of the system is defined to be

$$\tau_g=-\frac{dP(i\omega)}{d\omega}.$$

Input $f(t)$ and output $g(t)$ of the system are related by the convolution: $f(t)*b(t)=g(t)$. The Laplace transforms $F(p)$ and $G(p)$ of input and output respectively are related by the multiplication: $F(p)\,B(p)=G(p)$.

We shall call a function $w(t)$ a *wavelet* if it is one-sided (i.e. $w(t)=0$ for $t<0$) and L^2 (i.e. $\int_0^\infty |w(t)|^2\,dt<\infty$). The function space of all wavelets $w(t)$ (with measure dt) will be denoted by $L^2(0,\infty)$.

2. A purely non-deterministic, second-order, stationary stochastic process $x(t)$ (for continuous time parameter t) has a one-sided *moving-average representation*

$$x(t)=\int_{s=-\infty}^{t} w(t-s)\,dy(s), \tag{1}$$

where $w(t)$ is a wavelet and $y(s)$ is a process with orthogonal increments for which $E\{|dy(s)|^2\}=ds$. The corresponding *spectral representation* [Cramér, 1] is

$$x(t) = \frac{1}{2\pi} \int_{\omega=-\infty}^{\infty} e^{i\omega t} W(i\omega)\, dY(i\omega), \tag{2}$$

where $W(i\omega)$ is the L^2-Fourier transform of $w(t)$ and the formal derivative $dY(i\omega)/d\omega$ is the formal Fourier transform of the formal derivative $dy(s)/ds$. The $Y(i\omega)$ process has orthogonal increments for which

$$E\{|dY(i\omega)|^2\} = 2\pi\, d\omega.$$

For any given stationary stochastic process $x(t)$ there are infinitely many such representations; more precisely, there is a one-sided moving average representation, and a corresponding spectral representation, for every wavelet $w(t)$ that satisfies $\int_0^\infty w(t+s)\,\overline{w(s)}\, ds = \phi(t)$, where $\phi(t)$ is the autocovariance of the process. In this paper, we shall give necessary and sufficient conditions that a representation possess extremal properties.

3. A constant A_0 of modulus 1 is called a trivial all-pass transfer function. The function

$$A_1(p) = \prod_k \frac{p_k - p}{\overline{p}_k + p} \frac{|p_k - 1|}{p_k - 1} \frac{|p_k + 1|}{p_k + 1},$$

where $\{p_k\}$ is a non-empty set satisfying Re $p_k > 0$ and

$$\sum_k \frac{\text{Re } p_k}{1 + |p_k|^2} < \infty$$

is called a Type 1 all-pass transfer function. The function

$$A_2(p) = e^{-\alpha p} \quad (\alpha > 0)$$

is called a Type 2 all-pass transfer function. The function

$$A_3(p) = \exp\left[-\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1 - i\lambda p}{p - i\lambda}\, d\beta(\lambda)\right],$$

where $\beta(\lambda)$ is a non-decreasing function whose derivative vanishes almost everywhere and $0 < \beta(\infty) - \beta(-\infty) < \infty$ is called a Type 3 all-pass transfer function. The function

$$A(p) = A_0 A_1(p) A_2(p) A_3(p)$$

(where any or several of the factors on the right may be absent) is called an *all-pass transfer function.* If all the factors except A_0 are absent, then $A(p)$ is called *trivial*; otherwise $A(p)$ is called *non-trivial.*

4. If $M(i\omega) \geqslant 0$,

$$\int_{-\infty}^{\infty} |M(i\omega)|^2\, d\omega < \infty, \quad \int_{-\infty}^{\infty} \frac{\log M(i\omega)}{1+\omega^2}\, d\omega > -\infty,$$

then $$W_0(p) = \exp\left[\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1 - i\lambda p}{p - i\lambda} \frac{\log M(i\lambda)}{1+\lambda^2} d\lambda\right]$$

is called the *minimum-delay* transfer function for the gain $M(i\lambda)$. The wavelet $w_0(t)$ whose Laplace fransform is $W_0(p)$ is called the minimum-delay wavelet for the gain $M(i\lambda)$.

Krylov [4] and Karhunen [3] have established the following *canonical representation* for the transfer function of a system whose impulse response is a wavelet.

Lemma 1. *$W(p)$ is the L^2-Laplace transform of a wavelet if and only if, in the right half p-plane (i. e. Re $p > 0$),*

$$W(p) = A(p) W_0(p),$$

where $A(p)$ is an all-pass transfer function and $W_0(p)$ is the minimum-delay transfer function with the same gain as $W(p)$. This representation of $W(p)$ is unique.

In addition Karhunen [3] has established the following result.

Lemma 2. *Let $w(t)$ and $v(t)$ be wavelets and let*

$$\int_0^{\infty} w(t-r)\overline{v(t)}\,dt = 0 \quad \text{for} \quad r > 0.$$

Then $v(t) \equiv 0$ if and only if

$$w(t) = A_0 w_0(t)$$

where $|A_0| = 1$ and $w_0(t)$ is the minimum-delay wavelet with the same gain as $w(t)$.

5. From the canonical representation (Lemma 1) it can be shown that a system $A(p)$ is realizable and has gain $|A(i\omega)| = 1$ for almost all frequencies ω if and only if the system is an all-pass system. Moreover it is seen that in the right half p-plane (i. e. Re $p > 0$) an all-pass transfer function $A(p)$ is analytic and has zeros $\{p_k\}$ (where the set $\{p_k\}$ may be empty).

The following three lemmas can be established directly from the definition of the all-pass transfer function.

Lemma 3. *The group-delay τ_g of an all-pass transfer function is positive (resp. zero) for $-\infty < \omega < \infty$ if and only if the all-pass transfer function is non-trivial (resp. trivial).*

Lemma 4. *The modulus $|A(p)|$ of an all-pass transfer function satisfies $|A(p)| < 1$ (resp. $|A(p)| = 1$) in the right half p-plane if and only if the all-pass transfer function is non-trivial (resp. trivial).*

It follows from the canonical representation (Lemma 1) that if a wavelet $w(t)$ is the input to an all-pass system $A(p)$, then the output $v(t)$ is also a wavelet. Because $|A(i\omega)| = 1$, we have

$$|V(i\omega)| = |A(i\omega)\, W(i\omega)| = |W(i\omega)|$$

so the output wavelet $v(t)$ has the same gain as the input wavelet $w(t)$. Consequently, from Bessel's equality (for quadratic integrals of L^2-Fourier transforms) it follows that

$$\int_0^\infty |w(t)|^2\, dt = \int_0^\infty |v(t)|^2\, dt$$

which says that input and output wavelets have the same *total energy*. The following lemma, however, tells us that the *partial energy* of the output wavelet is delayed with respect to the partial energy of the input wavelet.

Lemma 5. *Let a wavelet $w(t)$ be the input to an all-pass system $A(p)$ and let the wavelet $v(t)$ be the resulting output. If the all-pass system is non-trivial, the partial energy of the input exceeds the partial energy of the output for some $\alpha > 0$:*

$$\int_0^\alpha |w(t)|^2\, dt > \int_0^\alpha |v(t)|^2\, dt.$$

If the all-pass system is trivial, the partial energy of the input equals the partial energy of the output for all $\alpha > 0$.

6. We now have the minimum-delay wavelet theorem.

Theorem 1. *Let $w(t)$ be a wavelet in the class of all wavelets with gain $M(i\omega)$. Then each of the following conditions is necessary and sufficient that*

$$w(t) = A_0\, w_0(t),$$

where $w_0(t)$ is the minimum-delay wavelet with gain $M(i\omega)$ and $|A_0| = 1$:

(*a*) *The set $\{w(t-r), r \geqslant 0\}$ is closed in $L^2(0, \infty)$.*
(*b*) *The group-delay of $w(t)$ is a minimum for $-\infty < \omega < \infty$.*
(*c*) *The modulus $|W(p)|$ is a maximum in the right half p-plane.*
(*d*) *The partial energy $\int_0^\alpha |w(t)|^2\, dt$ is a maximum for all $\alpha \geqslant 0$.*
(*e*) *For a purely non-deterministic, stationary stochastic process $x(t)$ with spectral density $\Phi(i\omega) = M^2(i\omega)$ and with moving-average representation* (1), *the least-squares linear prediction $z(t)$ of $x(t+\alpha)$, $\alpha > 0$, from the whole past $x(s)$, $s \leqslant t$, is*

$$z(t) = \int_{-\infty}^t w(t+\alpha-s)\, dy(s),$$

the minimum prediction error is

$$x(t+\alpha)-z(t)=\int_t^{t+\alpha} w(t+\alpha-s)\,d\,y(s)$$

and the minimum mean-square prediction error is given by the partial energy

$$E\{|x(t+\alpha)-z(t)|^2\}=\int_t^{t+\alpha}|w(t+\alpha-s)|^2\,d\,s=\int_0^{\alpha}|w(\tau)|^2\,d\,\tau.$$

(*f*) *For a purely non-deterministic, stationary stochastic process* $x(t)$ *with spectral density* $\Phi(i\omega)=M^2(i\omega)$ *and with moving-average representation* (1), *the closed linear manifold spanned by* $x(s)$, $s\leqslant t$, *is the same as the closed linear manifold spanned by* $y(s)$, $s\leqslant t$, *for all* t.

(*g*) *The function* $\Gamma(i\omega)$ *of the form*

$$\Gamma(i\omega)=\underset{N\to\infty}{\mathrm{l.i.m.}}\sum_{j=1}^{N} c_{Nj}\,e^{-i\omega r_{Nj}}$$

(*where* c_{Nj} *are complex constants and* $r_{Nj}\geqslant 0$) *that is determined by the method of least squares to approximate* $e^{i\omega\alpha}$ ($\alpha>0$) *with respect to measure* $M^2(i\omega)\,d\,\omega/2\pi$ *is*

$$\Gamma(i\omega)=\frac{1}{W(i\omega)}\int_0^{\infty} w(\alpha+t)\,e^{-i\omega t}\,d\,t$$

and the minimum mean-square error is given by the partial energy

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}|e^{i\omega\alpha}-\Gamma(i\omega)|^2 M^2(i\omega)\,d\,\omega=\int_0^{\alpha}|w(\tau)|^2\,d\,\tau.$$

Proof. Condition (*a*) follows immediately from Lemma 2 by noting that a set in Hilbert space is closed if and only if any element orthogonal to each member of the set identically vanishes. Conditions (*b*), (*c*), and (*d*) follow from Lemmas 3, 4, and 5, respectively, and from the canonical representation (Lemma 1). Conditions (*e*) and (*f*) follow from the work on prediction theory by Hanner [2], Karhunen [3], and Wiener [6]. Condition (*g*) follows from condition (*e*) by utilizing the isomorphism of the closed linear manifold spanned by $x(s)$, $s\leqslant t$ (probability measure) and the closed linear manifold spanned by $e^{i\omega s}$, $s\leqslant t$ (measure $M^2(i\omega)\,d\,\omega/2\pi$) such that $x(t)$ and $e^{i\omega t}$ ($-\infty<t<\infty$) are corresponding elements (see, e.g., Robinson [5], p. 83). Q.E.D.

7. Summing up, the spectral density $\Phi(i\omega)$ of a purely non-deterministic stationary stochastic process being given, we see that there exists a class of different spectral representations (2), such that the transfer function $W(p)$ satisfies $W(i\omega)=A(i\omega)\sqrt{\Phi(i\omega)}$, where $A(p)$ is an arbitrary all-pass transfer function. Alternatively, the autocovariance function $\phi(t)$ (which is the L^1-Fourier transform of $\Phi(i\omega)$) being given, there exists a corresponding class of moving-average representations (1), such that the wavelet $w(t)$ (which is the L^2-Fourier transform of $W(i\omega)$) satisfies $\int_0^{\infty} w(t+s)\,\overline{w(s)}\,d\,s=\phi(t)$. Among these representations there is one called the predictive (or Woldian) decomposition given by

$$x(t) = A_0 \int_{s=-\infty}^{t} w_0(t-s)\, d\, y_0(s) = \frac{A_0}{2\pi} \int_{\omega=-\infty}^{\infty} e^{i\omega t}\, W_0(i\omega)\, d\, Y_0(i\omega)$$

(where $|A_0| = 1$, and $w_0 \leftrightarrow W_0$ is minimum-delay for the gain $\sqrt{\Phi(i\omega)}$) which has extremal properties as given by Theorem 1.

The realizable system with input $x(t)$ and output $z(t)$ where $z(t)$ is the least-squares linear prediction of $x(t+\alpha)$, $\alpha > 0$, has transfer function

$$\Gamma(i\omega) = \underset{N\to\infty}{\text{l.i.m.}} \sum_{j=1}^{N} c_{Nj}\, e^{-i\omega r_{Nj}} = \frac{1}{W_0(i\omega)} \int_0^{\infty} w_0(\alpha+t)\, e^{-i\omega t}\, dt$$

(where c_{Nj} are complex constants and $r_{Nj} \geqslant 0$). The minimum mean-square prediction error has the decomposition

$$E\{|x(t+\alpha) - z(t)|^2\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} |e^{i\omega\alpha}\, W(i\omega) - \Gamma(i\omega)\, W(i\omega)|^2\, d\omega$$

$$= \int_{-\infty}^{\infty} |w(t+\alpha) - \underset{N\to\infty}{\text{l.i.m.}} \sum_{j=1}^{N} c_{Nj}\, w(t - r_{Nj})|^2\, dt$$

$$= \int_{-\alpha}^{0} |w(t+\alpha)|^2\, dt + \int_0^{\infty} |w(t+\alpha) - \underset{N\to\infty}{\text{l.i.m.}} \sum_{j=1}^{N} c_{Nj}\, w(t - r_{Nj})|^2\, dt.$$

We note that the first term is the partial energy of the wavelet $w(t)$, and the second term vanishes for arbitrary α if and only if the set $\{w(t-r),\ r \geqslant 0\}$ is closed, i.e. if and only if $w(t) = A_0\, w_0(t)$.

Uppsala, January, 1961.

REFERENCES

1. CRAMÉR, H., On harmonic analysis in certain functional spaces. *Ann. Math. 41*, pp. 215–230 (1942).
2. HANNER, O., Deterministic and non-deterministic stationary random processes. *Ark. Matematik 1*, pp. 161–177 (1949).
3. KARHUNEN, K., Über die Struktur stationärer zufälliger Funktionen. *Ark. Matematik 1*, pp. 141–160 (1949).
4. KRYLOV, V. I., On functions regular in a half-plane (Russian). *Mat. Sbornik* N.S. *6* (48), pp. 95–138 (1939).
5. ROBINSON, E. A., *Infinitely many variates*. Griffin, London, 1959.
6. WIENER, N., *Stationary time series*. Wiley, New York, 1949.
7. WOLD, H., *A study in the analysis of stationary time series*. Thesis, University of Stockholm. Uppsala, 1938 (second edition, 1954).

Tryckt den 30 januari 1962

Uppsala 1962. Almqvist & Wiksells Boktryckeri AB