Win Win Htay
Nagoya Math. J.
Vol. 149 (1998), 9–17

OPTIMALITIES FOR RANDOM FUNCTIONS LEE-WIENER'S NETWORK AND NON-CANONICAL REPRESENTATION OF STATIONARY GAUSSIAN PROCESSES

WIN WIN HTAY

Abstract. Representation of a Gaussian process in terms of a Brownian motion is a powerful tool in the investigation of its structure. Among various representations is the canonical representation which is viewed as the best one from the viewpoint of the prediction theory. We have discovered some significance of non-canonical representations and discuss their optimality in an information theoretical approach.

§1. Introduction

The probabilistic structure of a stationary Gaussian process $\{X(t)\}$ with $E(X(t)) = 0$ can well be illustrated by a representation which is expressed as a linear functional of a Brownian motion $\{B(t)\}$ in such a way that

$$X(t) = \int_{-\infty}^{t} F(t-u)\dot{B}(u)du, \qquad \dot{B}(u) = \frac{d}{du}B(u) \tag{1.1}$$

with a kernel function F which is non-random. There are various representations of the form (1.1); there exists, among others, the canonical representation which is unique for a given $\{X(t)\}$.

The main characteristic of the canonical representation of a Gaussian process is the fact that the Brownian motion has exactly the same information as $X(t)$ up to any time t. While, in the case where the representation (1.1) is non-canonical, the $\{X(t)\}$ has less information than the associated Brownian motion $\{B(t)\}$ up to some $t \in (-\infty, \infty)$, and hence the same property holds for every t by the stationarity. We are interested in giving detailed interpretation to non-canonical representations from the view point of information theory.

We restrict our attention to the cases where the representation of a given stationary Gaussian process can actually be realized as the output of

Received August 20, 1996.

linear networks with white noise input. More precisely, we shall deal with non-canonical representation for which the Fourier inverse transform $A(\lambda)$ of their kernel function involves a factor corresponding to the Lee-Wiener's networks formed by LCR circuits. The $A(\lambda)$ is the transmission function and generally it is expressed as a product of three functions $C(\lambda)$, $B(\lambda)$ and $S(\lambda)$, each of which is of particular type. Our restriction on the stationary process in question may be rephrased in such a way that $A(\lambda)$ is a product of the form

$$A(\lambda) = C(\lambda) \cdot B(\lambda), \tag{1.2}$$

where C corresponds to the canonical kernel and B is the Blaschke product. In fact the singular part $S(\lambda)$ would not appear in actual applications.

We shall discuss how a given $A(\lambda)$ can be factorized in the form (1.2) and how to form Lee-Wiener's networks that correspond to $B(\lambda)$.

Our main result is to show that for the non-canonical case how much extra information is contained in the white noise input $\{\dot{B}(t)\}$ compared to the observed process $\{X(t)\}$ up to time t. At the same time our results would give an actual explanation to the non-canonical representations, keeping the optimality in mind, when the theory is applied to communication system.

§2. Background

Let $\{X(t), t \in R^1\}$ be a centered, mean continuous, weakly stationary process. The covariance function is given by

$$\gamma(h) = E[X(t+h)X(t)].$$

We define Hilbert space H spanned by

$$\{\sum a_k X(t_k); \text{finite sum}, a_k \in C, t_k \in R\},$$

where the topology is introduced by the norm $|| \quad ||$ defined by

$$||X|| = [E(|X|^2)]^{1/2}.$$

Introduce an unitary operator U_t acting on H in such a way that

$$U_t X(s) = X(t+s).$$

https://doi.org/10.1017/S0027763000006528 Published online by Cambridge University Press

The system has the properties

$$U_t U_s = U_{t+s}$$
$$U_t \longrightarrow I \text{ as } t \longrightarrow 0 \text{ (strongly)}.$$

Thus $\{U_t; t \in R^1\}$ is a continuous one-parameter group of unitary operators. We can therefore appear to the Stone's theorem to obtain the spectral representation of the form

$$U_t = \int e^{it\lambda} dE(\lambda), \tag{2.1}$$

where $\{E(\lambda); \lambda \in R^1\}$ is a resolution of the identity. Since $X(t) = U_t X(0)$ holds, we have the spectral representation of $X(t)$:

$$X(t) = \int e^{it\lambda} dE(\lambda) X(0). \tag{2.2}$$

Assuming the purely nondeterministic property of $X(t)$, the covariance function $\gamma(h)$ is expressed in the form

$$\gamma(h) = \int e^{ih\lambda} f(\lambda) d\lambda,$$

where f is spectral density function, and we have

$$\int \frac{\log f(\lambda)}{1+\lambda^2} d\lambda > -\infty.$$

The formula (2.2) can be expressed in the form

$$X(t) = \int e^{it\lambda} A(\lambda) \dot{Z}(\lambda) d\lambda \tag{2.3}$$

where $\dot{Z}(\lambda)$ is the Fourier transform of $\dot{B}$, and we have

$$f(\lambda) = \frac{1}{2\pi} |A(\lambda)|^2.$$

We now remind that $\{X(t)\}$ is Gaussian and has a representation of the form (1.1). There is a relation between the kernel $F(t-u)$ and $A(\lambda)$ in (2.3):

$$|\tilde{F}(\lambda)|^2 = \frac{1}{2\pi} |A(\lambda)|^2 \quad (= f(\lambda)), \tag{2.4}$$

where $\tilde{F}$ is the Fourier inverse transform of F.

Another important tool of our approach is the theory of the Hardy class (see Hoffman [3]). The $A(\lambda)$ has the extension $A(\omega)$ which is holomorphic on the lower half plane C^- with boundary value $A(\lambda)$ and it is factorized in the form

$$A(\omega) = C(\omega) \cdot B(\omega) \cdot S(\omega). \tag{2.5}$$

The factor $C(\omega)$ corresponds to the canonical representation (see Hida Hitsuda [2]), $B(\omega)$ is the Blaschke product expressed in the form

$$B(\omega) = \prod_n \frac{\omega - \omega_n}{\omega - \bar{\omega}_n} \cdot \frac{\bar{\omega}_n - i}{\omega_n + i}, \tag{2.6}$$

where $\sum_n (|\operatorname{Im} \omega_n|)/(1 + |\omega_n|^2) < \infty$, $\operatorname{Im} \omega_n < 0$, and $S(\omega)$ is the singular factor which is ignored by the assumption mentioned in Section 1.

It should be noted that three factors can uniquely be determined up to constants. In particular $C(\omega)$ is expressed in the form

$$C(\omega) = \sqrt{2\pi} \exp\left[-\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{1 + \lambda\omega}{\lambda - \omega} \frac{\log f(\lambda)}{1 + \lambda^2} d\lambda\right], \tag{2.7}$$

which has no zero points in C^-. While, $B(\omega)$ is determined by zero points of $A(\omega)$.

The following assertion can easily be proved.

Proposition 1. *It holds that*

$$|A(\omega)| \leq |C(\omega)|, \qquad \omega \in C^-.$$

§3. The Blaschke product and the Lee-Wiener's networks

In this section, we discuss the correspondence between the Blaschke product $B(\omega)$, given by (2.6), and the Lee-Wiener's networks.

Proposition 2. *The zero points ω_n of $B(\omega)$ are either purely imaginary $-i\alpha$, $\alpha > 0$, or pairs of the form $\beta - i\alpha$ and $-\beta - i\alpha$, β real.*

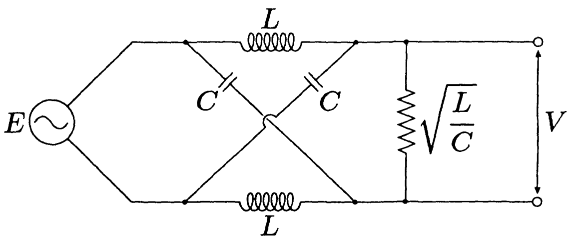
Proof. Since $X(t)$ is real valued, it is proved that $A(-\lambda) = \overline{A(\lambda)}$, and so is for $B(\lambda)$. This proves the assertion. □

Proposition 3. *Each factor of the Blaschke product corresponds to either a transmission function of Lee-Wiener's network (which is a lattice all-pass network) or its modification by shifts of constant frequency.*

Proof. First we consider a factor of the Blaschke product with purely imaginary root $-i\alpha$, $\alpha > 0$. For $\omega_n = -i\alpha$ we have

$$\frac{\omega - \omega_n}{\omega - \bar{\omega}_n} \cdot \frac{\bar{\omega}_n - i}{\omega_n + i} = (-1)\frac{\omega + i\alpha}{\omega - i\alpha}.$$

Now, we see that it is a transmission function that comes from the following lattice all-pass network (Fig.1).



$$\frac{V_\lambda}{E_\lambda} = \frac{1 - i\lambda\sqrt{LC}}{1 + i\lambda\sqrt{LC}}, \quad \frac{1}{\sqrt{LC}} = \alpha.$$

Fig.1

Namely, we have $V_\lambda / E_\lambda = (\lambda + i\alpha)(\lambda - i\alpha)$.

Second, we take a pair of factors of $B(\lambda)$ and have their product

$$\frac{\lambda - \beta + i\alpha}{\lambda - \beta - i\alpha} \cdot \frac{\beta + i\alpha - i}{\beta - i\alpha + i} \cdot \frac{\lambda + \beta + i\alpha}{\lambda + \beta - i\alpha} \cdot \frac{-\beta + i\alpha - i}{-\beta - i\alpha + i}.$$

If λ is shifted by β to have λ', then we have

$$\frac{\lambda' + i\alpha}{\lambda' - i\alpha} \cdot \frac{\lambda' + 2\beta + i\alpha}{\lambda' + 2\beta - i\alpha}.$$

Hence, we form a network that corresponds to $(\lambda' + i\alpha)/(\lambda' - i\alpha)$ and then we take the second factor which requests the shift of λ by 2β to have a new, but similar, network.

We now claim that we can actually form a chain of networks as many as we want, corresponding to the factors of the Blaschke product. If there are infinitely many factors, we need to prove existence of the limit. Our assumption on ω_n's given in (2.6) guarantees the existence of the limit. (See Hibino-Hitsuda-Muraoka [6]). ∎

§4. Optmality in an information theoretical approach

In this section, significant difference between the canonical and non-canonical representation may be observed from the view point of information theory.

To discriminate canonical and non-canonical representations, we use the following notations from now on. So far as our question is concerned, we may assume that two representations have the same Brownian motion $B(t)$.

$$(4.1)\qquad \begin{aligned} X_1(t) &= \int_{-\infty}^{t} F_1(t-u)\dot{B}(u)du, \quad \text{canonical representation,} \\ X_2(t) &= \int_{-\infty}^{t} F_2(t-u)\dot{B}(u)du, \quad \text{non-canonical representation.} \end{aligned}$$

Both $\{X_1(t)\}$ and $\{X_2(t)\}$ are of course the same Gaussian process as the given $\{X(t)\}$.

1^0) We first observe a characteristic of canonical representation.

Since all the variables that we are concerned with form a Gaussian system, the information quantity $I(Y, X_1(t))$, $t \geq 0$, is given by

$$(4.2)\qquad I(Y, X_1(t)) = -\frac{1}{2}\log[1 - \rho(Y, X_1(t))^2]$$

where $Y = \int_{-\infty}^{0} \beta(u)\dot{B}(u)du$, β being square integrable and $\rho(Y, X_1(t))$ is the correlation coefficient.

There is a particular Y_0 such that

$$(4.3)\qquad I(Y_0, X_1(t)) \geq I(Y, X_1(t)).$$

Indeed, $I(Y, X_1(t))$ is maximal if Y is taken to be

$$Y_0 = \text{const.} \int_{-\infty}^{0} F_1(t-u)\dot{B}(u)du$$

which turns out to be the best predictor for $X_1(t)$, $t > 0$, up to constant, based on $\{X_1(s), s \leq 0\}$, by the assumption of canonical property. It is noted that Y_0 is eventually expressed as a linear function of the $X_1(u)$, $u \leq 0$.

2^0) Next, consider a non-canonical representation where non-trivial Blaschke product is involved in (1.2):

$$X_2(t) = \int_{-\infty}^{t} F_2(t-u)\dot{B}(u)du.$$

Set

$$Y_\beta = \int_{-\infty}^{0} \beta(u)\dot{B}(u)du,$$

Then

$$\sup_\beta I(Y_\beta, X_2(t)) > I(Y_0, X_1(t)), \quad t > 0, \tag{4.4}$$

holds, where Y_0 is the one for the canonical case in (4.3). It is understood that the system $\{B(u), u \le 0\}$ has more information than $\{X_2(t), t \le 0\}$.

We can further discuss non-canonical representation also from quantitative view point of information theory.

LEMMA. *Let $C(\omega)$ be the factor of (2.5). Then its boundary value $C(\lambda)$ has an orthogonal expansion of the form*

$$C(\lambda) = \sum_0^\infty a_n \phi_n(\lambda), \quad \phi_n(\lambda) = \frac{(\lambda + i\alpha)^n}{(\lambda - i\alpha)^{n+1}}, \tag{4.5}$$

where $a_n \ne 0$ for every n.

Proof. $\phi_n(\lambda)$, $n \ge 0$, are the Fourier transforms of the Laguerre functions which form a complete orthonormal basis of $L^2(R^-)$. To conclude the assertion we use the fact that if $a_n = 0$ were true for some n, then the process $\int e^{it\lambda}\phi_n(\lambda)\dot{Z}(\lambda)d\lambda$ would be a stationary Gaussian process that is a linear function of $\dot{B}(u)$, $u \le t$, and is independent of $X_1(t)$ at every t. This is a contradiction to the canonical property. □

To fix the idea we start from a factor $(\lambda + i\alpha)/(\lambda - i\alpha)$, $\alpha > 0$, of the Blaschke product with purely imaginary zero point.

Take this α and have (4.5). A non-canonical transmission function is formed by

$$A(\lambda) = C(\lambda) \cdot \frac{\lambda + i\alpha}{\lambda - i\alpha}.$$

Then, we have

$$A(\lambda) = \sum_0^\infty a_n \phi_{n+1}(\lambda). \tag{4.6}$$

https://doi.org/10.1017/S0027763000006528 Published online by Cambridge University Press

Write

$$A(\lambda) = \sum_{0}^{\infty} b_n \phi_n(\lambda).$$

Then, for the expansion of $A(\lambda)$, we have $b_0 = 0$, $b_n = a_{n-1}$, $n \geq 1$. Note that formula (4.1) are expressed in the form

$$(4.1)' \qquad \begin{aligned} X_1(t) &= \int e^{it\lambda} C(\lambda) \dot{Z}(\lambda) d\lambda \quad \text{and} \\ X_2(t) &= \int e^{it\lambda} C(\lambda) \frac{\lambda + i\alpha}{\lambda - i\alpha} \dot{Z}(\lambda) d\lambda, \quad \text{respectively.} \end{aligned}$$

Remark. Let $\{l_n(u), n \geq 0\}$ be a complete orthonormal system $L^2(R^-)$ formed by Laguerre functions with parameter $\alpha > 0$. Then $\{\int_{-\infty}^{0} l_n(u) \dot{B}(u) \cdot du\}$ is an independent system. The Plancherel formula for stochastic integral proves

$$\int_{-\infty}^{0} l_n(u) \dot{B}(u) du = \int_{-\infty}^{\infty} \phi_n(\lambda) \dot{Z}(\lambda) d\lambda.$$

Hence, by using the expansion (4.6), we see that $\int A(\lambda) \dot{Z}(\lambda) d\lambda$ gives a sum of *independent* random variables; the leading term is missing compared to $\int C(\lambda) \dot{Z}(\lambda) d\lambda$.

The loss of information for non-canonical representation can be calculated as follows.

THEOREM. *Assume that the canonical representation of $X_1(t)$ is given by $C(\lambda)$ of the form* (4.5). *For a non-canonical representation of $X_2(t)$ given by $A(\lambda)$ in* (4.6), *we have*

$$\sup_{Y} I(Y, X_1(t)) = -\log |a_0| + \log \sigma$$

holds for any t, where $\sigma^2 = E(X_i(t)^2)$, $i = 1, 2$, $\sigma > 0$, and Y is any linear function of the $X_2(u)$, $u \leq t$.

Proof. Since the optimal Y is taken to be the projection of $X_1(t)$ down to the space spanned by the $\int e^{it\lambda} \phi_n(\lambda) \dot{Z}(\lambda) d\lambda$, $n \geq 1$, it is expresses as $\sum_{n=1}^{\infty} a_n \int e^{it\lambda} \phi_n(\lambda) \dot{Z}(\lambda) d\lambda$, so the Schwarz inequality proves that $\sup_Y \rho(Y, X_1(t))^2 = \sum_{n=1}^{\infty} (|a_n|^2)/\sigma^2$, $\sigma^2 = \sum_{n=0}^{\infty} |a_n|^2$. This proves the theorem. $\square$

We remind that the $\int e^{it\lambda}\phi_n(\lambda)\dot{Z}(\lambda)d\lambda$, $n \geq 1$, can be formed from $X_2(u)$, $u \leq t$, by the whitening technique, noting that $\dot{Z}(\lambda)$ in the canonical case is now replaced by $\dot{Z}_1(\lambda) = (\lambda + i\alpha)/(\lambda - i\alpha)\dot{Z}(\lambda)$. The technique may be illustrated by the following formula

$$X_2(t) = \int e^{it\lambda}\left\{C(\lambda)\frac{\lambda + i\alpha}{\lambda - i\alpha}\right\}\dot{Z}(\lambda)d\lambda = \int e^{it\lambda}C(\lambda)\dot{Z}_1(\lambda)d\lambda.$$

Conclusion Remark. For a general Blaschke product, we can repeat the same procedure as many times as the number of the factors in the Blaschke product and observe how much information is eventually lost. In this sense, canonical representation has optimality. Further results related to such an optimality will be discussed later in connection with other results.

Acknowledgements. The author wishes to express her gratitude to the Supervisor Professor Y. Shikata, Graduate School of Mathematics, Nagoya University. Thanks are due to Professor T. Hida, Professor S. Ihara and Professor L. Accardi for their kind comments. The author thanks also to the referee for his kind comments.

References

[1] T. Hida, Brownian Motion, Springer-Verlag, 1980.
[2] T. Hida and M. Hitsuda, Gaussian Processes, Amer. Math. Soc. Providence, Rhode Island, 1993.
[3] K. Hoffman, Banach Spaces of Analytic Functions, Prentice-Hall, Inc., 1962.
[4] S. Ihara, Information Theory for Continuous Systems, World Scientific, 1993.
[5] Yu. A. Rozanov, Stationary Random Processes, Holden-Day, 1967.
[6] Y. Hibino, M. Hitsuda and H. Muraoka, Remarks on a non-canonical representation for a stationary Gaussian process.
[7] N. Wiener, Nonlinear Problems in Random Theory, M.I.T. Press, Second edition, 1963.

Graduate School of Mathematics
Nagoya University
Chikusa-ku, Nagoya 464-8602
Japan

CURRENT ADDRESS:
Department of Mathematics
University of Yangon
Yangon
Myanmar