

Expectation of nonlinear functions of Gaussian processes

G. TERDIK (Debrecen)

1. *Introduction, notations.* This paper has two aims one is to show that the formula of Bonnet [1] for calculating the expectation of nonlinear function without memory of a Gaussian stochastic process is valid under weak assumption; the other one is to examine — as a consequence — the Generalized Appel—Wick polynomial system of several variables.

The vector space of rapidly decreasing functions in the n -dimensional euclidean space R^n will be denoted by $\mathcal{S}_n$ i.e.

$$\mathcal{S}_n = \{g | g \in \mathcal{C}^\infty(R^n), \sup_{|\alpha| \leq N} \sup_{x \in R^n} (1 + |x|^2)^N |(D_\alpha g)(x)| < \infty, N = 0, 1, \dots\}.$$

Here $|x|^2 = \sum_{i=1}^n x_i^2$, α multi-index is an n -tuple of nonnegative integers α_i , $|\alpha| = \sum_{i=1}^n \alpha_i$, D_α is a differential operator defined by

$$D_\alpha = i^{-|\alpha|} D^\alpha = \left(\frac{1}{i} \frac{\partial}{\partial x_1}\right)^{\alpha_1} \cdots \left(\frac{1}{i} \frac{\partial}{\partial x_n}\right)^{\alpha_n}$$

and $C^\infty(R^n)$ is the set of complex function g defined in R^n for which $D_\alpha g$ is continuous for every multi-index α . The elements of the dual space $\mathcal{S}'_n$ of $\mathcal{S}_n$ are called tempered distribution, they are the continuous linear functionals on $\mathcal{S}_n$. It is known that every $f \in L^p(R^n)$, $1 \leq p \leq \infty$, every polynomial and every measurable function whose absolute value is majorized by some polynomial is tempered distribution. It is customary to identify the tempered distribution u_f with the function f if

$$u_f(g) = \int_{R^n} g f dm_n, \quad g \in \mathcal{S}_n$$

(m_n denotes the normalized Lebesgue measure on R^n defined by $dm_n(x) = (2\pi)^{-n/2} dx$) and to say that such distributions are functions. The Fourier transform of a function $g \in L^1(R^n)$ is the function $\hat{g}$ defined by

$$\hat{g}(t) = \int_{R^n} g(x) e^{-i(t, x)} dm_n(x).$$

Note that if $g \in \mathcal{S}_n$ then $g \in L^1(R^n)$ and the Fourier transform is a continuous, linear, one-to-one mapping of $\mathcal{S}_n$ onto $\mathcal{S}'_n$ of period 4, whose inverse is also continuous ([2] 7.7). Associate with each tempered distribution u its Fourier transform $\hat{u}$ by

$$\hat{u}(g) = u(\hat{g}).$$

It is known ([2] 7.15) that the Fourier transform is a continuous, linear, one-to-one mapping of $\mathcal{S}'_n$ onto $\mathcal{S}'_n$ of period 4, whose inverse is also continuous. The space $\mathcal{C}^\infty(R^n)$ with the topology defined by the uniform convergence of differentials $D_\alpha g_k$ for every α on every compact set $K(\subseteq R^n)$ will be denoted by $\mathcal{C}_n$. The dual space of $\mathcal{C}_n$ is $\mathcal{C}'_n$.

The n dimensional random variable ξ is Gaussian with expectation μ and covariance matrix G if its characteristic function has the form

$$\varphi_\xi(t) = \exp \{i(t, \mu) - 1/2 t' G t\}.$$

It is easy to see that $\varphi_\xi(t) \in \mathcal{S}_n$ and so the associated distribution function $g_\xi \in \mathcal{S}_n$ too, if it exists.

2. Expectation of nonlinear functions

Lemma. a) *The convergence*

$$(2.1) \quad \lim_{n \rightarrow \infty} \sum_{k=0}^n \left(-\frac{1}{2}\right)^k \frac{(x' C x)^k}{k!} e^{-1/2 x' \Gamma x} = \exp \left\{-\frac{1}{2} x' G x\right\}$$

is valid in $\mathcal{C}_n$, where $G = \Gamma + C$ is the covariance matrix of a nonconstant random variable ξ and Γ is the diagonal matrix of diagonal elements of G . b) If $\inf_{|x|=1} |x' G x| > 0$, and

$$\max \left(\sup_{|x|=1} \left| \frac{x' C x}{x' G x} \right|, \sup_{|x|=1} \left| \frac{x' C x}{x' \Gamma x} \right| \right) < 1$$

then the convergence (2.1) is valid in $\mathcal{S}_n$ too.

PROOF. It is enough to show that for each multi-index α

$$\lim_{n \rightarrow \infty} \sup_{x \in R^n} |x^\alpha| e^{-1/2 x' \Gamma x} e^{-1/2 x' C x} \frac{|-1/2 x' C x|^n}{n!} = 0.$$

This, using the Stirling's formula follows from the inequality

$$\begin{aligned} & \sup_{|x|=1} \sup_{\lambda \geq 0} \sup_{0 \leq \vartheta \leq 1} |x^\alpha| \lambda^{|\alpha|+2n} e^{-(\lambda^2/2)x'(\Gamma+\vartheta C)x} \frac{|x' C x|^n}{2^n n!} \cong \\ & \cong \sup_{|x|=1} \sup_{\lambda \geq 0} |x^\alpha| \lambda^{|\alpha|+2n} \frac{|x' C x|^n}{2^n n!} (e^{-1/2 \lambda^2 x' \Gamma x} + e^{-1/2 \lambda^2 x' (\Gamma+C)x}) \cong \\ & \frac{e^{-(n+|\alpha|/2)}}{n! \varepsilon_1^{|\alpha|/2}} \left(n + \frac{|\alpha|}{2} \right)^{n+|\alpha|/2} 2^{|\alpha|/2} \left(\sup_{|x|=1} \left| \frac{x' C x}{x' (\Gamma+C)x} \right|^n + \sup_{|x|=1} \left| \frac{x' C x}{x' \Gamma x} \right|^n \right) \end{aligned}$$

where

$$\varepsilon_1 = \min \left(\inf_{|x|=1} |x' G x|, \inf_{|x|=1} |x' \Gamma x| \right).$$

Let the function f be tempered distribution and the density function g_ξ of the Gaussian random variable ξ exist then the expectation of $f(\xi)$ exists and

$$\begin{aligned} Ef(\xi) &= (2\pi)^{n/2} \int_{\mathbb{R}^n} f(x) g_\xi(x) dm_n(x) = (2\pi)^{n/2} u_f(g_\xi) = \\ &= (2\pi)^{n/2} \hat{u}_f((\hat{g}_\xi)^\sim) = (2\pi)^{n/2} \hat{u}_f(\varphi_\xi) \end{aligned}$$

where $\check{g}(x) = g(-x)$. So we can define for each $u \in \mathcal{S}'_n$ and Gaussian random variable ξ the expectation $E(u, \xi)$ by the formula

$$(2.2) \quad E(u, \xi) = (2\pi)^{n/2} \hat{u}(\varphi_\xi).$$

As the Fourier transform of $v \in \mathcal{C}$ is function ([3], Theorem 4. § 3, VI) i.e.

$$\hat{v}(y) = (2\pi)^{-n/2} v_x(e^{-i(x, y)})$$

we get that the expectation (2.2) equals

$$E(v, \xi) = (2\pi)^{n/2} \hat{v}(\varphi_\xi) = \int_{\mathbb{R}^n} \varphi_\xi(y) v_x(e^{-i(x, y)}) dm_n(y).$$

Formula (2.2) leads to the following useful calculation of the expectation.

Theorem. Under assumptions

- a) $u \in \mathcal{C}'_n$ and ξ is an arbitrary Gaussian random variable with mean 0.
- b) $\hat{u} \in \mathcal{S}'_n$ and ξ is such a Gaussian random variable $E\xi = 0$, that assumptions b) of Lemma are fulfilled.

The expectation $E(u, \xi)$ can be calculated by the formula

$$(2.3) \quad E(u, \xi) = \sum ED^\alpha(u, \xi^*) \prod_{i < j} \frac{G_{ij}^{k_{ij}}}{k_{ij}!}, \{ \alpha \equiv 0 | \alpha_i = \sum_{j=1}^{i-1} k_{ji} + \sum_{j \equiv i+1} k_{ij}, k_{ij} \equiv 0 \}.$$

Where ξ^* denotes such a random variable whose components are uncorrelated and $E\xi^* = E\xi$, $\text{var } \xi_i^* = \text{var } \xi_i = G_{ii}$.

PROOF. Since for each multi-index α

$$\begin{aligned} \hat{u}(x^\alpha \varphi_{\xi^*}) &= \hat{u}(x^\alpha (\hat{g}_{\xi^*})^\sim) = \hat{u}(((x)^\alpha \hat{g}_{\xi^*})^\sim) = \\ &= \hat{u}((-D_\alpha g_{\xi^*})^\sim) = (D_\alpha u)^\wedge(\varphi_{\xi^*}). \end{aligned}$$

We get

$$\begin{aligned} E(u, \xi) &= (2\pi)^{n/2} \hat{u} \left(\varphi_{\xi^*} \sum x^\alpha \prod_{i < j} \frac{G_{ij}^{k_{ij}}}{k_{ij}!} \right) = \\ &= \sum (2\pi)^{n/2} (D^\alpha u)^\wedge(\varphi_{\xi^*}) \prod_{i < j} \frac{C_{ij}^{k_{ij}}}{k_{ij}!} \end{aligned}$$

where sum is over $\{ \alpha \equiv 0 | \alpha_i = \sum k_{ji} + \sum k_{ij} \}$.

Corollary. One can compute the expectation $EF(\xi)$ by the formula (2.3) for an arbitrary Gaussian random variable ξ with mean 0, $F(y)$ is an analytical function of complex variables and for some constants C , N and B

$$(2.4) \quad |F(y)| \leq C(1+|y|)^N e^{B|I_m y|}.$$

PROOF. By the Theorem of Paley—Wiener ([3] § 4, VI) under the assumptions of the corollary there exists such a $v \in \mathcal{C}'_n$ that

$$F(y) = (2\pi)^{-n/2} v_x(e^{-i(x,y)}) = \hat{v}(y).$$

In the following we shall show some illustrative examples for the use of the above theorem i.e. formula (2.3).

Examples. Let ξ denote a Gaussian random variable $(0, G)$.

1. If $F \in \mathcal{C}'_n$ and $F(y) = \prod_{i=1}^n F_i(y_i)$ then by (2.3)

$$EF(\xi) = \sum_{\{\alpha \geq 0 | \alpha_i = \sum_{j=1}^n k_{ji} + \sum_{j=1}^n k_{ij}\}} \prod_{r=1}^n EF_r^{(\alpha_r)}(\xi_r) \prod_{i < j} \frac{G_{ij}^{k_{ij}}}{k_{ij}!}.$$

If $n=2$, then

$$EF_1(\xi_1) F_2(\xi_2) = \sum_{k=0}^{\infty} EF_1^{(k)}(\xi_1) EF_2^{(k)}(\xi_2) \frac{G_{12}^k}{k!}.$$

If $n=3$ then

$$EF_1(\xi_1) F_2(\xi_2) F_3(\xi_3) = \sum_{k, l, h=0}^{\infty} EF_1^{(k+l)}(\xi_1) EF_2^{(k+h)}(\xi_2) EF_3^{(h+l)}(\xi_3) \frac{G_{12}^k}{k!} \frac{G_{13}^l}{l!} \frac{G_{23}^h}{h!}.$$

2. As there is no restriction on the components of ξ in the Theorem (if $F \in \mathcal{C}'_n$) they can be equal, $E\xi=0$, $D^2\xi=\sigma^2$ that is why then

$$EF(\xi, \xi, \dots, \xi) = \sum_{\alpha \geq 0} \frac{(\sigma^2)^{|\alpha|}}{\prod_{i < j} k_{ij}!} ED^\alpha F(\xi_1^*, \dots, \xi_n^*).$$

In this case if $F(x) = \prod_i F_i(x_i)$ then

$$E \prod_i F_i(\xi) = \sum_{\alpha \geq 0} \frac{(\sigma^2)^{|\alpha|}}{\prod_{i < j} k_{ij}!} \prod_i EF_i^{(\alpha_i)}(\xi).$$

3. The polynomial

$$P(x) = \sum_{0 \leq |\alpha| \leq N} C_\alpha x^\alpha$$

belongs to $\mathcal{C}'_n$, its Fourier transform can be easily calculated

$$\hat{P} = \left(\sum_{0 \leq |\alpha| \leq N} C_\alpha (-D_\alpha) \right) \delta$$

where δ is the Dirac measure on R^n . Thus

$$EP(\xi) = (2\pi)^{n/2} \hat{P}(\varphi_\xi) = (2\pi)^{n/2} \sum_{0 \leq |\alpha| \leq N} C_\alpha (-D_\alpha \varphi_\xi)(0)$$

or

$$EP(\xi) = \sum_{0 \leq |\alpha| \leq N} ED^\alpha P(\xi^*) \prod_{i < j} \frac{G_{ij}^{k_{ij}}}{k_{ij}!}.$$

4. If $z(t)$, $t \in R_1$ is a stationary Gaussian process with mean 0 and covariance function $\Gamma(\tau)$, $\hat{F} \in \mathcal{C}_1'$ then $Y(t) = F(z(t))$ is stationary too with mean $EF(z(0))$ and covariance function

$$\begin{aligned} \Gamma_{YY}(\tau) &= \sum_{k=1}^{\infty} E^2 F^{(k)}(z(0)) \frac{\Gamma_{\tau}^k}{k!} \\ \Gamma_{Yz}(\tau) &= \Gamma(\tau) EF'(z(0)). \end{aligned}$$

From this follows that the normed cross-covariance function $\gamma_{Yz}(\tau)$ does not depend on F i.e.

$$\gamma_{Yz}(\tau) = \frac{\Gamma_{Yz}(\tau)}{\Gamma_{Yz}(0)} = \frac{\Gamma(\tau)}{\Gamma(0)}.$$

5. If $z(t)$, $t \in R_1$ is a vector valued stationary Gaussian process with mean 0 and covariance matrix $\Gamma(\tau) = (\Gamma_{ij}(\tau))$, $\hat{F} \in \mathcal{C}_n'$ then $Y(t) = F(z(t))$ is stationary too with mean $EF(z(0))$ and covariance function

$$\Gamma_{YY}(\tau) = \sum_{1 \leq |\alpha_1|, |\alpha_2|} \prod_{i < j} \frac{\Gamma_{ij}^{l_{ij} + l'_{ij}}(0)}{l_{ij}!} \prod_{i, j} \frac{\Gamma_{ij}^{k_{ij}}}{k_{ij}!} ED^{\alpha_1} F(z(0)) ED^{\alpha_2} F(z(0))$$

where

$$\begin{aligned} \alpha_{1i} &= \sum_{j \neq i} (l_{ij} + l_{ji}) + \sum_j (k_{ij} + k_{ji}) \\ \alpha_{2i} &= \sum_{j \neq i} (l'_{ij} + l'_{ji}) + \sum_j (k_{ij} + k_{ji}). \end{aligned}$$

The cross-covariance function and the normed one are

$$\Gamma_{Yz_k}(\tau) = \Gamma_{kk}(\tau) \sum_{\alpha} \prod_{i < j} \frac{\Gamma_{ij}^{l_{ij}}}{l_{ij}!} ED^{\alpha} F(z(0))$$

where

$$\alpha_k = 1 + \sum_{j \neq k} l_{kj} + l_{jk}, \quad \alpha_i = \sum_{j \neq i} (l_{ij} + l_{ji}), \quad i \neq k.$$

$$\gamma_{Yz_k}(\tau) = \frac{\Gamma_{Yz_k}(\tau)}{\Gamma_{Yz_k}(0)} = \frac{\Gamma_{kk}(\tau)}{\Gamma_{kk}(0)}.$$

3. *The Generalized Appel—Wick polynomial system.* The Appel polynomials play an important role in the examination of nonlinear stochastic systems such as Uryson and Zadeh models [4], [5]. The one-variable Appel polynomials of degree n which differ from the Hermite polynomials only in a constant factor were mentioned by BONNET [1], CAMPBELL [2] and SHUTTERLY [5]. The n -dimensional Hermite polynomials can be defined by the same way as the standard Hermite polynomials i.e.

differentiating the joint density function (GRAD [3]) of n independent standard Gaussian random variables. This method does not work in the dependent case. We define the n -variable Generalized Appel—Wick (GAW) polynomials of degree n by the following rules:

a) $A_0 = 1$

b) $\frac{\partial}{\partial x_i} A_n(x_1, \dots, x_n) = A_{n-1}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$

c) $EA_n(x_1, \dots, x_n) = 0$, $n = 1, 2, \dots$ where $(x_1, \dots, x_n)$ is a Gaussian random vector with mean 0 and covariance matrix $G = (G_{i,j})$.

One can get the GAW polynomials by the following recursive formulas

$$A_n(x_1, \dots, x_n) = x_k A_{n-1}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) - \sum_{i \neq k} G_{ik} A_{n-2}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{k-1}, x_{k+1}, \dots, x_n).$$

Indeed $A_0 = 1$, $A_1(x_1) = x_1$, $A_2(x_1, x_2) = x_1 x_2 - G_{12}$ so for $n = 1, 2$ b) is valid and by induction from $n-2$ and $n-1$ follows that

$$\frac{\partial}{\partial x_i} A_n(x_1, \dots, x_n) = A_{n-1}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n).$$

Assumption c), fulfills as well

$$\begin{aligned} EA_n(x_1, \dots, x_n) &= Ex_k A_{n-1}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) - \\ &- \sum_{i \neq k} G_{ik} EA_{n-2}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{k-1}, x_{k+1}, \dots, x_n) = \\ &= \sum_{i \neq k} G_{ik} E \frac{\partial}{\partial x_i} A_{n-1}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n) - \\ &- \sum_{i \neq k} G_{ik} EA_{n-2}(x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_{k-1}, x_{k+1}, \dots, x_n) = 0. \end{aligned}$$

Let us consider now the second order moments of the GAW polynomials when the variable $(x, z) = (x_1, \dots, x_n, z_1, \dots, z_m)$ is Gaussian with mean 0 and cov $(x_i, z_j) = G_{x_i z_j}$. Using formula (2.3) we get

$$\begin{aligned} (3.1) \quad EA_n(x_1, \dots, x_n) A_m(z_1, \dots, z_m) &= \sum_{\alpha_1, \alpha_2} ED^{\alpha_1} A_n(x) ED^{\alpha_2} A_m(z) \prod_{i,j} \frac{G_{x_i z_j}^{\alpha_{ij}}}{\alpha_{ij}!} = \\ &= \delta_n^m \sum_{n!}^* \prod_{l=1}^n G_{x_l z_{i_l}}, \end{aligned}$$

where

$$\alpha_{1i} = \sum_j \gamma_{ij}, \quad \alpha_{2i} = \sum_j \gamma_{ji}.$$

From this and the definition of GAW polynomials it follows that the expectation (3.1) is zero except $|\alpha_1| = |\alpha_2| = n$ i.e. $\gamma_{li} = 1$, $l = 1, 2, \dots, n$ where $i_1, \dots, i_n$ is a permutation of numbers $1, 2, \dots, n$, $\gamma_{ij} = 0$ in other cases, so the summation $\sum^*$ has to be extended to all possible permutations $i_1, i_2, \dots, i_n$ of numbers $1, 2, \dots, n$.

These propositions for the GAW polynomials remain true also in the case when the variables are not different. Let us introduce the following notation for the k variable GAW polynomial of degree n ($k \leq n$):

$$A_{n_1, n_2, \dots, n_k}(x_1, \dots, x_k) = A_{\Sigma n_i}(\underbrace{x_1 \dots x_1}_{n_1}, \dots, \underbrace{x_k \dots x_k}_{n_k}).$$

As a special case we get for the Gaussian stationary process $z(t)$ with mean 0 and covariance function $\Gamma(\tau)$ the GAW polynomials are quasi orthogonal i.e.

$$\begin{aligned} EA_{n_1, n_2, \dots, n_k}(z(t_1), \dots, z(t_k)) &= 0 \\ EA_{n_1, n_2, \dots, n_k}(z(t_1), \dots, z(t_k)) A_{m_1, \dots, m_l}(z(s_1), \dots, z(s_l)) &= \\ = \delta_{\Sigma n_i}^{\Sigma m_i} \prod_{i=1}^k n_i! \sum_{j_q^i \geq 0} \binom{m_i}{f_1, \dots, f_k} \prod_{i=1}^k \Gamma_{(s_i - t_i)}^{f_i} \prod_{i=1}^{l-1} \binom{m_i}{j_1^i, \dots, j_k^i} \times \prod_{i=1}^{l-1} \prod_{q=1}^k \Gamma_{(t_q - s_i)}^{j_q^i}. \end{aligned}$$

Where

$$m_i = \sum_{q=1}^k j_q^i, \quad f_i = n_i - \sum_{q=1}^{l-1} j_q^i \geq 0 \quad \text{and} \quad \binom{m}{k_1, \dots, k_l} = \frac{m!}{\prod_i k_i!}.$$

In a special way

$$\begin{aligned} EA_{k,l}(z(t), z(s)) A_{m,n}(z(u), z(w)) &= \\ = \delta_{k+l}^{m+n} k! l! \sum_{j=\max(0, n-m)}^{\min(n, m)} \binom{n}{m-j} \Gamma_{(w-t)}^{n-j} \Gamma_{(w-s)}^{n-m+j} \binom{m}{j} \Gamma_{(u-t)}^j \Gamma_{(u-s)}^{m-j} \end{aligned}$$

and

$$EA_k(z(t)) A_l(z(s)) = \delta_k^l k! \Gamma_{(t-s)}^k.$$

References

- [1] G. BONNET, Transformation des signaux aleatoires, *Ann. Télécom.*, t. 19, n 9—10. (1964), 203—220.
- [2] L. L. CAMPBELL, On a class of polynomials useful in probability calculations, *I.E.E.E. Trans. Inf. Theory*, IT-10, 4 (1964), 255.
- [3] H. GRAD, Note on N -dimensional Hermite Polynomials, *Communications on Pure and Appl. Maths.*, 2 (Dec. 1949), 325—330.
- [4] W. RUDIN, Functional Analysis, *New York*, 1973.
- [5] H. B. SHUTTERLY, General results in mathematical theory of random signals and noise in nonlinear devices, *I.E.E.E. Trans. on Inf. Theory*, IT-9 (1963), 74—84.
- [6] P. VÁRLAKI, G. TERDIK, J. BOKOR, Estimation of the basic nonlinearity characteristics for identification of second order dynamic systems, *Intern. Conf. of Apl. Cont. and Identification, Kopenhagen*, 1983.
- [7] P. VÁRLAKI, G. TERDIK, Identification of Raibman kernels in Zadeh functional series representation of nonlinear systems, *Probl. of Cont. and Inf. Theory*, 1985, № 3.
- [8] K. YOSIDA, Functional Analysis, *Springer Verl.* 1965.
- [9] И. А. Ибрагимов, Ю. А. Розанов, Гауссовские случайные процессы, *Наука, Москва*, 1970.
- [10] П. Варлаки, Оценка степени нелинейности Изв. Акад. Наук СССР, Тех. Киб (1980), И° I, 156—160.

(Received September 12, 1984)