ON ZERO MEMORY NONLINEAR TRANSFORMATIONS OF GAUSSIAN PROCESSES

GARY L. WISE
University of Texas

JOHN B. THOMAS
Princeton University

Summary. This paper treats the second moment properties of a zero memory nonlinearity, given that the input is a stationary Gaussian process. The output autocorrelation function is shown to be expressed conveniently in terms of the input autocorrelation function and a set of coefficients describing the ZNL. Two theorems are proved concerning the output process bandwidth. The first shows that the output bandwidth is generally greater than the input bandwidth. The second gives necessary and sufficient conditions for the output process to be strictly bandlimited.

Introduction. From the viewpoint of second moment theory, the study of linear systems with random inputs is very well developed. However, such is not the case in general for nonlinear systems. In this paper, we consider the class of (time invariant) zero memory nonlinearities (ZNL's); that is, the class of systems where the output at time t is a function only of the input at time t. Notice that this class encompasses many commonly encountered nonlinearities; for example: quantizers, limiters, rectifiers, power law devices, etc. Since many results are known based on the second moment characterization of random processes, we will be interested in the second moment properties of the output of the ZNL.

The input will be taken to be a zero mean, stationary, Gaussian random process X(t). The variance of X(t) will be denoted by $\sigma^2 > 0$, and the autocorrelation function by $R_X(\tau) = E\{X(t)X(t+\tau)\}$. The class of ZNL's will be the class $g(\cdot)$ of all Baire functions [1] such that

$$\int_{-\infty}^{\infty} [g(u)]^2 \exp\left(-\frac{u^2}{2\sigma^2}\right) du < \infty \quad .$$

Thus the output $Y(t) = g[X(t)]$ of the ZNL is a well-defined second-order random process. Notice that this class of ZNL's includes virtually any ZNL of engineering interest. In the next section we discuss the form of the output autocorrelation function. Then, in the final section, we discuss the bandwidth of the output of the ZNL.

This research was supported by the National Science Foundation under Grant GK-24187 and by the U.S. Air Force Office of Scientific Research under Grant AFOSR-71-2101.

The Output Autocorrelation Function. The ZNL g(·) transforms the random process X(t) to Y(t), which will be a stationary, second-order random process. Let $R_y(\tau)$ denote its autocorrelation function; that is,

$$R_Y(\tau) = E\{Y(t)Y(t+\tau)\} = E\{g[X(t)]g[X(t+\tau)]\}.$$

By transforming the random process X(t), the ZNL g(·) induces a transformation on the second moment properties; that is, the output autocorrelation function $R_Y(\tau)$ can be represented as a transformation of $R_X(\tau)$ such that

$$R_Y(\tau) = T_g[R_X(\tau)].$$

An expression for the output autocorrelation function is given by the following:

$$R_Y(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(u)g(v)\ f[u,v,R_X(\tau)]\ du\ dv, \tag{1}$$

where the input bivariate density is

$$f[u,v\ R_X(\tau)] = \frac{1}{2\pi\sigma^2\sqrt{1-\rho_X{}^2(\tau)}} \exp\left[\frac{-[u^2-2uv\rho_X(\tau)+v^2]}{2\sigma^2[1-\rho_X{}^2(\tau)]}\right] \tag{2}$$

and $\rho_X(\tau) = R_X(\tau)/\sigma^2$. The integral in Eq. (1) is difficult to evaluate. Also, using this integral, it is not easy to obtain general statements concerning properties of the output autocorrelation function $R_Y(\tau)$.

As a general approach to the analysis of the output autocorrelation function, we will expand the bivariate density of Eq. (2) in a series of orthonormal functions. To simplify the analysis, the orthonormal functions will be chosen such that the expansion reduces to a diagonal series. In this case the orthonormal functions become the Hermite polynomials, orthonormalized with respect to the univariate density of X(t). The resulting series is a classical expansion known as Mehler's formula [2,3], given by

$$f[u,v,R_X(\tau)] = \frac{1}{2\pi\sigma^2}\exp\left[\frac{-(u^2+v^2)}{2\sigma^2}\right]\sum_{n=0}^{\infty}[\rho_X(\tau)]^n\phi_n(u)\phi_n(v), \tag{3}$$

where

$$\phi_n(u) = \frac{H_n(u/\sigma)}{\sqrt{n!}}.$$

It can be shown (4] that the right side of Eq. (3) is pointwise convergent, as well as convergent in an L_2 sense, to the bivariate density $f[.,.,R_X(\tau)]$. Now consider expanding g(·) in terms of the sequence $\{\phi_n(\tau)\}$. We have

$$g(u) = \sum_{n=0}^{\infty} b_n\ \phi_n(u), \tag{4}$$

where the coefficients are given by

$$b_n = \int_{-\infty}^{\infty} g(u)\,\phi_n(u) \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left[\frac{-u^2}{2\sigma^2}\right] du, \tag{5}$$

and where the convergence is in an L_2 sense.

By using the series expansions in Eqs. (3) and (4), the double integral in Eq. (1) can be expressed as the following sum:

$$R_Y(\tau) = \sum_{n=0}^{\infty} [\rho_X(\tau)]^n\, b_n^2. \tag{6}$$

Since $|\rho_X(\tau)| \le 1$, it follows in a straightforward fashion that the series of Eq. (6) converges uniformly in τ. We see that the transformation $T_g(\cdot)$ is given by Eq. (6).

If the infinite sum in Eq. (6) is truncated, the truncation error can be uniformly bounded in a straightforward fashion. Also, the truncated sum is nonnegative definite. Another advantage of Eq. (6) over Eq. (1) is that general properties of the output autocorrelation function $R_Y(\tau)$ can be obtained.

As a simple example of the series expansion, consider the “expander” given by

$$g(u) = u^3 + u.$$

In this case, Eq. (4) simplifies to

$$g(u) = (3\sigma^3 + \sigma)\,\phi_1(u) + \sqrt{6}\,\sigma^3\,\phi_3(u).$$

Therefore, from Eq. (6), the autocorrelation function of the output

$$Y(t) = [X(t)]^3 + X(t)$$

is given by

$$R_Y(\tau) = (3\sigma^2 + 1)^2 R_X(\tau) + 6[R_X(\tau)]^3.$$

From Eq. (6) one can readily establish the well-known fact that the modulus of the correlation coefficient of the output of the ZNL is upper bounded by the modulus of the correlation coefficient of the Gaussian input process. One can also see that, for $R_X(\tau)$ positive, the transformation of the autocorrelation function is convex and monotonic.

Consider the problem of determining the input autocorrelation function $R_X(\tau)$ from a knowledge of the output autocorrelation function $R_Y(\tau)$. Obviously, if $g(\cdot)$ is an odd function, this can always be done, that is, we see from Eq. (5) that $b_{2n} = 0$ and thus the transformation in Eq. (6) is monotonic.

Suppose that the autocorrelation function $R_X(\tau)$ is an analytic function. In the neighborhood of zero in which $R_Y(\tau) > b_0^2$, it is also true that $R_X(\tau) > 0$. Thus, for τ in this region, $R_X(\tau)$ can be determined from a knowledge of the output autocorrelation function $R_Y(\tau)$. Then, since $R_X(\tau)$ is analytic, it is uniquely determined for all τ by the process of analytic continuation. Therfore, if the input autocorrelation function is analytic, it is uniquely determined by the output autocorrelation function. Notice that if X(t) is bandlimited, then $R_X(\tau)$ is analytic.

Output Bandwidth. In this section we consider how the ZNL affects the bandwidth of X(t). We will avoid the trivial case by assuming that the ZNL is not characterized by a constant function. This is equivalent to assuming that, in Eq. (5), for $n \geq 1$, at least one b_n is nonzero. First we treat the second moment bandwidth, and then we consider strict band-limitedness of the output Y(t).

Since we are not interested in the d-c component of the output, we will assume that the output random process has been centered about its mean. This is equivalent to working with the autocovariance function of the output. It is readily seen that this autocovariance function is given by

$$\sum_{n=1}^{\infty} b_n^2 [\rho_X(\tau)]^n$$

and the correlation coefficient $\rho_Y(\tau)$ by

$$\rho_Y(\tau) = \frac{\sum_{n=1}^{\infty} b_n^2 [\rho_X(\tau)]^n}{\sum_{n=1}^{\infty} b_n^2} \tag{7}$$

For a zero mean, stationary, second-order random process with a normalized spectral density given by $S(\omega)$, the second moment bandwidth of this random process, if it exists, is

$$\int_{-\infty}^{\infty} \omega^2 S(\omega)\, d\omega \;.$$

It is easily shown that

$$\int_{-\infty}^{\infty} \omega^2 S(\omega)\, d\omega = -\rho''(0),$$

where $\rho(\tau)$ is the correlation coefficient of the random process.

Assume X(t) has a finite second moment bandwidth given by $BW[X(t)] = -\rho_X''(0)$ and that the second moment bandwidth of Y(t) is $BW[y(t)] = -\rho_Y''(0)$. Let the function r(u) be defined on [-1, 1] by

$$r(u) = \sum_{n=1}^{\infty} b_n^2 \, u^n \Big/ \sum_{n=1}^{\infty} b_n^2 .$$

Then we have

$$\rho_Y(\tau) = r[\rho_X(\tau)].$$

On taking the second derivative, we get

$$\rho_Y''(\tau) = r'[\rho_X(\tau)][\rho_X'(\tau)]^2 + r'[\rho_X(\tau)]\rho_X''(\tau).$$

Recall that $\rho_X'(0) = 0$, and we have

$$\rho_Y''(0) = r'(1) \; \rho_X''(0),$$

or equivalently ,

$$BW[Y(t)] = r'(1) \; BW[X(t)].$$

Notice that

$$r'(1) = \frac{\sum_{n=1}^{\infty} n \, b_n^2}{\sum_{n=1}^{\infty} b_n^2} \geq 1. \qquad (8)$$

Therefore, we conclude that $BW[Y(t)] \geq BW[X(t)]$. 2

We see from Eq . (8) that r'(1) = 1 if and only if $b_n^2 = 0$ for n 2 $\geq$ 2. This is equivalent to the ZNL g(·) being of the form g(u) = au + b. Thus r'(1) = 1 if and only if the graph of g(u) is a straight line, or, equivalently, if g(·) is affine. This result is summarized in the following theorem.

Theorem 1: Let X(t) be a zero mean, stationary, Gaussian random process with a finite second moment bandwidth. Then g[X(t)] has a sencond moment bandwidth that is greater than or equal to that of X(t). Equality holds if and only if g(·) is affine.

Consider the "expander"

$$g(u) = u^3 + u,$$

given in the earlier example. In this case

$$r(u) = u\left[\frac{9\sigma^4 + 6\sigma^2 + 1}{15\sigma^4 + 6\sigma^2 + 1}\right] + u^3\left[\frac{6\sigma^4}{15\sigma^4 + 6\sigma^2 + 1}\right],$$

and

$$r'(1) = \frac{27\sigma^4 + 6\sigma^2 + 1}{15\sigma^4 + 6\sigma^2 + 1}.$$

The quantity r'(1) is the ratio of the second moment bandwidth of $[X(t)]^3 + X(t)$ to the second moment bandwidth of X(t).

We see that this ratio increases monotonically as the variance of X(t) increases and that it ranges in the interval (1, 1-8).

Now we consider a different aspect of the output bandwidth, namely the strict band-limitedness of the output Y(t). Assume that the spectral density of X(t) exists. Since we are considering strict bandlimitedness, the appropriate concept will be the Lebesgue measure of the support of the spectral density. A random process is strictly bandlimited if the spectral density has bounded support. Consider the normalized spectral density of Y(t), given by the Fourier transform of $\rho_Y(\tau)$. It follows from the Lebesgue Convergence Theorem [5] that the series in Eq. (7) may be transformed term by term. The Fourier transform of $[\rho_X(\tau)]^n$ is given by $(\frac{1}{2\pi})^{n-1}$ times the (n-1)-fold convolution of the normalized spectral density of X(t). Let M be the Lebesgue measure of the support of the Fourier transform of $\rho_X(T)$. Then it is not hard to show that the Lebesgue measure of the support of the Fourier transform of $[\rho_X(\tau)]^n$ is greater than or equal to nM. Thus it follows that, if the series in Eq. (7) does not truncate, then the support of the spectral density of Y(t) will be unbounded. If the series does truncate, the support of the spectral density of Y(t) will be bounded only if the support of the Fourier transform of $\rho_X(\tau)$ is bounded. Thus we see that, if the series in Eq. (7) truncates, and if the spectral density of X(t) has bounded support, then the spectral density of Y(t) will have bounded support. Also, we see that if the series does not truncate, or if the spectral density of X(t) does not have bounded support, then, in either case, Y(t) will not be strictly bandlimited. Notice that truncation of the series in Eq. (7) is equivalent to the ZNL being characterized by a polynomial. This result is summarized in the following theorem.

Theorem 2: Let X(t) be a zero mean, stationary, Gaussian random process. Then g[X(t)] is strictly bandlimited if and only if:

a) X(t) is strictly bandlimited, and
b) g(·) is a polynomial.

Notice that many common ZNL's are not polynomials. In particular, it follows from Theorem 2 that if X(t) is passed through any type of limiter, then the output cannot be strictly bandlimited. As a general statement, it follows from Theorems 1 and 2 that the loss of normality of X(t), caused by a ZNL, is always accompanied by a spreading of the spectrum.

References

1. W. Feller, An Introduction to Probability Theory and its Applications, Vol. II, New York: John Wiley, 1971, pp. 104-106.

2. J.F. Barrett and D.G. Lampard, "An Expansion for Some Second-order Probability Distributions and Its Application to Noise Problems", IRE Trans. Information Theory, Vol. IT-1, pp. 10-15, March 1955.

3. J.B. Thomas, An Introduction to Statistical Communication Theory, New York: John Wiley, 1969, p. 314.

4. R. Leipnik, "Integral Equations, Biorthonormal Expansions, and Noise", SIAM J. Appl. Math., Vol. 7, pp. 6-30, March 1959.

5. H.L. Royden, Real Analysis, London: Macmillan, 1968, p. 88.