

A Useful Theorem for Nonlinear Devices Having Gaussian Inputs*

ROBERT PRICE†

Summary—If and only if the inputs to a set of nonlinear, zero-memory devices are variates drawn from a Gaussian random process, a useful general relationship may be found between certain input and output statistics of the set. This relationship equates partial derivatives of the (high-order) output correlation coefficient taken with respect to the input correlation coefficients, to the output correlation coefficient of a new set of nonlinear devices bearing a simple derivative relation to the original set. Application is made to the interesting special cases of conventional cross-correlation and autocorrelation functions, and Busgang's theorem is easily proved. As examples, the output autocorrelation functions are simply obtained for a hard limiter, linear detector, clipper, and smooth limiter.

IN THE COURSE of investigating the asymptotic frequency behavior of power spectra resulting from the passage of noise through zero-memory nonlinear devices, an interesting, unique property of Gaussian processes has been encountered, which does not appear to have been previously reported.

STATEMENT OF THE THEOREM

Assume x_1, x_2, \dots, x_n to be random variables from a Gaussian process whose n th order joint probability density is given by:¹

$$p(x_1, x_2, \dots, x_n) = (2\pi)^{-n/2} |M_n|^{-1/2} \cdot \exp \left\{ -\frac{1}{2} \sum_{r=1}^n \sum_{s=1}^n \frac{M_{rs}}{|M_n|} (x_r - \bar{x}_r)(x_s - \bar{x}_s) \right\} \quad (1)$$

where $|M_n|$ is the determinant of $M_n = [\rho_{rs}]$ and $\rho_{rs} = \frac{\overline{x_r x_s} - \bar{x}_r \bar{x}_s}{\sigma_r \sigma_s} = \rho_{sr}$ is the correlation coefficient of x_r and x_s . The means of x_r and x_s are \bar{x}_r and \bar{x}_s , respectively. M_{rs} is the cofactor of ρ_{sr} in M_n .

Let there be n zero-memory nonlinear devices (linearity of course being included as a special case) specified by the input-output relationship $f_i(x)$, $i = 1, 2, \dots, n$. Let each x_i be the single input to a corresponding $f_i(x)$, and designate the n th-order correlation coefficient of the outputs as:

$$R = \overline{\prod_{i=1}^n f_i(x_i)} \quad (2)$$

where the bar denotes the average taken over all x_i . Then, with weak restrictions on the $f_i(x)$, we have the following theorem for the partial derivatives of R with respect to the input correlation coefficients:

* Manuscript received by the PGIT, January 3, 1958. The research in this paper was supported jointly by the Army, Navy, and Air Force under contract with Mass. Inst. Tech.

† Lincoln Lab., M.I.T., Lexington, Mass.

¹ H. Cramér, "Mathematical Methods of Statistics," Princeton University Press, Princeton, N. J., sec. 24.2; 1946.

$$\frac{\partial^k R}{\prod_{m=1}^N (\partial \rho_{r_m s_m})^{k_m}} = \left(\frac{1}{2} \right)^{\sum_{m=1}^N k_m \delta_{r_m s_m}} \left[\prod_{i=1}^n f_i^{(\sum_{m=1}^N \epsilon_{im} k_m)}(x_i) \right] \quad (3)$$

where r_m and s_m , $m = 1, 2, \dots, N$, are integers lying between 1 and n , inclusive, and are not necessarily distinct. The k_m are positive integers, with $k = \sum_{m=1}^N k_m$. ϵ_{im} is the number of times i appears in (r_m, s_m) . $\delta_{r_m s_m}$ is the Kronecker δ function, $\delta_{r_m s_m} = 1$ for $r_m = s_m$, 0 for $r_m \neq s_m$. The symbol $f_i^{(q)}(x_i)$ denotes the q th derivative of $f_i(x)$, taken at x_i .

Furthermore, not only is the above theorem true for inputs having an n th-order joint Gaussian distribution, but it holds true *only* for such inputs if the $f_i(x)$ are allowed to be of general form.

Proof

We now prove that in order for (3) to be satisfied it is both sufficient and necessary that the x_i have the joint probability density given by (1). Assume that each $f_i(x)$ can be represented by the sum of two Laplace transforms,²

$$f_i(x) = \frac{1}{2\pi j} \int_{C_{i+}} h_{i+}(u) e^{jux} du + \frac{1}{2\pi j} \int_{C_{i-}} h_{i-}(u) e^{jux} du \quad (4)$$

where

$$\left. \begin{aligned} h_{i+}(u) &= \int_0^{\infty} f_i(x) e^{-jux} dx \\ h_{i-}(u) &= \int_{-\infty}^0 f_i(x) e^{-jux} dx \end{aligned} \right\} \quad (5)$$

and the C_{i+} and C_{i-} are appropriate contours. Without assuming any particular form for $p(x_1, x_2, \dots, x_n)$ for the present,

$$R = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \cdot \prod_{i=1}^n f_i(x_i) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \quad (6)$$

Substituting (4) in (6) and inverting the order of integration, following Rice's characteristic function method,³

² D. V. Widder, "The Laplace Transform," Princeton University Press, Princeton, N. J., ch. 6; 1946.

³ S. O. Rice, "Mathematical analysis of random noise," *Bell Sys. Tech. J.*, vol. 23, pp. 282-332, July, 1944; and vol. 24, pp. 46-156; January, 1945. See sec. 4.8.

$$R = \frac{1}{(2\pi j)^n} \sum' \int_{C_{1*}} \int_{C_{2*}} \cdots \int_{C_{n*}} \cdot \prod_{i=1}^n h_{i*}(u_i) \theta(u_1, u_2, \dots, u_n) du_1 du_2 \cdots du_n \quad (7)$$

where \sum' denotes a summation over all possible \pm combinations and $\theta(u_1, u_2, \dots, u_n)$ is the n th-order characteristic function:

$$\theta(u_1, u_2, \dots, u_n) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} p(x_1, x_2, \dots, x_n) \cdot \exp\left(j \sum_{i=1}^n u_i x_i\right) dx_1 dx_2 \cdots dx_n \quad (8)$$

with $j = \sqrt{-1}$.

We find a necessary condition for (3) to be satisfied by setting $N = 1 = k = k_1$. The partial derivative of the left-hand side of (3) is taken on θ in the integrand of (7), and the derivatives of the right-hand side are taken using (4). Thus the necessary condition:

$$\sum' \int_{C_{1*}} \int_{C_{2*}} \cdots \int_{C_{n*}} \prod_{i=1}^n h_{i*}(u_i) \left\{ \frac{\partial \theta(u_1, u_2, \dots, u_n)}{\partial \rho_{r_1 s_1}} + \left(\frac{1}{2}\right)^{\delta_{r_1 s_1}} u_{r_1} u_{s_1} \theta(u_1, u_2, \dots, u_n) \right\} du_1 du_2 \cdots du_n = 0 \quad (9)$$

is obtained. The term in braces must be zero in order to satisfy (9) for arbitrary $f_i(x)$ and hence $h_{i*}(u)$. Integrating the resulting equation for all (r_1, s_1) (but taking into account that $\rho_{rs} = \rho_{sr}$),

$$\log \theta(u_1, u_2, \dots, u_n) = -\frac{1}{2} \sum_{r=1}^n \sum_{s=1}^n \rho_{rs} u_r u_s + g(u_1, u_2, \dots, u_n) \quad (10)$$

where g is some function which must now be found.

Let $\rho_{rs} = 1$ for all (r, s) . Then all the x_i are completely correlated, and $p(x_1, x_2, \dots, x_n)$ can be written:

$$p(x_1, x_2, \dots, x_n) = p(x_1) \prod_{i=2}^n \delta(x_i - x_1 + \bar{x}_1 - \bar{x}_i) \quad (11)$$

where $\delta(x)$ is the Dirac δ function. Substituting (11) in (8), θ is of the form:

$$\theta(u_1, u_2, \dots, u_n) = \exp\left(j \sum_{i=1}^n u_i \bar{x}_i\right) g_1\left(\sum_{i=1}^n u_i\right), \quad \text{for all } \rho_{rs} = 1 \quad (12)$$

where

$$g_1(u) = \int_{-\infty}^{+\infty} p_1(x_1 - \bar{x}_1) e^{ju(x_1 - \bar{x}_1)} d(x_1 - \bar{x}_1). \quad (13)$$

Similarly, when $\rho_{11} = 1$, $\rho_{1r} = \rho_{r1} = -1$ for all $r \neq 1$, and $\rho_{rs} = 1$ for all r or $s \neq 1$, then x_2, x_3, \dots, x_n are completely correlated with $(-x_1)$ and we obtain:

$$\theta(u_1, u_2, \dots, u_n) = \exp\left(j \sum_{i=1}^n u_i \bar{x}_i\right) g_1\left(u_1 - \sum_{i=2}^n u_i\right), \quad \text{for } \rho_{11} = 1, \rho_{1r} = \rho_{r1} = -1 \text{ for all } r \neq 1, \text{ and } \rho_{rs} = 1 \text{ for all } r, s \neq 1. \quad (14)$$

Substituting (12) in (10), we find:

$$g(u_1, u_2, \dots, u_n) = j \sum_{i=1}^n u_i \bar{x}_i + g_2\left(\sum_{i=1}^n u_i\right) \quad (15)$$

where $g_2(u) = \log g_1(u) + u^2/2$. On the other hand, substituting (14) in (10) yields

$$g(u_1, u_2, \dots, u_n) = j \sum_{i=1}^n u_i \bar{x}_i + g_2\left(2u_1 - \sum_{i=1}^n u_i\right). \quad (16)$$

Since u_1 and $\sum_{i=1}^n u_i$ may be considered as independent variables, the only solution which renders (15) and (16) compatible is $g_2(u) = K$, a constant. Thus, finally, we have from (10) and (15) the necessary condition:

$$\theta(u_1, u_2, \dots, u_n) = \exp\left[-\frac{1}{2} \sum_{r=1}^n \sum_{s=1}^n \rho_{rs} u_r u_s + j \sum_{i=1}^n u_i \bar{x}_i + K\right]. \quad (17)$$

This is recognized to be the characteristic function of the n -dimensional Gaussian distribution⁴ of (1) ($K = 0$ for proper normalization).

It is now a simple matter to prove the sufficiency of (17), and hence (1), for satisfying (3). Using (17) in (7), and remembering that $\rho_{rs} = \rho_{sr}$,

$$\frac{(-1)^k \partial k_R}{\prod_{m=1}^N (\partial \rho_{r_m s_m})^{k_m}} = \left(\frac{1}{2}\right)^{\sum_{m=1}^N k_m \delta_{r_m s_m}} \sum' \int_{C_{1*}} \int_{C_{2*}} \cdots \int_{C_{n*}} \cdot \prod_{i=1}^n u_i^{\sum_{m=1}^N \epsilon_{i m k_m}} h_{i*}(u_i) \theta(u_1, u_2, \dots, u_n) du_1 du_2 \cdots du_n. \quad (18)$$

By analogy to (6) and (7), and differentiating (4) with respect to x , the right side of (18) is seen to be equal to

$$(-1)^k \left(\frac{1}{2}\right)^{\sum_{m=1}^N k_m \delta_{r_m s_m}} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \cdot \prod_{i=1}^n f_i^{\left(\sum_{m=1}^N \epsilon_{i m k_m}\right)}(x_i) p(x_1, x_2, \dots, x_n) dx_1 dx_2 \cdots dx_n \quad (19)$$

thus yielding (3).

A SPECIAL CASE AND ITS APPLICATIONS

Consider the familiar situation where $n = 2$, and let ρ denote the crosscorrelation coefficient of x_1 and x_2 . Then (3) yields

$$\frac{\partial^k R}{\partial \rho^k} = \overline{f_1^{(k)}(x_1) f_2^{(k)}(x_2)}. \quad (20)$$

Suppose that x_1 and x_2 are values of a stationary Gaussian time series $x(t)$ whose autocorrelation function is $\rho(\tau)$. x_1 is taken at time t and x_2 at time $(t + \tau)$. $R(\tau)$ will denote the crosscorrelation function between the outputs

⁴ Cramér, *op. cit.*, sec. 24.1.

of two zero-memory nonlinear devices whose inputs are $x(t)$ and $x(t + \tau)$, respectively. We find easily

$$R(\tau) = \rho(\tau) \int_{-\infty}^{+\infty} \frac{x f_2(x) e^{-x^2/2}}{\sqrt{2\pi}} dx \quad (25)$$

$$\frac{\partial^k R(\tau)}{\partial \rho(\tau)^k} = \overline{f_1^{(k)}[x(t)] f_2^{(k)}[x(t + \tau)]} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{f_1^{(k)}(x_1) f_2^{(k)}(x_2) \exp\{-[x_1^2 + x_2^2 - 2\rho(\tau)x_1x_2]/2[1 - \rho^2(\tau)]\}}{2\pi\sqrt{1 - \rho^2(\tau)}} dx_1 dx_2. \quad (21)$$

Eq. (21) is particularly simple when the $f_i(x)$ are piecewise-polynomial functions and k is sufficiently high. Then the $f_i^{(k)}(x)$ consist entirely of δ functions of various orders and the integral can be easily evaluated.

It is often of interest to obtain the derivatives of a crosscorrelation function with respect to τ . It is convenient to break down such τ derivatives into a series of products of derivatives of $R(\tau)$ with respect to $\rho(\tau)$, and $\rho(\tau)$ with respect to τ , using

$$\frac{dR(\tau)}{d\tau} = \frac{\partial R(\tau)}{\partial \rho(\tau)} \frac{d\rho(\tau)}{d\tau}. \quad (22)$$

This enables the nonlinear devices to be treated independently of the shape of the input correlation function $\rho(\tau)$, using (21). Similarly, the derivatives of $\rho(\tau)$ with respect to τ do not involve the $f_i(x)$.

As an example, Cohen⁶ shows that in general, for autocorrelation functions $R(\tau)$, the limiting behavior of the corresponding power spectrum $\Phi(\omega)$ is given by:

$$\lim_{\omega \rightarrow \infty} \omega^2 \Phi(\omega) = -\frac{1}{\pi} \left. \frac{dR(\tau)}{d\tau} \right|_{\tau=0+} \quad (23)$$

$$\lim_{\omega \rightarrow \infty} \omega^2 \left[\omega^2 \Phi(\omega) + \frac{1}{\pi} \left. \frac{dR(\tau)}{d\tau} \right|_{\tau=0+} \right] = -\frac{1}{\pi} \left. \frac{d^3 R(\tau)}{d\tau^3} \right|_{\tau=0+}$$

and so on, where the derivatives are with respect to τ .

Another application of (20) is in deriving Bussgang's interesting result⁷ that the crosscorrelation function between the input and the output of a nonlinear device driven by Gaussian noise has the same shape as the input autocorrelation function. In this case $f_1(x) = x$ and $f_2(x)$ is arbitrary. Then $f_1'(x)$ is unity, and all higher derivatives of $f_1(x)$ are zero. Putting this into (21) and evaluating the integral,

$$\frac{\partial^k R(\tau)}{\partial \rho(\tau)^k} = \begin{cases} \int_{-\infty}^{+\infty} \frac{f_2'(x) \exp(-x^2/2)}{\sqrt{2\pi}} dx & k = 1 \\ 0 & k > 1. \end{cases} \quad (24)$$

⁶ I. N. Amiantov and V. I. Tikhonov, "The effect of normal fluctuations on typical nonlinear elements," *Bull. Acad. Sci. USSR*, pp. 33-42; April, 1956. Here, an autocorrelation case [$f_1(x) = f_2(x)$] is studied by expanding the second-order joint Gaussian probability density of $x(t)$ and $x(t + \tau)$ in powers of $\rho(\tau)$, using Mehler's formula; see Bateman Manuscript Project, "Higher Transcendental Functions," McGraw-Hill Book Co., Inc., New York, N. Y., vol. 2, p. 194, (22); 1953. They then integrate by parts to obtain the output correlation function in a similar series but do not recognize the simple form (21) for this series. Using this method it is not required that $f_i(x)$ be Laplace transformable, rather than our proof.

⁷ R. Cohen, "Some Analytical and Practical Aspects of Wiener's Theory of Prediction," Res. Lab. of Electronics, M.I.T., Cambridge, Mass., Tech. Rep. No. 69, ch. 4, sec. 2; June 2, 1948.

⁸ J. J. Bussgang, "Crosscorrelation Functions of Amplitude-Distorted Gaussian Signals," Res. Lab. of Electronics, M.I.T., Cambridge, Mass., Tech. Rep. 216, sec. 3; March 26, 1952.

thus yielding Bussgang's result. Unlike Bussgang's theorem, (20) cannot be generalized to hold for probability distributions other than Gaussian.⁸⁻¹⁰

SOME SIMPLE AUTOCORRELATION EXAMPLES [FOR $\overline{x(t)} = 0, \overline{x^2(t)} = 1$]

Hard Limiter

Van Vleck's well-known result on the autocorrelation function of the output of a hard limiter¹¹ can be derived very simply, using (21). If

$$f_1(x) = f_2(x) = \begin{cases} 1; & x \geq 0 \\ -1; & x < 0 \end{cases} \quad (26)$$

then $f_1^{(1)}(x)$ and $f_2^{(1)}(x)$ are first-order δ functions of area 2, at $x = 0$.

Substituting in (21) and integrating,

$$\frac{\partial R(\tau)}{\partial \rho(\tau)} = \frac{2}{\pi\sqrt{1 - \rho^2(\tau)}}. \quad (27)$$

When $\rho(\tau) = 0, R(\tau) = 0$. Thus

$$R(\tau) = \frac{2}{\pi} \int_0^{\rho(\tau)} \frac{d\rho(\tau)}{\sqrt{1 - \rho^2(\tau)}} = \frac{2}{\pi} \sin^{-1} [\rho(\tau)] \quad (28)$$

which is Van Vleck's result.

Linear Detector

Similarly, the autocorrelation function of the output of a linear detector can be easily found. If

$$f_1(x) = f_2(x) = \begin{cases} x; & x \geq 0 \\ 0; & x < 0 \end{cases} \quad (29)$$

then $f_1^{(2)}(x)$ and $f_2^{(2)}(x)$ are first-order δ functions of area unity at $x = 0$. Substituting in (21) and integrating:

$$\frac{\partial^2 R(\tau)}{\partial \rho(\tau)^2} = \frac{1}{2\pi\sqrt{1 - \rho^2(\tau)}}. \quad (30)$$

Doubly-integrating (30) with the boundary conditions:

⁸ J. F. Barret and D. G. Lampard, "An expansion for some second-order probability distributions and its application to noise problems," *IRE TRANS. ON INFORMATION THEORY*, vol. IT-1, pp. 10-15; March, 1955.

⁹ J. L. Brown, Jr., "On a cross-correlation property for stationary random processes," *IRE TRANS. ON INFORMATION THEORY*, vol. IT-3, pp. 28-31; March, 1957.

¹⁰ A. H. Nuttall, "Invariance of Correlation Functions under Nonlinear Transformations," Res. Lab. of Electronics, M.I.T., Cambridge, Mass., Quart. Progress Rep., p. 63; October 15, 1957.

¹¹ J. L. Lawson and G. E. Uhlenbeck, "Threshold Signals," McGraw-Hill Book Co., Inc., New York, N. Y., p. 58; 1950.

$$\left. \begin{aligned} \frac{\partial R(\tau)}{\partial \rho(\tau)} &= \left[\int_0^\infty f_1^{(1)}(x) \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \right]^2 = \frac{1}{4} \\ R(\tau) &= \left[\int_0^\infty \frac{x e^{-x^2/2}}{\sqrt{2\pi}} dx \right]^2 = \frac{1}{2\pi} \end{aligned} \right\} \text{for } \rho(\tau) = 0 \quad (31)$$

we obtain:

$$\begin{aligned} R(\tau) &= \int_0^{\rho(\tau)} \left[\frac{1}{4} + \int_0^y \frac{dt}{2\pi \sqrt{1-t^2}} \right] dy + \frac{1}{2\pi} \\ &= \frac{1}{2\pi} \left\{ \rho(\tau) \cos^{-1} [\rho(\tau)] + \sqrt{1-\rho^2(\tau)} \right\} \end{aligned} \quad (32)$$

which is in agreement with Rice's result.¹²

Clipper

The relations derived independently by Robin¹³ and Laning and Battin¹⁴ for the autocorrelation function of the output of a clipper may also be found by this method. With a clipper characteristic:

$$f_1(x) = f_2(x) = \begin{cases} l; & l \leq x \\ x; & -l \leq x \leq l \\ -l; & x \leq -l \end{cases} \quad (33)$$

and $f_1^{(2)}(x)$ and $f_2^{(2)}(x)$ each are a pair of first-order δ functions at $x = -l$ and $x = l$, with areas 1 and -1 , respectively. Substituting in (21) and integrating,

$$\frac{\partial^2 R(\tau)}{\partial \rho(\tau)^2} = \frac{\exp \left[-\frac{l^2}{1+\rho(\tau)} \right] - \exp \left[-\frac{l^2}{1-\rho(\tau)} \right]}{\pi \sqrt{1-\rho^2(\tau)}} \quad (34)$$

which is Robin's result, for input noise of unit variance.

Smooth Limiter

Finally, Baum's recent interesting result¹⁵ for the

¹² Rice, *op. cit.*, eq. (4.7-5).
¹³ L. Robin, "The autocorrelation function and power spectrum of clipped thermal noise. Filtering of simple periodic signals in this noise," *Ann. Telecomm.*, vol. 7, pp. 375-387; September, 1952.

¹⁴ J. H. Laning, Jr. and R. H. Battin, "Random Processes in Automatic Control," McGraw-Hill Book Co., Inc., New York, N. Y., p. 362, eq. (B-8); 1956.

¹⁵ R. F. Baum, "The correlation function of smoothly limited Gaussian noise," *IRE TRANS. ON INFORMATION THEORY*, vol. IT-3, pp. 193-197; September, 1957.

autocorrelation function of the output of a device having an error-function characteristic will be derived. With

$$f_1(x) = f_2(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2l^2} dt \quad (35)$$

we have

$$f_1^{(1)}(x) = f_2^{(1)}(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2l^2}. \quad (36)$$

Substituting in (21):

$$\begin{aligned} \frac{\partial R(\tau)}{\partial \rho(\tau)} &= \frac{1}{2\pi} \sqrt{\frac{\rho_1^2 - \rho_2^2}{1 - \rho^2(\tau)}} \\ &\cdot \left\{ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{\exp \left[\frac{-\rho_1 x_1^2 + \rho_1 x_2^2 - 2\rho_2 x_1 x_2}{2(\rho_1^2 - \rho_2^2)} \right]}{2\pi \sqrt{\rho_1^2 - \rho_2^2}} dx_1 dx_2 \right\} \end{aligned} \quad (37)$$

where

$$\left. \begin{aligned} \rho_1 &= \frac{\{l^{-2}[1 - \rho^2(\tau)] + 1\} [1 - \rho^2(\tau)]}{\{l^{-2}l^2 - \rho^2(\tau) + 1\}^2 - \rho^2(\tau)} \\ \rho_2 &= \frac{\rho(\tau)[1 - \rho^2(\tau)]}{\{l^{-2}[1 - \rho^2(\tau)] + 1\}^2 - \rho^2(\tau)} \end{aligned} \right\} \quad (38)$$

The term in braces in (37) must equal unity, since it is the integral of a second-order Gaussian probability density. Thus, from (38),

$$\begin{aligned} \frac{\partial R(\tau)}{\partial \rho(\tau)} &= \frac{1}{2\pi} \sqrt{\frac{\rho_1^2 - \rho_2^2}{1 - \rho^2(\tau)}} \\ &= \frac{1}{2\pi \sqrt{(1 + l^{-2})^2 - l^{-4} \rho^2(\tau)}}. \end{aligned} \quad (39)$$

Integrating and using the condition that when $\rho(\tau) = 0$, $R(\tau) = 0$,

$$\begin{aligned} R(\tau) &= \frac{l^2}{2\pi} \int_0^{\rho(\tau)} \frac{d\rho(\tau)}{\sqrt{(l^2 + 1)^2 - \rho^2(\tau)}} \\ &= \frac{l^2}{2\pi} \sin^{-1} \left[\frac{\rho(\tau)}{1 + l^2} \right] \end{aligned} \quad (40)$$

which is in agreement with Baum's result.

