Communications

A Random Number Generator for Ocean Noise Statistics

Roger J. Webster

Abstract—**Ocean noise is characterized by both temporal and specpg 134tral properties. Time-domain statistics are often non-Gaussian, and spectra are typically non-white. Both attributes can be simulated at the same time by applying a spectral shaping filter to random numbers having a specified kurtosis. A method for generating random numbers with arbitrary kurtosis is presented.**

I. Introduction

Experimental measurements of the undersea acoustic environment show that the distribution of ambient noise, by which is meant the distribution of instantaneous pressure (as recorded by a hydrophone, in units of dynes/cm^2) can be significantly non-Gaussian under certain conditions. This departure from the normal can be observed in the kurtosis β_2, a fourth moment statistical measure (in fact, the ratio of the fourth to the square of the second moments) reflecting the "flatness" of a probability density function (pdf). Relative to the gaussian with a kurtosis value equal to three, high kurtosis distributions are sharply peaked and heavy-tailed, while low kurtosis distributions are flatter and centrally concentrated. High kurtosis is associated with impulsive sources, like the sound of snapping shrimp, and can assume values greatly in excess of three. Low kurtosis is associated with sinusoidal sources, like machinery noise, with a pure tone having a kurtosis value of 3/2. A statistical analysis of ambient noise data from several ocean acoustic environments is presented in [1]. Many different functions have been employed as pdf's to fit the observed frequency distributions, such as the generalized Gaussian, Middleton Class-A, Johnson S-system, ϵ-mixture, and so fourth [2]–[4], [5, Ch. 3]. Any of these pdf's could be used as a basis for a random number generator if a numerical inversion of the cumulative distribution function is practical [6, 3.4.1]. In this note we emphasize formal simplicity in presenting a simple transformation of the uniform distribution as a means of generating random numbers having distributions with arbitrary kurtosis. If these random numbers are filtered, the kurtosis is changed in a predictable way. This allows one to simulate ocean noise with specified non-gaussian statistics as well as a specified spectral shape.

II. An Ocean Noise Model

It is usually considered that ocean noise is constituted as the sum of a number of different acoustic sources having, for example, sinusoidal, Gaussian, impulsive, and possibly other distributions. The result is presented as a waveform which, putting aside other criteria of randomness, can be characterized by the distribution of its sampled values, and specifically by the kurtosis of this distribution. A sinusoidal waveform is, of course, not random but the distribution of deterministically sampled values of a sine wave is formally equivalent to the probability distribution of a sine wave with uniformly distributed random phase. Fig. 1(a) reproduces a diagram

Manuscript received May 20, 1993; revised May 5, 1993.

The author is with Computing Devices Canada, Ottawa, Ont. Canada K1G 3M9.

IEEE Log Number 9215409.

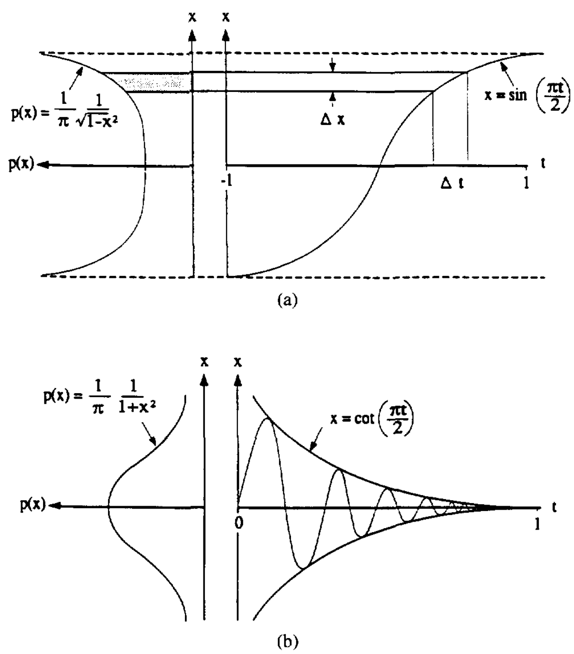

Fig. 1. Mapping a sampled waveform into a distribution. (a) sine wave. (b) damped sine wave.

from Hamming [7, 38.3] which illustrates the connection. If the continuous waveform $x(t)$ is monotonic, then the number of samples lying between $x(t_0)$ and $x(t_0+\Delta t)$ is proportional to the magnitude $|\Delta x| = |x(t_0) - x(t_0+\Delta t)|$, that is $|\Delta x| = \Delta t.|dx/dt|_{t=t_0}$. With uniform sampling, Δt is constant and simply scales the distribution. Thus, properly normalized, the distribution is:

$$p(x) = \frac{1}{2\left|\frac{dx}{dt}\right|} \tag{1}$$

which is the usual transformation formula for random variables. The same method of using a sampled deterministic function to generate a pseudorandom probability distribution was explored in [8]. For the sine wave $x(t) = \sin(\pi t/2)$ on the interval $-1 < t < 1$, the distribution is the U-shaped function:

$$p(x) = \frac{1}{\pi}\frac{1}{\sqrt{1-x^2}}. \tag{2}$$

Hamming's comment, that the "favorite" values are at the extremes of the range, is pertinent. It is equivalent to saying that the sampled values are dominated by the envelope of the waveform. Consequently, we speculate that nonsinusoidal distributions, which we associate with random noise perhaps only because they are not U-shaped, or because they resemble the pdf's of familiar random variables, may in fact arise from the deterministic sampling of sinusoids with non-uniform envelopes, that is to say, from damped sinusoids of relatively

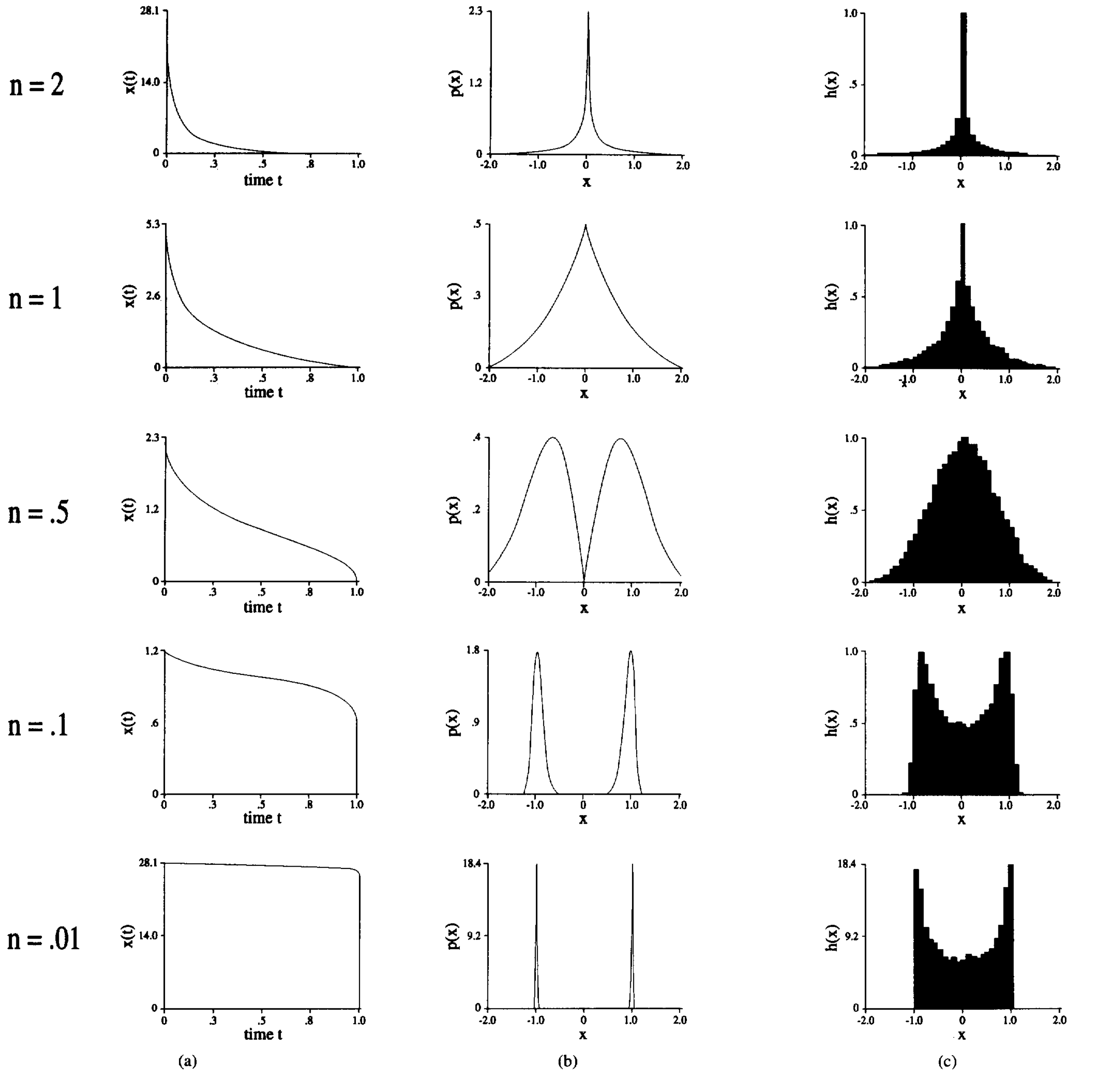

Fig. 2. For range of n. (a) envelope function $[\log(1/t)]^n$. (b) envelope pdf (4). (c) sample distribution from (9).

high frequency by which the contribution of the envelope is made relatively smooth. For example, what is called impulsive noise may arise from sampling an impulse like $x(t) = \sin(\omega t)/t$, $t \geq 0$, which, with envelope $1/t$, will give rise to a distribution proportional to $1/x^2$. The conditions of the transformation (1) are not violated as long as we consider only monontonic envelopes. To give a properly normalized example, consider the very similar impulse $x(t) = \cot(\pi t/2)\sin(\omega t)$, now confined to the interval $0 \leq t \leq 1$, as diagrammed in Fig. 1(b). The distribution of the envelope $\cot(\pi t/2)$ is, from (1), the cauchy density

$$p(x) = \frac{1}{\pi}\frac{1}{1+x^2} \tag{3}$$

typical of the heavy-tailed distributions associated with impulsive noise.

In this model of ocean noise, all sources are considered to be sinusoidal and different distributions arise according to the different kinds of damping the sinusoids are subject to.

III. Log $(1/t)$ Damping

Although familiar damping functions like $1/t$ and e^{-t} could be employed, the most suitable choice for generating sea noise distributions seems to be the function log $(1/t)$ raised to an arbitrary power. Thus, if the envelope is $[\log(1/t)]^n$, $0 \leq t \leq 1$, then the

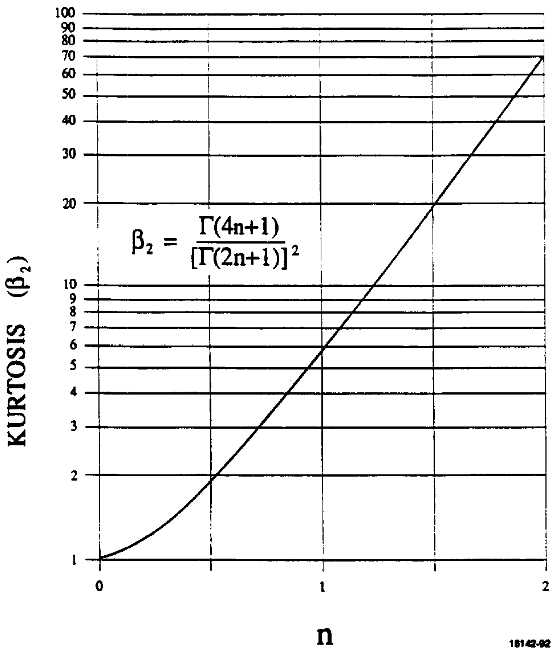

Fig. 3. Kurtosis of the envelope pdf (7) as a function of the index n. Multiply by 3/2 to get kurtosis for simulated ocean noise (14).

distribution of the envelope is from (1)

$$p(x) = \frac{1}{2n}|x|^{1/n-1}e^{-|x|^{1/n}}, \qquad -\infty < x < \infty. \tag{4}$$

These two functions are shown in the first two columns of Fig. 2 for various values of n. At $n = 1$ we have the Laplace distribution commonly used to model impulsive ocean noise; at $n = 1/2$ we have the Rayleigh distribution, a density with Gaussian tails; as $n \to 0$ the distribution tends to the discrete form $p(x) = 0.5\delta(|x| - 1)$. The envelope distribution (4) will approximate the distribution of sampled values of the waveform,

$$x(t) = \left[\log \frac{1}{t}\right]^n \sin(\omega t + \phi) \tag{5}$$

given a sufficiently large number of samples on the interval $0 < t \leq 1$, and that ω is also large; ϕ is an arbitrary initial phase. However, when (5) is sampled, the envelope distribution (4) will be "filled out" in the centre by sinusoidal oscillations. Thus we would expect the Rayleigh distributed envelope at $n = 1/2$ to become more Gaussian-like, and at $n = 0$ we would expect to obtain the sinusoidal distribution (2) precisely.

The r-th moment of (4) is easily found from [9, p. 313, (15)] as

$$m_r = \int_{-\infty}^{\infty} x^r p(x)\,dx = \Gamma(nr + 1) \tag{6}$$

and so the kurtosis is simply

$$\beta_2 = \frac{m_4}{(m_2)^2} = \frac{\Gamma(4n+1)}{[\Gamma(2n+1)]^2}. \tag{7}$$

As can be seen from the graph, Fig. 3, the range $n = 0$ to $n = 2$ covers the range of kurtosis values most likely to be encountered in ocean noise distributions.

IV. Random Number Generation

We have seen that deterministically sampled values of a damped sinusoid can produce distributions with arbitrary kurtosis. However, the sample sequence is not itself random and, except for the case of a pure sinusoid, ocean noise would normally be modeled as a random process. This could be accomplished by regarding (5) as a random variable transformation, the random sequence being generated as

$$x_m = \left[\log \frac{1}{t_m}\right]^n \sin(\omega t_m + \phi) \tag{8}$$

$m = 1, 2, 3, \cdots$, where now t_m is a sequence of independent random variables each uniformly distributed on the interval $0 < t_m \leq 1$. Again, many cycles of the sinusoid in (8) would be required to smoothly extract both the negative and positive contributions of the envelope. A single cycle over the entire duration of the envelope would lead to a skewed distribution. However, the arbitrary phase ϕ could be regarded as a second random variable, allowing the envelope and sinusoid to be sampled independently. Then, only one cycle would be needed. In other words, we could use two uniform variates to generate each sample, as in

$$x_m = \left[\log \frac{1}{t_{2m-1}}\right]^n \sin(2\pi t_{2m}). \tag{9}$$

Now each x_m is equal to the product of two independent random variables with pdf's (4) and (2). We can use the result [10]: if z_1 and z_2 are random variables whose pdf's have Mellin transforms $M_1(s)$ and $M_2(s)$ then the Mellin transform of the pdf of the product $z_1 z_2$ is just the product of the Mellin transforms $M_1(s)M_2(s)$, where the Mellin transform $M(s)$ of pdf $p(z)$ is defined as

$$M(s) = \int_{-\infty}^{\infty} z^{s-1} p(z)\,dz. \tag{10}$$

The Mellin transform of (4) is from [9, p. 313, (15)]

$$M_1(s) = \Gamma(ns - n + 1) \tag{11}$$

and the Mellin transform of (2) is from [p. 3.1, (31)]

$$M_2(s) = \frac{1}{\pi}\frac{\Gamma(\frac{1}{2})\Gamma(\frac{s}{2})}{\Gamma(\frac{1+s}{2})}. \tag{12}$$

Consequently, the Mellin transform of the pdf of x_m is

$$M(s) = \frac{1}{\pi}\frac{\Gamma(\frac{1}{2})\Gamma(\frac{s}{2})}{\Gamma(\frac{1+s}{2})}\Gamma(ns - n + 1) \tag{13}$$

and the kurtosis is

$$\beta_2 = \frac{M(5)}{[M(3)]^2} = \frac{3}{2}\frac{\Gamma(4n+1)}{[\Gamma(2n+1)]^2}. \tag{14}$$

This is just the kurtosis of the envelope distribution (7) multiplied by the factor 3/2, the kurtosis of (4); by the Mellin transform theorem, the kurtosis of a product is equal to the product of the kurtoses. The third column in Fig. 2 presents histograms, each formed from 20 000 samples of x_m generated by (9). The results are entirely as expected; note how the central part of the distribution is "filled in," producing more realistic shapes than the envelope distribution alone provides. The kurtosis (14) is the exact value for the distribution of the x_m, and the histograms are exactly specified in terms of moments. It does not seem possible to find the general distribution for the x_m in closed form. However, at $n = 1/2$ the Mellin transform (13) reduces to

$$M(s) = \frac{\Gamma(\frac{s}{2})}{\sqrt{\pi}} \tag{15}$$

for which the distribution is the Gaussian form [9, p. 313, (15)]

$$p(x) = \frac{1}{\sqrt{\pi}}e^{-x^2}. \tag{16}$$

Thus, Gaussian noise is included precisely. In fact, this case will be recognized as being one of the components of the Box–Muller transformation [11] for generating Gaussian random variables.

V. Colored Noise

The random numbers generated by (9) are statistically independent and so represent the idealized case of white noise. For many simulations this will be acceptable. However, real ocean noise is colored. The usual method for simulating colored noise is to pass white noise through an appropriate filter to implement the desired spectral shaping. But because filtering is an additive process, this tends to make non-Gaussian data more Gaussian by the central limit theorem. However, the precise quantitative relationship for kurtosis is as follows [12]: if x_n are iid random variables with kurtosis $\beta_2\{x\}$, and a_n are scalars, then the kurtosis of

$$y = \sum_{n=1}^{N} a_n x_n \tag{17}$$

is

$$\beta_2\{y\} = 3 + [\beta_2\{x\} - 3]\frac{\sum_{n=1}^{N} a_n^4}{\left(\sum_{n=1}^{N} a_n^2\right)^2}. \tag{18}$$

Thus, for a finite impulse response filter

$$y_m = \sum_{n=1}^{N} a_n x_{m-n+1} \tag{19}$$

or a first-order infinite impulse response filter, with $0 < a < 1$,

$$\begin{aligned} y_m &= a y_{m-1} + (1-a) x_m \\ &= (1-a)\sum_{n=1}^{m} a^{m-n} x_n \end{aligned} \tag{20}$$

we can specify the kurtosis of the output filtered noise y_m independent of the filter coefficients by requiring the kurtosis of the input unfiltered noise x_m to satisfy (18). Obviously the output kurtosis would have to be closer to 3 than the input kurtosis. For the recursive filter (20) one would naturally take the steady state condition $m = \infty$ and the relationship becomes, solving for the input kurtosis,

$$\beta_2\{x\} = 3 + \frac{\beta_2\{y\} - 3}{\left(\frac{1-a^2}{1+a^2}\right)}. \tag{21}$$

The procedure is therefore as follows: given the desired spectral shaping filter coefficients and the desired output kurtosis $\beta_2\{y\}$, calculate the required input kurtosis $\beta_2\{x\}$ from (18) or (21); then solve (14) for n, either directly, or using Fig. 3 (remembering to account for the 3/2 factor); with this value of n generate the random numbers from (9).

Appendix

Theorem: If x_n are iid zero-mean random variables with kurtosis $\beta_2\{x_n\} = \beta_2\{x\}$ for all n, and third moment zero, and if a_n are scalars, then

$$\beta_2\left\{\sum a_n x_n\right\} = 3 + [\beta_2\{x\} - 3]\frac{\sum a_n^4}{(\sum a_n^2)^2}$$

Proof: Let $\mu_2\{x\}$ and $\mu_4\{x\}$ denote the variance and fourth moments of x, then

$$\begin{aligned} \mu_2\{a_n x_n\} &= a_n^2 \mu_2\{x_n\} \\ \mu_4\{a_n x_n\} &= a_n^4 \mu_4\{x_n\} \end{aligned}$$

The second and fourth cumulants are [13, 26.1.13]

$$\begin{aligned} \kappa_2\{a_n x_n\} &= \mu_2\{a_n x_n\} = a_n^2 \mu_2\{x_n\} \\ \kappa_4\{a_n x_n\} &= \mu_4\{a_n x_n\} - 3[\mu_2\{a_n x_n\}]^2 \\ &= a_n^4 \mu_4\{x_n\} - 3a_n^4[\mu_2\{x_n\}]^2. \end{aligned}$$

The kurtosis is defined as [13, 16.1.18]

$$\beta_2\{a_n x_n\} = \frac{\mu_4\{a_n x_n\}}{[\mu_2\{a_n x_n\}]^2} = \frac{\mu_4\{x_n\}}{[\mu_2\{x_n\}]^2}.$$

Now use the result that the cumulant of a sum equals the sum of the cumulants. Thus,

$$\begin{aligned} \kappa_2\left\{\sum a_n x_n\right\} &= \sum \kappa_2\{a_n x_n\} = \mu_2\{x_n\}\sum a_n^2 \\ \kappa_4\left\{\sum a_n x_n\right\} &= \sum \kappa_4\{a_n x_n\} \\ &= \mu_4\{x_n\}\sum a_n^4 - 3[\mu_2\{x_n\}]^2 \sum a_n^4. \end{aligned}$$

Then

$$\begin{aligned} \beta_2\left\{\sum a_n x_n\right\} &= 3 + \frac{\kappa_4\{\sum a_n x_n\}}{[\kappa_2\{\sum a_n x_n\}]^2} \\ &= 3 + \frac{(\mu_4\{x_n\} - 3[\mu_2\{x_n\}]^2)\sum a_n^4}{[\mu_2\{x_n\}]^2(\sum a_n^2)^2} \\ &= 3 + [\beta_2\{x\} - 3]\frac{\sum a_n^4}{(\sum a_n^2)^2} \end{aligned}$$

References

[1] F. W. Machell, C. S. Penrod and G. E. Ellis, "Statistical characteristics of ocean acoustic noise processes," in *Topics in Non-Gaussian Signal Processing.* New York: Springer-Verlag, 1988, pp. 29–57.

[2] M. Bouvet and S. C. Schwartz, "Underwater noises: Statistical modeling, detection and normalization," *J. Acouts. Soc. Amer.,* vol. 83, no. 3, pp. 1023–1033, Mar. 1988.

[3] P. A. Nielsen and J. B. Thomas, "A comparison of parametric and non-parametric detector performance levels in underwater noise," *J. Acoust. Soc. Amer.,* vol. 87, no. 1, pp. 225–236, June 1990.

[4] W. C. Torrez, "On modeling the distribution of non-gaussian ambient noise," in *Proc. 19th Symp. Computer Science, Statistics,* Philadelphia, PA, Alexandria, VA, pp. 497–501.

[5] S. A. Kassam, *Signal Detection in Non-Gaussian Noise.* New York: Springer-Verlag, 1988.

[6] D. Knuth, The Art of Computer Programming, Vol. 2: Seminumerical Algorithms. New York: Addison-Wesley, 1981, sec. ed.

[7] R. W. Hamming, *Numerical Methods for Scientists and Engineers.* New York: Dover, 1986.

[8] R. J. Webster, "Spectral line profiles generated by deterministic frequency modulation," *IEEE Trans. Signal Proc.,* vol. 39, no. 4, pp. 1012–1017, Apr. 1991.

[9] A. Erdelyi, Ed., Tables of Integral Transforms, Vol. 1. New York: McGraw-Hill, 1954.

[10] B. A. Dolan, "The Mellin transform for moment generation and for the probability density of products and quotients of random variables," *Proc. IEEE,* pp. 1745–1746, Dec. 1964.

[11] G. E. P. Box and M. E. Muller, "A note on the generation of random normal deviates," *Ann. Math. Stat.,* vol. 29, no. 2, pp. 610–611, 1958.

[12] R. J. Webster, "Ambient noise statistics," *IEEE Trans. Signal Proces.,* vol. 41, no. 6, pp. 2249–2253, June 1993.

[13] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions.* New York: Dover, 1965.