Frequency Domain Modeling for Classification of Signals

Kanungo Barada Mohanty
Electrical Engineering Department
National Institute of Technology
Rourkela, India
E-mail: kbmohanty@nitrkl.ac.in

***Abstract*—A probability distribution model is proposed in this paper. Fourier Transform of a unit rectangular pulse, whose width is a random variable with Gaussian distribution, is used to derive the probability density function (p.d.f.) in the frequency domain. Result of the mathematical derivation is an exponential mathematical function involving an infinite summation over all integers. The projection theorem is used to arrive at the exact probability density function. To verify this experimentally, a randomly generated sample of Gaussian numbers, representing the pulse width is mapped onto the frequency domain, and the resulting points have a certain probability distribution, which matches with the theoretically proposed function.**

Keywords— Probability Density Functions, Gaussian distribution, Fourier Transform, sinc function.

I. INTRODUCTION

Mathematical analysis for transformation of a signal from time domain to frequency domain using the probability model is an important research topic. Ref. [1] provides a new derivation for the power received by an antenna in a reverberation chamber using Probability Density Function (PDF). Probability Density Function is analytically computed in [2] for the output noise of an interferometer. Conditional probability density function of a time-varying random signal in the presence of additive Gaussian noise is examined in [3]. The probability density function is used in [4] for the field propagation in wall tunnel. In [5] uncertainty bounds in frequency response function measurements are calculated using analytic expression of the probability density function. Though probability density function is used in different applications, its application in frequency domain modeling is yet to be explored.

Some original techniques for frequency domain modeling are described in [6,7,8]. A novel technique for frequency domain modeling is given in [9]. An efficient frequency-domain modeling and simulation method of a coupled interconnect system using scattering parameters is described in [10]. Application of frequency domain modeling are described in [11] for STATCOM is, and in [12] for HVDC systems. An approach to direct frequency-domain representation of an external system of any size or complexity is presented in [13].

Exact mathematical analysis for transformation of a signal from time domain to frequency domain using the probability density function is done in this paper, based on [6]-[8]. This provides insight into mapping of the statistical variation of a parameter of a signal, on to the variation of its frequency components in the frequency domain. Also the idea is to find out whether exact analysis can yield techniques that have a feasible computational complexity, with respect to existing techniques for estimation of statistical parameters. This in turn helps to classify signals or parts of signals (as in their frequency components) according to different application specific probability patterns. For this a unit rectangular pulse is chosen as a test signal. Its width is chosen as the statistically varying parameter, with Gaussian distribution.

II. THEORETICAL PROPOSITION

A. Problem Formulation

A unit rectangular wave is considered, represented by the function:

$$\mathrm{Rect}(t/\tau) = \begin{cases} 0 & |t| > \tau/2 \\ 1 & |t| < \tau/2 \\ 0.5 & |t| = \tau/2 \end{cases}$$

Where, τ is the pulse width and t is time.

This rectangular wave has a width τ which is a random variable. The distribution of this random variable is Gaussian and is given by the probability density function (p.d.f.):

$$\eta(\tau,\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \times \exp\left(-0.5 \times \frac{(\tau-\mu)^2}{\sigma^2}\right) \qquad (1)$$

Here, μ is the mean width of the square wave, and σ is the standard deviation of the random variable τ (i.e., the standard deviation of the width of the rectangular pulse). The frequency domain description of this pulse is given by its Fourier Transform:

$$G(\omega) = \int_{-\infty}^{\infty} \mathrm{rect}\,(t/\tau).e^{-j\omega t}dt$$

Therefore

$$G(\omega) = \frac{2}{\omega}\sin(\omega\tau/2) = \tau\,\mathrm{sinc}(\omega\tau/2) \qquad (2)$$

978-0-7695-3593-7/09 $25.00 © 2009 IEEE
DOI 10.1109/UKSIM.2009.22

Equation (2) is of the form G = Func(τ), the p.d.f. of random variable τ being Gaussian. The p.d.f. of function G is derived in the following section.

B. Derivation

Since value of sine function varies from −1 to +1, the range of function G is obtained as follows.

$$-1 \le \sin(\omega\tau/2) \le +1$$

$$\Rightarrow -\frac{2}{\omega} \le \frac{2}{\omega}\sin(\omega\tau/2) \le \frac{2}{\omega}$$

$$\Rightarrow -\frac{2}{\omega} \le G \le \frac{2}{\omega} \qquad (3)$$

From (2) the inverse trigonometric relation is obtained as follows.

$$\sin(\omega\tau/2) = \frac{\omega}{2}G$$

$$\Rightarrow \frac{\omega\tau}{2} = n\times\pi + (-1)^n \sin^{-1}(\frac{\omega G}{2})$$

$$\Rightarrow \tau_n = \frac{2}{\omega}\{n\times\pi + (-1)^n \sin^{-1}(\frac{\omega G}{2})\} \qquad (4)$$

In (4) n is an integer (zero, positive or negative). Here the sine inverse term refers to the principal values of the angle $\frac{\omega G}{2}$.

This value $\ni$ (0 , $\pi/2$) if $\frac{\omega G}{2} > 0$

and $\ni$ (0, $-\pi/2$) if $\frac{\omega G}{2} < 0$

Hence, (4) represents the inverse circular functions those map from G to τ. It is to be noted the mapping between G and τ is a one many correspondence. So, there exist infinite number of mutually exclusive inverse functions mapping from G to τ. Keeping this relation in mind the probability density function (p.d.f.) of G is derived by applying the projection theorem for transformation of random variables.

Suppose X is a continuous random variable with probability distribution f(x). Let Y = u(x) define a transformation between values of X and Y, that is NOT a one to one. If the interval over which X is defined, can be partitioned into k mutually disjoint sets such that each of the inverse functions:

$$x_1 = w_1(y),\ x_2 = w_2(y) \ldots\ x_k = w_k(y),$$

of y=u(x), defines a one to one correspondence, then the probability distribution of Y is:

$$g(y) = \sum_{i=1}^{k} f(w_i(y))\times|J_i|$$

J_i is the Jacobian of each inverse function and is defined as:

$$J_i = w_i^{'}(y),\quad i = 1, 2, 3, \ldots\ldots k$$

Here, $\tau_n = \text{InvFunc}(G,n)$ as in (4) represents the infinite inverse functions such that each is a one to one correspondence for infinite mutually disjoint sets of values over which τ is defined.

$$\frac{d\tau_n}{dG} = (-1)^n \times \frac{2}{\sqrt{4-(\omega G)^2}}$$

$$\Rightarrow J_n = \left|\frac{d\tau_n}{dG}\right| = \frac{2}{\sqrt{4-(\omega G)^2}} \qquad (5)$$

It is to be noted that J_n is independent of n and hence can be taken out of the $\sum$ sign. Also,

$$4-(\omega G)^2 > 0$$

$$\Rightarrow -\frac{2}{\omega} < G < \frac{2}{\omega} \qquad (6)$$

Hence, the probability distribution function of G is formulated.

$$\eta_2(G) = \frac{2}{\sqrt{4-(\omega G)^2}}\sum_n \frac{1}{\sqrt{2\pi\sigma^2}}\exp(Z) \qquad (7)$$

$$Z = -0.5\times(\frac{2}{\omega}(n\times\pi + (-1)^n \sin^{-1}(\omega G/2) - \mu)^2 \times \frac{1}{\sigma^2} \qquad (8)$$

ω is real varying from $-\infty$ to $+\infty$.

Equation (7) is the required p.d.f. It must be noted that (7) represents a summation of terms over integer variable n for each frequency ω (which varies from $-\infty$ to $+\infty$).

C. Plotting of Probability Function

The probability distribution is plotted using MATLAB. Equation (7) represents an infinite number of probability distributions one for each frequency, which itself can vary continuously from $-\infty$ to $+\infty$.

Issues:

1. The p.d.f. is plotted keeping the frequency fixed at 2π rad/s (or 1hz).
2. The 2nd issue is that the function to be plotted consists of an infinite summation over integer n. Hence all values of n for which there is a significant contribution are added. This value of n turns out to be from –4 to +4 (for 1hz frequency) after which the order of probability density becomes very small and hence can be safely ignored.

Under these two stipulations the function to be plotted is given by:

$$\eta_2(G) = \frac{2}{\sqrt{4-(2\pi G)^2}}\sum_n \frac{1}{\sqrt{2\pi\sigma^2}}\exp(Y) \qquad (9)$$

$$Y = -0.5\times(\frac{2}{2\pi}(n\times\pi + (-1)^n \sin^{-1}(2\pi G/2) - 1)^2 \times \frac{1}{1^2} \qquad (10)$$

The plot of probability density versus G with unity standard deviation is shown in Fig. 1.

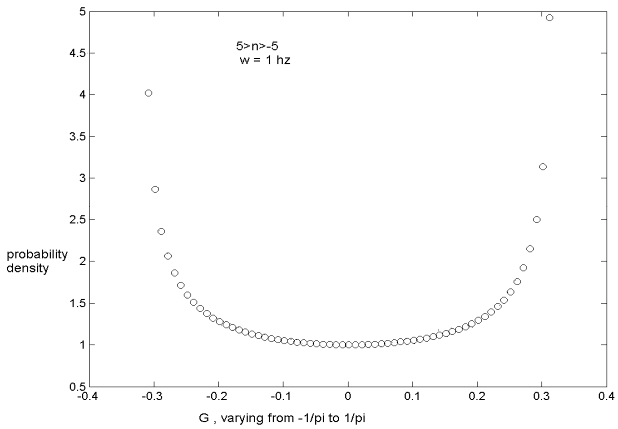

Figure 1. Probability density function vs. G with standard deviation = 1 [equation (9)-(10)]

D. *Inferences and Implications*

A number of interesting points are observed from the mathematical expression as well as the plot. Those are listed below.

1. From (2), G is purely real. It implies the p.d.f. expression (as well as plot) is for the probability distribution of the magnitude of a particular frequency component (1hz).
2. The most interesting part is the nature of this curve. The probability of G is minimum at the center of its range and increases as the limits on either side are approached. The area under the curve is approximately unity.
3. The range of values of G, divided into class intervals, is taken to be $-\frac{2}{\omega} < G < \frac{2}{\omega}$. This means the range of variation in the magnitude of G is inversely proportional to the frequency ω. Hence at higher frequencies the oscillation in the amplitude of G decreases. Also the magnitude of G decreases as the frequency increases.
4. For the 1hz frequency, summation over 9 values of n gave sufficiently accurate results. For other frequency components, the range of values of n over which the contributions are significant, are added.
5. This is the p.d.f. of Fourier transform of a function, one of whose parameter τ, is normally distributed with standard deviation equal to one. Distribution of G with standard deviation of τ as 2 and 4 are plotted in Fig.2 and Fig.3.
6. It is seen from Fig.2 and Fig.3 that as σ increases, the curve becomes flatter at the center. The magnitude of G at the range limits increase. This can be interpreted with greater fluctuations in the values of G as the variable τ becomes more and more flat in its normal distribution. This is shown in the plots.

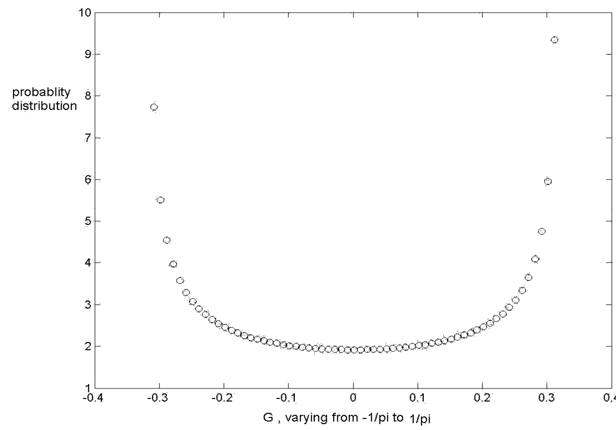

Figure 2. Probability density vs. G with standard deviation =2

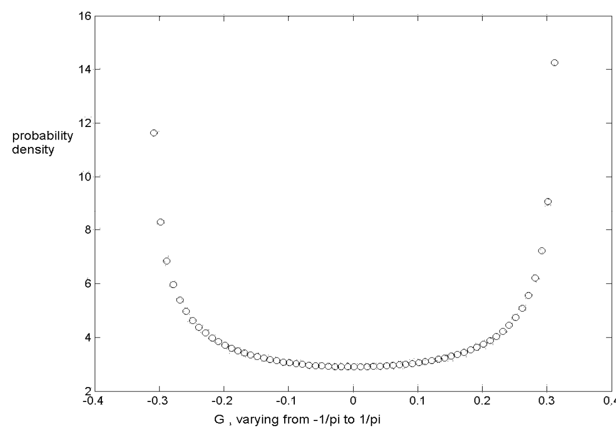

Figure 3. Probability density vs. G with standard deviation = 4.

III. EXPERIMENTAL ANALYSIS

A. *Experimental Design*

Subsequent to above theoretical derivation, an experiment is done in order to confirm the analysis. For this the basic algorithm is as follows.

1. A sample of random numbers that are Gaussian distributed is generated. This is because the width of the rectangular pulse is assumed to be Gaussian distributed. The set of these generated data points called X_i represent the statistically random variable τ, the rectangular pulse width.
2. In the next step, the Fourier Transform of the pulse at these generated data points is computed. The value of

$G(\tau)$ over the entire range of data points that represent the statistically varying width τ keeping the frequency fixed is computed. For conducting the random trial the frequency ω is arbitrarily fixedat 2π rad/s (or 1hz).

3. This set of computed values is denoted as $\tilde{G}$. The data points of $\tilde{G}$ are distributed with a certain probability pattern. The experimental probability density of $\tilde{G}$ is plotted.
4. The range of $\tilde{G}$ is divided into 20 equal class intervals. The MATLAB program computes the frequency of the data points in each interval.

The plot obtained experimentally for 1hz frequency is shown in Fig. 4 as histogram.

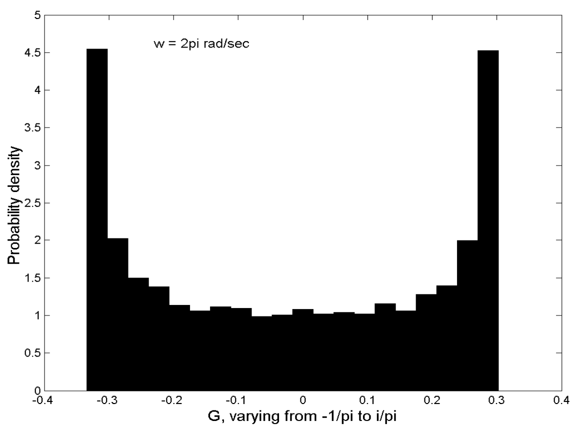

Figure 4. Histogram showing distribution of Fourier Transform of randomly generated Gaussian data points

B. *Comparative Analysis*

The histogram of Fig. 4 is compared with the curve that is obtained from the theoretical calculations, and shown in Fig.5.

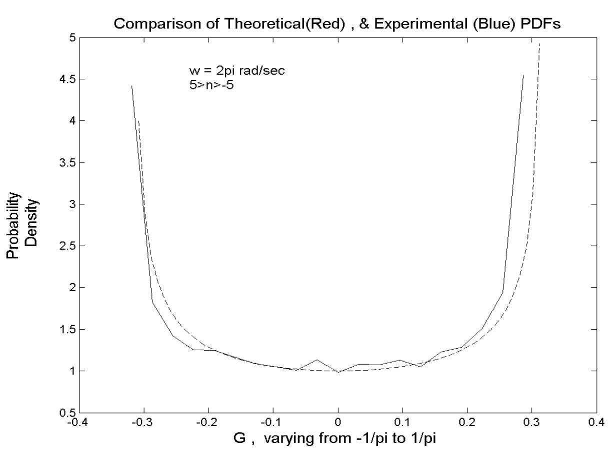

Figure 5. Comparison of experimental (solid line) and theoretical (dotted line) p.d.f.

IV. CONCLUSIONS

1. The outcome of the experimental plot when superimposed on the theoretical plot is seen to match remarkably, which confirms that the mathematical analysis is pretty accurate.
2. It is found that the probability density function is U shaped at the 1Hz frequency as well as for higher frequency components that are investigated. It can be concluded that the probability of the frequency component taking on the two extreme values of the magnitude range is maximum and taking on values near center is minimum.
3. As the pulse width becomes more and more deviant statistically (σ increases), the p.d.f. acquires a flatter shape at the center and higher magnitude at the edges. This can be interpreted as increased fluctuations in the magnitude.
4. As frequency is increased, the range of variation of magnitude ($-2/\omega$ to $2/\omega$) decreases. This is consistent with the fact that in case of a Sinc Function, the amplitude of frequency response decreases as frequency increases.
5. One important issue is the factor n. The p.d.f. is a summation of a complicated exponential function over the integer n. It turns out that the series is convergent and in this case the sum is taken for 9 values of n (from –4 to +4). The plot for a frequency of 10Hz is observed, and found that the sum is to be taken over greater values of n from –40 to +40. This is a clear indication of the fact that as frequency of interest increases, so does the contribution of higher value n terms. In order to obtain a theoretical p.d.f. plot for a high frequency component, the summation over a greater value of n is to be taken.

The exact analysis technique that is used to study the probability patterns of a signal magnitude in the frequency domain, when one of its parameters is a random variable can be further extended. In actual applications where the variations are unknown and at the same time it is not possible to study a large sample set (such as biomedical signals), then suitable bi-variate or multivariate probability models can be constructed for their frequency domain representation. This is particularly important because, not just the pulse width, but its magnitude and phase should also be considered as random variables because they too show probabilistic variations. Hence, given such a non-deterministic signal, the uncertainty is translated to the frequency domain. This is done to get an idea of exact fluctuation of different frequency components in their 'probability space'. This in turn can be utilized for better design of systems that are required to deal with such signals. If the signal under study is the output response of a control system, then such an analysis could lead to a fresh approach to diverse topics like robust stability of control systems.

REFERENCES

[1] D. A. Hill, "Probability Density Function of Power Received in a Reverberation Chamber," IEEE Trans. Electromagnetic Compatibility, vol. 50, no. 4, pp. 1019-1019, Nov. 2008.

[2] M. D. Todd, "On the Probability Structure of Output Noise From a Digital Phase Demodulation System Subject to Biased Intensity-Based Input Noise," Journal of Lightwave Technology, vol. 26, no. 14, pp. 2291-2300, Jul. 2008.

[3] Y. S. Shmaliy, "On the Multivariate Conditional Probability Density of a Vector Perturbed by Gaussian Noise," IEEE Trans. Information Theory, vol. 53, no. 12, pp. 4792-4797, Dec. 2007.

[4] H. Y. Pao, "Probability density function for waves propagating in a straight rough wall tunnel," IEEE International Symposium of Antennas and Propagation Society, pp. 2975- 2978, June 2004.

[5] R. Pintelon, Y. Rolain, W. Van Moer, "Probability density function for frequency response function measurements using periodic signals," IEEE Trans. Instru. and Meas., vol. 52, no. 1, pp. 61-68, Feb. 2003.

[6] Poor H., An Introduction to Signal Detection and Estimation, Springer-Verlag: New York, 1985, ch. 4.

[7] S. Chen, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks," IEEE Trans. Neural Networks, vol. 4, no. 4, pp. 570–579, July 1993.

[8] R. W. Lucky, "Automatic equalization for digital communication," Bell Syst. Tech. J., vol. 44, no. 4, pp. 547–588, Apr. 1965.

[9] M. Silveira, I. Elfadel, J. White, M. Chilukuri, and K. Kenneth, "Efficient frequency-domain modeling and circuit simulation of transmission lines," *IEEE Trans. Comp., Hybrids, Manufact. Technol.*, vol. 17, pp. 505-513, Nov. 1994.

[10] W. T. Beyene, J. Schutt-Aine, "Accurate frequency-domain modeling and efficient circuit simulation of high-speed packaging interconnects," IEEE Trans. Microwave Theory and Technique*s*, vol. 45, no. 10, pp. 1941-1947, Oct 1997.

[11] A. R. Wood, C. M. Osauskas, "A linear frequency-domain model of a STATCOM," IEEE Trans. Power Delivery, vol. 19, no. 3, pp. 1410-1418, July 2004.

[12] D. J. Hume, A. R. Wood, and C. M. Osauskas, "Frequency domain modeling of interharmonics in HVdc systems," Proc. Inst. Elect. Eng., Gen., Transm. Dist., vol. 150, no. 1, pp. 41-48, Jan. 2003.

[13] A. Semlyen, M. R. Iravani, "Frequency domain modeling of external systems in an Electro-Magnetic Transients Program," IEEE Trans. Power Systems, vol. 8, no. 2, pp. 527-533, May 1993.