On Defining White Noise

R. M. Howard
Department of Electrical and Computer Engineering
Curtin University of Technology
GPO Box U1987, Perth, Western Australia, 6845
r.howard@exchange.curtin.edu.au

Abstract--**The mathematical background necessary to rigorously define white noise is detailed. It is shown that it is necessary to accept an arbitrarily small correlation time, and a flat power spectral density approximation over a finite, but arbitrarily large, frequency range, to adequately define a white noise random process consistent with the Wiener-Khintchine relationships. Dichotomous random processes are used to illustrate appropriate principles and problems. The results are generalized to Gaussian white noise.**

Keywords--white noise; power spectral density; autocorrelation; Wiener-Khintchine; dichotomous noise.

I. INTRODUCTION

The history of random phenomena is diverse and involves many famous physicists and mathematicians; an excellent overview is detailed in [1]. In electrical engineering the term 'noise' is widely associated with the random fluctuations that arise in electronic components and devices due to the random nature of electron movement. Due to the extremely fast nature of electron movement a large class of noise phenomenon are characterized as being 'white'. White noise is ubiquitious and limits the performance of many electronic/photonic based systems including most communication systems.

Mathematicians and Physicists usually define white noise in terms of the properties of one dimensional Brownian motion, usually called the Wiener process, and a useful introduction is given in [2]. Ito and Stratonovich stochastic calculus have been developed to rigorously underpin this approach e.g. [3]. A Wiener process can be defined as the limit of a sequence of random walk processes where, for the ith random walk and as illustrated in Fig. 1, the time step is scaled down by a factor i^2 and the amplitude change is scaled down by a factor of i. The rate of change increases according to i. White noise is defined as the rate of change of the Wiener process and in the limit becomes uncorrelated between pairs of time instants and is of infinite magnitude.

In contrast, in Engineering disciplines it is usual to define a white noise random process mathematically in terms of a specified autocorrelation and/or power spectral density function with the underlying physical random process being left undefined. Specifically, white noise is usually defined, e.g. [4] [5], as one which is uncorrelated at all arbitrarily chosen pairs of time instants, has an autocorrelation function, $R(\tau)$, which is an impulse and a constant power spectral density, $G(f)$, i.e.

$$R(\tau) = k\delta(\tau), \qquad G(f) = \eta/2. \tag{1}$$

Here δ is the Dirac delta. The autocorrelation function and power spectral density functions are assumed to satisfy the Wiener-Khintchine theorem, e.g. [6]. Further, in communication theory it is usual, e.g. [6], to define white noise by its power spectral density according to $G(f) = \eta/2$.

With the Engineering approach, and definitions, for white noise an infinite power random process, i.e. a random process which physically does not exist, is implied. The issue, is to establish clearly, and rigorously, finite power random processes which exhibit characteristics consistent with the above definitions and which have a clear physical form. This is important pedagogically. Section II details the necessary mathematical background. In section III, dichotomous random processes are used to illustrate the problems with defining white noise random processes. In section IV the results are generalized to Gaussian white noise.

II. BACKGROUND THEORY

A. Mathematical Results

First, Lebesgue integration, [7] and [8], provides an elegant theoretical framework for integration theory. A function $f:\mathbf{R} \to \mathbf{R}$ is said to be (Lebesgue) integrable on a set E if $\int_E f(t)dt$ is finite. A bounded function, which is Riemann integrable on an interval, is Lebesgue integrable and the two integrals are equal. Lebesgue integration handles the integral of a function at a point best with the general result:

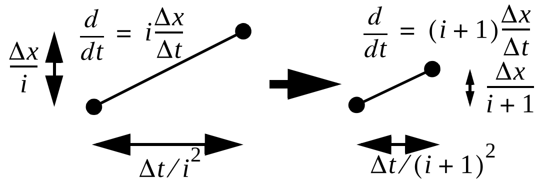

Figure 1. Illustration of the change in scaling of one step in a random walk between the ith and the $(i+1)th$ random walks used to define a Wiener random process.

$$M(E) = 0 \qquad \Rightarrow \qquad \int_E f(t)dt = 0 . \tag{2}$$

Here M is the measure operator (measure can be understood here as 'length') and this equation states: 'zero length implies zero area'.

Second, two important results, which give sufficient conditions for the interchange of limit and integral operations, are the monotone and dominated convergence theorems, e.g. [7]. The dominated convergence theorem is more powerful and for reference is stated as follows:

If $\{f_n:\boldsymbol{R} \to \boldsymbol{R}\}_{n=1}^{n=\infty}$ is a sequence of integrable functions such that $\lim_{n \to \infty} f_n(t) = f(t)$, and there exists an integrable function $g:\boldsymbol{R} \to \boldsymbol{R}$ such that $|f_n(t)| < |g(t)|$ for all n and over the region of integration, then

$$\lim_{n \to \infty} \int f_n(t)dt = \int \lim_{n \to \infty} f_n(t)dt = \int f(t)dt . \tag{3}$$

Consistent with these results it is not possible, in general, to interchange the order of the limit and integral operations for a sequence of functions which, in the limit, become impulsive. Consider a sequence of functions $\{\delta_i\}_{i=1}^{\infty}$ which is such that

$$\lim_{i \to \infty} \delta_i(t) = \begin{cases} \infty & t = 0 \\ 0 & t \neq 0 \end{cases} \qquad \lim_{i \to \infty} \int_{-\infty}^{\infty} \delta_i(t)dt = 1 . \tag{4}$$

With the definition $\delta(t) = \lim_{i \to \infty} \delta_i(t)$ it is the case, consistent with (2), that

$$\lim_{i \to \infty} \int_{-\infty}^{\infty} \delta_i(t)dt = 1 \neq \int_{-\infty}^{\infty} \lim_{i \to \infty} \delta_i(t)dt = \int_{-\infty}^{\infty} \delta(t)dt = 0 . \tag{5}$$

Importantly, the Dirac delta is associated with the limit of the sequence of functions $\{\delta_i\}_{i=1}^{\infty}$ and not with the function definition $\delta(t) = \lim_{i \to \infty} \delta_i(t)$. For notational convenience $\lim_{i \to \infty} \int_{-\infty}^{\infty} \delta_i(t)dt = 1$ is written as $\int_{-\infty}^{\infty} \delta(t)dt = 1$. However, it is very common for this notation to be overlooked and for $\int_{-\infty}^{\infty} \delta(t)dt = 1$ to be used as a mathematical relationship (and sometimes with erroneous outcomes [9]).

Third, infinity is notation for the limit of an unbounded sequence and, in general, must be considered as such. Accordingly, the formulation that follows is based on the finite interval $[0, T]$.

B. *Power Spectral Density Theory* [8]

Consider a basis set $\{b_1, \ldots\}$ for the interval $[0, T]$. Any signal $x:\boldsymbol{R} \to \boldsymbol{C}$, which is square integrable, i.e. $x \in L^2$ (see [10]), can be decomposed according to

$$x(t) = \sum_{i=1}^{\infty} c_i b_i(t), \qquad t \in [0, T] . \tag{6}$$

The power spectral density of x, with respect to the basis set $\{b_1, \ldots\}$, defines the power of the individual basis signals $c_i b_i(t)$, $i \in \{1, \ldots\}$. By far the most general basis set used is a sinusoidal basis set. For such a basis set, and for the interval $[0, T]$, Parseval's theorem leads to

$$P(T) = \frac{1}{T}\int_0^T |x(t)|^2 dt = \frac{1}{T}\int_{-\infty}^{\infty} |X(T, f)|^2 df \tag{7}$$

and, hence, the definition of the power spectral density according to

$$G(T, f) = |X(T, f)|^2 / T . \tag{8}$$

Here $X(T, f)$ is the Fourier transform of x on the interval $[0, T]$. The time averaged autocorrelation function of x is defined according to

$$R(T, \tau) = \frac{1}{T} \int_{\max\{0, \tau\}}^{T + \min\{0, \tau\}} x(t)x^*(t - \tau)dt \tag{9}$$

where $*$ is the conjugation operator and $-T < \tau < T$.

For a random process defined by an ensemble (set) of signals the autocorrelation and power spectral density are defined as a weighted average according to

$$R(T, \tau) = \sum_{i=1}^{\infty} p_i R_i(T, \tau), \qquad G(T, f) = \sum_{i=1}^{\infty} p_i G_i(T, f) \tag{10}$$

where the subscript i denotes the ith waveform of the random process and p_i is the probability of occurrence of the ith waveform.

It can then be shown that the power spectral density-autocorrelation function relationships, for either an individual signal or a random process, are:

$$R(T, \tau) = \int_{-\infty}^{\infty} G(T, f)e^{j2\pi f\tau} df, \qquad G(T, f) = \int_{-T}^{T} R(T, \tau)e^{-j2\pi f\tau} d\tau. \tag{11}$$

The relationships defined by (11) are the Wiener-Khintchine relationships for a signal or a random process. It is common to assume that stationarity is necessary for the Wiener-Khintchine relationships (11). This is not the case as these relations, as specified by (11), follow from (8) and (9). Finally, consistent with (7) and (9), the average power on the interval $[0, T]$ is

$$P(T) = R(T, 0) = \int_{-\infty}^{\infty} G(T, f)df \qquad \textbf{(12)}$$

i.e. the sum of the power in the constituent signals equals the total power.

C. Power Spectral Density Theory - Infinite Interval

For the infinite interval the autocorrelation function is defined according to

$$R_{\infty}(\tau) = \lim_{T \to \infty} R(T, \tau) . \qquad \textbf{(13)}$$

Subtleties exist, in general, for the infinite interval case for defining the power spectral density and these are detailed in ch. 3 of [8]. For the following discussion, however, the results stated above for the finite interval case will suffice.

III. DICHOTOMOUS WHITE NOISE

To illustrate appropriate principles, consider a dichotomous random process; a random process defined by associating one of two possible values, e.g. A_o or $-A_o$, with each signalling interval of a set duration t_o. One waveform from a dichotomous random process is illustrated in Fig. 2. For the case of equally probable amplitudes, and with the amplitudes being independent from one signalling interval to the next, it is the case that the autocorrelation and power spectral density functions for the dichotomous random process are as defined in Fig. 3.

First, consider an infinite sequence of dichotomous random processes where the ith random process is defined by a signalling period of $t_o = 1/i$ seconds and with amplitudes from the set $\{-A, A\}$. For the ith random process the autocorrelation and power spectral density functions are defined in Fig. 3 for the case of $A_o = A$, $t_o = 1/i$ and $f_o = i$. For all random processes in the sequence the Wiener-Khintchine relationships (11) hold. Consider the limit of the sequence of random processes and the definitions

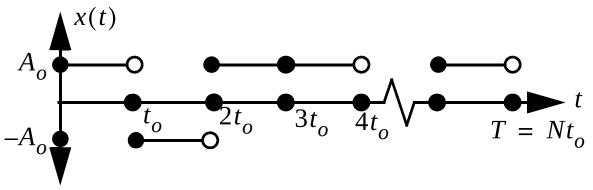

Figure 2. One waveform from a dichotomous random process with amplitudes from the set $\{-A_o, A_o\}$ and with t_o being the signalling interval duration.

$$R(T, \tau) = \lim_{i \to \infty} R_i(T, \tau) = A^2\delta_k(\tau),$$
$$G(T, f) = \lim_{i \to \infty} G_i(T, f) = 0. \qquad \textbf{(14)}$$

Here $\delta_k(\tau)$ is the Kronecker delta. The limiting autocorrelation function has the required property of white noise, namely, uncorrelatedness for all pairs of points. However, $R(T, \tau)$ and $G(T, f)$ do not satisfy the Wiener-Khintchine relationships as stated by (11). Further,

$$P(T) = \lim_{i \to \infty} \int_{-\infty}^{\infty} G_i(T, f)df = A^2 \neq \int_{-\infty}^{\infty} G(T, f)df = 0 \qquad \textbf{(15)}$$

i.e. the order of limit and integral cannot be interchanged as, for example, the conditions specified in the dominated convergence theorem cannot be satisfied. The power is not equal to the integral of the power spectral density as required by a power spectral density function and consistent with (12).

Second, consider an infinite sequence of dichotomous random process where the ith random process is defined by a signalling period of $t_o = 1/i$ seconds and with amplitudes from the set $\{-\sqrt{i}A, \sqrt{i}A\}$. For the ith random process the autocorrelation and power spectral density functions are defined in Fig. 3 for the case of $A_o = \sqrt{i}A$, $t_o = 1/i$ and $f_o = i$. Again all random processes in the sequence satisfy the Wiener-Khintchine relationships (11). In the limit:

$$R(T, \tau) = \lim_{i \to \infty} R_i(T, \tau) = \begin{cases} \infty & t = 0 \\ 0 & t \neq 0 \end{cases} \qquad \textbf{(16)}$$

$$G(T, f) = \lim_{i \to \infty} G_i(T, f) = A^2 \qquad \textbf{(17)}$$

and uncorrelatedness at all pairs of points, consistent with white noise, results. However,

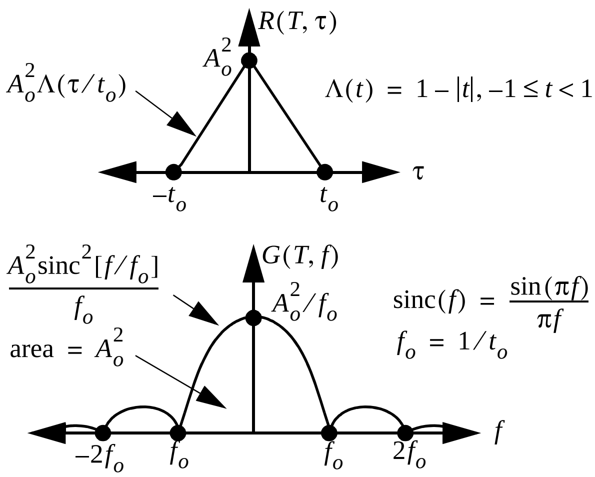

Figure 3. Autocorrelation and power spectral density function of the dichotomous random process defined in Fig. 2.

$$P(T) = \lim_{i \to \infty} \int_{-\infty}^{\infty} G_i(T, f) df = \infty \quad \lim_{i \to \infty} \int_{-\infty}^{\infty} R_i(T, \tau) d\tau = A^2 \quad \textbf{(18)}$$

and the limiting random process has infinite power for all finite values of T. It is the case that

$$R(T, \tau) = \lim_{f_o \to \infty} \int_{-f_o}^{f_o} G(T, f) e^{j2\pi f\tau} df = \begin{cases} \infty & t = 0 \\ 0 & t \neq 0 \end{cases} \quad \textbf{(19)}$$

as required by the Wiener-Khintchine relationship. However, consistent with (2)

$$\int_{-\infty}^{\infty} R(T, \tau) e^{-j2\pi f\tau} d\tau = 0 \neq G(T, f) \quad \textbf{(20)}$$

which is inconsistent with the Wiener-Khintchine relationship.

Note, for both cases, all random processes in the sequences exhibit finite power and a finite specified correlation time. The problem arises by taking the limit and attempting to define a random process which is uncorrelated for all pairs of time instants. For the first case a random process with zero power spectral density results; for the second a random process with infinite power results. For both cases the limiting autocorrelation functions and power spectral densities do not satisfy the Wiener-Khintchine relationships.

The unsolved problem is: Is there a limiting random process corresponding to the limit autocorrelation and power spectral density functions? This problem ultimately raises issues about the transition from the discrete to the continuum and the definition of the real line [11]. The following is a partial solution and is illustrated for the first case noted above: First, the sequence of random processes is a sequence without end. Accordingly, and consistent with a sequence, and limit operation, the following can be stated: $\forall \tau \in [-T, T]$, $\forall \varepsilon > 0$, $\exists I_o$ such that for all $i > I_o$ it is the case that

$$\left| R_i(T, \tau) - A^2 \delta_k(\tau) \right| < \varepsilon . \quad \textbf{(21)}$$

That is, for a specified maximum correlation time τ_o, and level $\varepsilon > 0$, it follows that all random processes in the sequence, after the $I_o th$ random process, exhibit autocorrelations within ε of zero for all $\tau_o < |\tau| < T$. Hence, excluding an arbitrarily chosen small correlation time τ_o 'white noise' can be defined consistent with a physical process and such that the Wiener-Khintchine relationships hold.

To further illustrate, consider the common communication context where noise, with a set power spectral density level of η, is hypothesized to exist over a fixed and finite frequency range of $-f_{BW} < f < f_{BW}$. Such a power spectral density can be approximated, arbitrarily accurately, by a finite power dichotomous random process with amplitudes from the set $\{-\eta f_o, \eta f_o\}$, and a signalling period $t_o = 1/f_o$, provided $f_o \gg f_{BW}$. The autocorrelation and power spectral density functions are defined in Fig. 3 for the case of $A_o = \sqrt{\eta f_o}$ and $t_o = 1/f_o$. The following approximations then hold:

$$\begin{aligned} G(T, f) &\approx \eta & |f| \ll f_o \\ R(T, \tau) = \eta f_o \Lambda(\tau/t_o) &= 0 & |\tau| > 1/f_o \end{aligned} \quad \textbf{(22)}$$

where f_o is chosen, arbitrarily large, such that a set accuracy level in the approximations is attained. The power in the random process is given by

$$P(T) = R(T, 0) = \int_{-\infty}^{\infty} G(T, f) df = \eta f_o . \quad \textbf{(23)}$$

For all finite values of T, and f_o, the Wiener-Khintchine relationships, given by (11), hold.

Thus, with the acceptance of an arbitrarily small non-zero correlation time, and a flat power spectral density approximation over a finite, but arbitrarily large, frequency range, white noise can be rigorously defined consistent with the Wiener-Khintchine relationships and a physical random process.

IV. GENERALIZATION

A. From Dichotomous White Noise to Gaussian White Noise

By far the most common noise encountered in electronic, and photonic, systems is Gaussian white noise. The term 'Gaussian' refers to the probability density function, at a set time, of the random process. In many instances the Gaussian probability density function is accurately approximated, and the approximation, in general, arises from the sum of independent random processes according to the central limit theorem. For a summation of independent dichotomous random processes the De-Moivre Laplace theorem specifies the approximation to a Gaussian form, e.g. p. 259 of [8] and [12]. Such a sum can arise as an on/off approximation to the random movement of a large number of entities.

B. Non-Dichotomous Based White Noise Random Processes

Many random processes are defined by signals based on the association of a waveform, chosen from a defined signal set $S = \{\phi_1, \dots\}$, with a set time interval, or a randomly chosen time, as illustrated in Fig. 4, and defined according to (e.g. ch. 5 [8]):

$$x_1(t) = \sum_{i=1}^{N} \phi_i(t - it_o) \qquad \phi_i \in S, t \in [0, T], T = Nt_o \quad \textbf{(24)}$$

$$x_2(t) = \sum_{i=1}^{N} \phi(t - t_i) \qquad S = \{\phi(t)\}, \quad t, t_i \in [0, T] \quad \textbf{(25)}$$

or

$$x_3(t) = \sum_{i=1}^{N} \phi_i(t - t_i) \qquad \phi_i \in S;\ t, t_i \in [0, T]. \qquad \textbf{(26)}$$

The first form is a signalling waveform and is consistent with many communication signals. The second form, for the case where each of the times $t_1, \ldots, t_N$ are chosen at random from the interval $[0, T]$ with a uniform distribution, is one waveform from a shot noise process. The final form defines a signal which is a generalization of the shot and signalling waveforms. For the case of independence between signal waveforms, and where the mean of the signals in the signalling set is zero, the power spectral density of these forms is given by

$$G(T, f) = r\sum_i p_i |\Phi_i(f)|^2 \qquad \textbf{(27)}$$

where r is the mean waveform rate, $\Phi_i(f)$ is the Fourier transform of the ith signal in the signal set (the assumption here is that signals in the signal set have negligible energy outside the interval $[0, T]$), and p_i is the probability of occurrence of the ith signal in the signal set.

For the case where each signal in the signal set has finite energy, and a power spectral density which is approximately constant over the interval $-f_o < f < f_o$, it follows that each of the random processes defined by (24) to (26) are consistent with white noise with a correlation time of less than $1/f_o$ seconds. Thus, for example, shot noise processes generally lead to random processes which are white (but not necessarily Gaussian).

A new random process defined by an independent sum of such processes will lead (approximately) to Gaussian white noise when the sum is sufficiently large. Thus, for example, noise arising from the sum of many individual shot noise processes associated with the rapid transit of electrons across a PN junction yields Gaussian white noise.

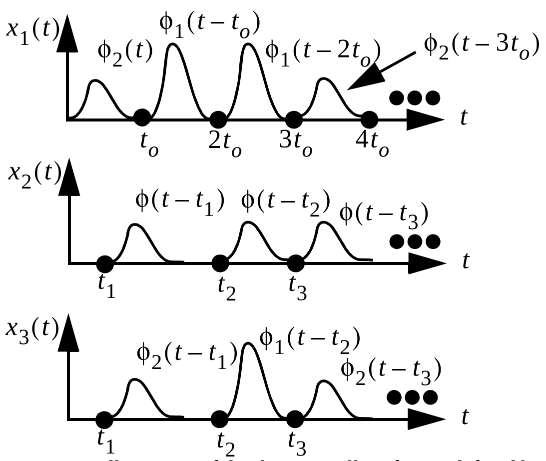

Figure 4. Illustration of the three signalling forms, defined by (24) to (26), assuming a signal set $S = \{\phi_1, \phi_2\}$.

C. *Infinite Interval*

Results have been presented for the finite interval $[0, T]$. Results can be extended to the infinite interval $[0, \infty]$ by taking the limit of $T \to \infty$. For the random processes discussed above this is unproblematic as $R(T, \tau)$ and $G(T, f)$ are independent of T. In general, however, and in particular for the case of random processes containing periodic components the approach detailed in [8] is appropriate.

D. *$1/f$ Noise*

The approach of defining the power spectral density for the finite interval, and treating the infinite interval as a limit of the interval $[0, T]$, has been successfully applied to models for $1/f$ noise [13]. It is through such an approach that the infinite power problem associated with the $1/f$ spectral form can be clarified. Importantly, the approach leads to physical models for $1/f$ noise.

V. CONCLUSION

This paper has carefully detailed the necessary mathematical background to provide, consistent with the Engineering approach, a rigorous definition for white noise. It is was shown that it is necessary to accept an arbitrarily small correlation time, and a flat power spectral density approximation over a finite, but arbitrarily large, frequency range, to adequately define a physically realizable white noise random process consistent with the Wiener-Khintchine relationships. Dichotomous random processes were used to illustrate appropriate principles and problems. The generalization to Gaussian white noise was detailed.

REFERENCES

[1] L. Cohen, ‘The history of noise’, IEEE Signal Processing Magazine, pp. 20-45, Nov. 2005.

[2] D. T. Gillespie, ‘The mathematics of Brownian motion and Johnson noise’, America Journal of Physics, vol. 64, pp. 225-240. 1996.

[3] C. W. Gardner, Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences, 2nd Ed., Springer, 1985, ch.4.

[4] A. Papoulis & S. U. Pillai, Probability, Random Variables and Stochastic Processes, McGraw Hill, 2002, p. 413.

[5] J. H. Park, ‘White noise and the delta function’, Proceedings of the IEEE, vol. 56, pp. 114-115, 1968.

[6] S. Haykin, Communication Systems, 4th Ed., Wiley, 2001, pp. 46, 61.

[7] D. C. Champeney, A Handbook of Fourier Theorems, Cambridge University Press, 1987, p. 26.

[8] R. M. Howard, Principles of Random Signal Analysis and Low Noise Design: The Power Spectral Density and its Applications, Wiley, 2002, ch. 3.

[9] R. M. Howard, ‘Dirac delta and singular distributions: The general non-good function case’, unpublished.

[10] E. C. Titchmarch, Introduction to the Theory of Fourier Integrals, Oxford, 1948, p. 10.

[11] D. A. Sprecher, Elements of Real Analysis, Dover, 1970, ch. 8.

[12] W. Feller, An Introduction to Probability Theory and its Applications, 3rd Ed., Wiley, 1968, pp. 182-186.

[13] R. M. Howard & L. A. Raffel, 'General models for 1/f noise', in Proceedings of SPIE Vol. 5113, 'Noise in Devices and Circuits', Ed. by M. Jamal Deen, Zeynep Celik-Butler, Micheal E. Levinshtein, (SPIE, Bellingham, WA, 2003), pp. 282-293, 2003.