Chapter 12: Mean Square Calculus

Many applications involve passing a random process through a system, either dynamic (*i.e.,* one with memory that is described by a differential equation) or one without memory (for example, $Y = X^2$). In the case of dynamic systems, we must deal with derivatives and integrals of stochastic processes. Hence, we need a *stochastic calculus,* a calculus specialized to deal with random processes.

It should be clear that such a calculus must be an extension of the concepts covered in Chapter 11. After all, any definition of a derivative must contain the notion of limit (the definition can be stated as a limit of a function *or* a sequence). And, an integral is just a limit of a sequence (for example, recall the definition of the Riemann integral as a limit of a sum).

One might ask if "ordinary" calculus can be applied to the sample functions of a random process. The answer is yes. However, such an approach is overtly restrictive and complicated, and it is not necessary in many applications. It is restrictive and complicated because it must deal with *every* possible sample function of a random process. In most applications, the complications of an "ordinary" calculus approach are not required since only statistical averages (such as means and variances) are of interest, not individual sample functions.

Many applications are served well by a calculus based on mean-square convergence concepts similar to those introduced in Chapter 11. Such a *mean-square calculus* discards the difficulties of having to deal with *all* sample functions of a random process; instead it "uses" only the "important" sample functions, those that influence a statistical average of interest (the average power, for example). Also, the mean-square calculus can be developed adequately using "ordinary" calculus concepts; measure-theoretic techniques are not required. The development of mean-square calculus parallels the development of "ordinary" calculus (concepts in m.s calculus have counterparts in "ordinary" calculus and vice versa). For these reasons, the mean square calculus is included in most advanced books on applied random processes, and it is the topic of this chapter.

The major "pillars" of the mean-square calculus are the notions of *mean-square limit*, *mean-square continuity*, *mean square differentiation* and *mean-square integration*. From a limited perspective, one could say that these notions are only simple applications of functional analysis (more specifically, *Hilbert space* theory), a field of study involving vector spaces that serves as the basis, and "unifying force", of many electrical engineering disciplines. This chapter introduces the above-mentioned "pillars", and we give "hints" at the broader vector space interpretation of these concepts and results.

Finite Average Power Random Processes

In this chapter, we consider only real-valued random processes with

$$E[X^2(t)] < \infty \tag{12-1}$$

for all t. Such processes are said to have *finite average power*, or they are said to be *second-order* random processes. We deal with real-valued processes only in order to simplify the notation and equations. Excluding complex-valued processes, and eliminating complex notation, does not restrict coverage/discussion/understanding of the *basic* concepts of mean-square calculus. Note that every result in this chapter can be generalized easily to the complex-valued random process case.

For every fixed t, finite-second-moment random variable X(t) is in the vector space $\boldsymbol{L}_2$ discussed in Chapter 11. As a result, we can apply to random processes the inner product and norm notation that was introduced in the previous chapter. Let $X(t_1)$ and $Y(t_2)$ be two finite-power random processes. The *inner product* of $X(t_1)$ and $Y(t_2)$ is denoted as $\langle X(t_1), Y(t_2)\rangle$, and it is defined by

$$\langle X(t_1), Y(t_2)\rangle \equiv E[X(t_1)Y(t_2)]. \tag{12-2}$$

For each t, the norm of X(t), denoted in the usual manner by $\|X(t)\|$, is defined by

$$\|X(t)\|^2 = \langle X(t), X(t)\rangle = E[X^2(t)]\,. \tag{12-3}$$

Finally, note that the correlation function Γ can be expressed as

$$\Gamma(t_1, t_2) \equiv E\big[X(t_1)X(t_2)\big] = \big\langle X(t_1), X(t_2)\big\rangle\,. \tag{12-4}$$

We assume that all finite-power processes have zero mean. This assumption imposes no real limitation. Since (12-1) implies that E[X] < ∞ (use the Cauchy-Schwarz inequality to show this), we can form the new random process

$$Y(t) = X(t) - E[X(t)] \tag{12-5}$$

that has zero mean. Hence, without loss of generality, we limit ourselves to zero-mean, finite-average power random processes.

The theory of m.s. limits, m.s. continuity, m.s. differentiation and m.s. integration of a stochastic process can be given using the inner product and norm notation introduced above. Alternatively, one can use the equivalent "expectation notation". Both notational methodologies have advantages and disadvantages; in this chapter, we will use both.

Limits

The limit of a random process can be defined in several different ways. Briefly, we mention some possibilities before focusing on the mean-square limit.

Surely (Everywhere): As $t' \to t$, $X(t',\omega)$ approaches $Y(t,\omega)$ for **every** $\omega \in \mathbf{S}$. This is the "ordinary Calculus" limit; it is very restrictive and rarely used in the theory of random processes.

Almost Surely (Almost Everywhere): There exists $A \subset \mathbf{S}$, $\mathbf{P}(A) = 1$, such $X(t',\omega) \to Y(t,\omega)$ as $t' \to t$ for **every** $\omega \in A$. This is only slightly less restrictive ("weaker") then requiring that the limit exist everywhere (the former case), and it is rarely used in applications.

In Probability: As $t' \to t$, $X(t',\omega)$ approaches $Y(t,\omega)$ *in probability* (i.p.) if, for all $\varepsilon > 0$, we have

$$\underset{t' \to t}{\text{limit}}\ \mathbf{P}\Big[\big|X(t') - Y(t)\big| > \varepsilon\Big] = 0\,. \tag{12-6}$$

Often, this is denoted as

$$Y(t) = \underset{t' \to t}{\text{l.i.p}}\ X(t')\,. \tag{12-7}$$

This form of limit is "weaker" than the previously-defined surely and almost surely limits. Also, it is "weaker" than the mean-square limit, defined below.

Limit in the Mean

For finite-power random processes, we adopt the limit-in-the mean notation that was introduced for sequences in Chapter 11. Specifically, as t' approaches t (*i.e.*, $t' \to t$), we state that process $X(t')$ has the *mean-square limit* $Y(t)$ if

$$\underset{t' \to t}{\text{limit}}\ \big\|X(t') - Y(t)\big\|^2 = \underset{t' \to t}{\text{limit}}\ E\Big[\{X(t') - Y(t)\}^2\Big] = \underset{\varepsilon \to 0}{\text{limit}}\ E\Big[\{X(t+\varepsilon) - Y(t)\}^2\Big] = 0\,. \tag{12-8}$$

To express this, we use the l.i.m notation introduced in Chapter 11. The symbolic statement

$$Y(t) = \underset{t' \to t}{\text{l.i.m}}\ X(t') = \underset{\varepsilon \to 0}{\text{l.i.m}}\ X(t+\varepsilon) \tag{12-9}$$

should be interpreted as meaning (12-8). In (12-9), we have employed a variable t' that approaches t; equivalently, we have used a variable ε that approaches 0, so that $t + \varepsilon$ approaches t. While mathematically equivalent, each of the two notation styles has its advantages and disadvantages, and we will use both styles in what follows.

Completeness and the Cauchy Convergence Criteria (Revisited)

Please review Theorem 11-6, the completeness theorem for the vector space of finite-second-moment random variables. This theorem states that the vector space of finite-second-moment random variables is *complete* in the sense that a sequence of finite-second-moment random variables converges to a *unique* limit if, and only if, sequence terms come arbitrarily "close together" (in the m.s. sense) as you "go out" in the sequence (this is the *Cauchy criteria*).

Theorem 11-6, stated for sequences of finite-second-moment random variables, has a counterpart for finite-power random processes. Let t_n, $n \ge 0$, be *any* sequence that approaches zero as *n* approaches infinity (otherwise, the sequence is arbitrary). Directly from Theorem 11-6, we can state that

$$\underset{n\to\infty}{\text{l.i.m}}\ X(t+t_n) \tag{12-10}$$

exists as a *unique* (in the mean-square sense) random process if, and only if, the double limit

$$\underset{\substack{n\to\infty\\ m\to\infty}}{\text{l.i.m}}\,[\,X(t+t_n)-X(t+t_m)\,]=0 \tag{12-11}$$

exists. Hence, we need not know the limit (12-10) to prove convergence of the sequence; instead, we can show (12-11), the terms come arbitrarily "close together" as we "go out" in the sequence. For random processes, the Cauchy Convergence Theorem is stated most often in the following manner.

Theorem 12-1 (Cauchy Convergence Theorem): Let X(t) be a real-valued, finite-power random process. The mean-square limit

$$Y(t) \equiv \underset{t'\to t}{\text{l.i.m}}\ X(t') = \underset{\varepsilon\to 0}{\text{l.i.m}}\ X(t+\varepsilon) \tag{12-12}$$

exists as a *unique* (in the mean-square sense) random process if, and only if,

$$\underset{\substack{t_1 \to t \\ t_2 \to t}}{\text{limit}}\, E\Big[\{X(t_1) - X(t_2)\}^2\Big] = \underset{\substack{\varepsilon_1 \to 0 \\ \varepsilon_2 \to 0}}{\text{limit}}\, E\Big[\{X(t+\varepsilon_1) - X(t+\varepsilon_2)\}^2\Big] = 0\,. \tag{12-13}$$

In terms of the l.i.m notation, Equation (12-13) can be stated as

$$\underset{\substack{t_1 \to t \\ t_2 \to t}}{\text{l.i.m}}\, [\, X(t_1) - X(t_2)\,] = \underset{\substack{\varepsilon_1 \to 0 \\ \varepsilon_2 \to 0}}{\text{l.i.m}}\, [\, X(t+\varepsilon_1) - X(t+\varepsilon_2)\,] = 0\,. \tag{12-14}$$

The result Y(t) defined by (12-12) is not needed to know that the limit exists; m.s. limit (12-12) exists if, and only if, (12-14) holds. When using this result, one should remember that (12-13) and (12-14) must hold regardless of *how* t_1 and t_2 approach t (alternatively, ε_1 and ε_2 approach zero); this requirement is implied by the Calculus definition of a limit. The Cauchy Convergence Theorem plays a central role in the mean-square calculus of finite-power random processes.

Continuity of the Expectation Inner Product (Revisited)

Theorem 11-9 was stated for sequences, but it has a counterpart when dealing with finite-power random processes. Suppose X(t) and Y(t) are finite-power random processes with

$$\begin{aligned} &\underset{t' \to t_1}{\text{l.i.m}}\, X(t') = X_0(t_1) \\ &\underset{t' \to t_2}{\text{l.i.m}}\, Y(t') = Y_0(t_2). \end{aligned} \tag{12-15}$$

Then we have

$$E\big[X_0(t_1)Y_0(t_2)\big] = E\Bigg[\underset{t_1'\to t_1}{\text{l.i.m}}\ X(t_1')\cdot \underset{t_2'\to t_2}{\text{l.i.m}}\ Y(t_2')\Bigg] = \underset{\substack{t_1'\to t_1\\ t_2'\to t_2}}{\text{limit}}\ E\big[X(t_1')Y(t_2')\big]. \tag{12-16}$$

Written using inner product notation, Equation (12-16) is equivalent to

$$\big\langle X_0(t_1)Y_0(t_2)\big\rangle = \Big\langle \underset{t_1'\to t_1}{\text{l.i.m}}\ X(t_1')\cdot \underset{t_2'\to t_2}{\text{l.i.m}}\ X(t_2')\Big\rangle = \underset{\substack{t_1'\to t_1\\ t_2'\to t_2}}{\text{limit}}\ \big\langle X(t_1')Y(t_2')\big\rangle. \tag{12-17}$$

As was pointed out in the coverage given Theorem 11-9, the inner product is continuous. In the context of random processes, *continuity of the expectation* is expressed by Equations (12-15) and (12-16) (the expectation of a product is referred to as an *inner product* so we can say that the *inner product is continuous*). As t_1' and t_2' get "near" t_1 and t_2, respectively (so that $X(t_1')$ and $Y(t_2')$ get "near" $X_0(t_1)$ and $Y_0(t_2)$, respectively), we have $E[X(t_1')Y(t_2')]$ coming "near" $E[X_0(t_1)Y_0(t_2)]$.

Existence of the Correlation Function

A random process that satisfies (12-1) has a correlation function $\Gamma(t_1,t_2) = E[X(t_1)X(t_2)]$ that exists and is finite. We state this fact as the following theorem.

Theorem 12-2: For all t, zero-mean process X(t) has finite average power (*i.e.*, $X(t) \in \mathbf{L_2}$) if, and only if, its correlation function $\Gamma(t_1,t_2) = E[X(t_1)X(t_2)]$ exists as a finite quantity.

Proof: Suppose zero mean X(t) has finite average power (*i.e.*, satisfies (12-1)). Use the Cauchy-Schwarz inequality to see

$$\big|\Gamma(t_1,t_2)\big| = \big|E[X(t_1)X(t_2)]\big| \le \sqrt{E[X^2(t_1)]\,E[X^2(t_2)]} < \infty, \tag{12-18}$$

so $\Gamma(t_1,t_2)$ exists as a finite quantity. Conversely, suppose that $\Gamma(t_1,t_2)$ exists as a finite quantity. As a result of (12-18), we have

$$E[X^2(t)] = E[X(t)X(t)] = \Gamma(t,t) < \infty, \tag{12-19}$$

so that X(t) has finite average power.♥

This theorem is important since it assures us that finite-power random processes are synonymous with those that possess correlation functions.

Continuity of Random Processes

For random processes, the concept of continuity is based on the existence of a limit, just like the concept of continuity for "ordinary", non-random functions. However, as discussed previously, for random processes, the required limit can be defined in several ways (*i.e*, everywhere, almost everywhere, in probability, mean square sense, etc). In what follows, we give simple definitions for several types of continuity before concentrating on the type of continuity that is most useful in applications, mean-square continuity.

Sample-function continuity (a.k.a. *continuity* or *continuity everywhere*) at time t requires that *each and every* sample function be continuous at time t. We say that X(t) is *sample function continuous* at time t if

$$\underset{t'\to t}{\text{limit}}\, X(t',\rho) \equiv \underset{\varepsilon\to 0}{\text{limit}}\, X(t+\varepsilon,\rho) = X(t,\rho) \tag{12-20}$$

for ***all*** $\rho \in \mathbf{S}$. This is the "strongest" type of continuity possible. It is too restrictive for many applications.

A "weaker", less restrictive, form of continuity can be obtained by "throwing out" a set of sample functions that are associated with an event whose probability is zero. We say that the random process X(t,ρ) is *almost surely sample function* continuous (a.k.a. *continuous almost everywhere*) at time t if (12-20) holds everywhere except on an event whose probability is zero. That is, X(t,ρ) is almost surely sample function continuous if there exists an event *A*, **P**(*A*) = 1, for which (12-20) holds for all $\rho \in A$. This form of continuity requires the use of *measure* *theory*, an area of mathematics that most engineers are not conversant with. In addition, it is too restrictive for most applications, and it is not needed where only statistical averages are of interest, not individual sample functions.

Continuity in probability, or *p-continuity*, is based on the limit-in-probability concept that was introduced in Chapter 11, and it is even weaker that a.s. continuity. We say that X(t,ρ) is *p-continuous* at time t if

$$\underset{t'\to t}{\text{limit}}\,\mathbf{P}\Big[\,\big|X(t')-X(t)\big|>\alpha\,\Big]=\underset{\varepsilon\to 0}{\text{limit}}\,\mathbf{P}\Big[\,\big|X(t+\varepsilon)-X(t)\big|>\alpha\,\Big]=0 \tag{12-21}$$

for all $\alpha > 0$.

Mean Square Continuity

A stochastic process X(t) is *mean-square (m.s.) continuous* at time t if

$$X(t)=\underset{t'\to t}{\text{l.i.m}}\,X(t')\equiv\underset{\varepsilon\to 0}{\text{l.i.m}}\,X(t+\varepsilon)\,, \tag{12-22}$$

which is equivalent to

$$\underset{t'\to t}{\text{l.i.m}}\big[X(t')-X(t)\big]\equiv\underset{\varepsilon\to 0}{\text{l.i.m}}\big[X(t+\varepsilon)-X(t)\big]=0 \tag{12-23}$$

or

$$\underset{t'\to t}{\text{limit}}\,\big\|X(t')-X(t)\big\|^2=\underset{t'\to t}{\text{limit}}\;E\Big[\{X(t')-X(t)\}^2\Big]\equiv\underset{\varepsilon\to 0}{\text{limit}}\;E\Big[\{X(t+\varepsilon)-X(t)\}^2\Big]=0\,. \tag{12-24}$$

Mean square continuity does not imply continuity at the sample function level. A simple test for mean-square continuity involves the correlation function of the process.

Theorem 12-3: At time t, random process X(t) is mean-square continuous if, and only if, correlation $\Gamma(t_1,t_2)$ is continuous at $t_1 = t_2 = t$.

A simple proof of this theorem can be based on Theorem 12-1, the Cauchy Convergence Theorem. Basically, the requirement

$$X(t) = \underset{t'\to t}{\text{l.i.m}}\, X(t') \tag{12-25}$$

for m.s. continuity is equivalent to the Cauchy convergence requirement (12-13). Hence, the proof of Theorem 12-3 boils down to establishing that (12-13) is equivalent to $\Gamma(t_1,t_2)$ being continuous at $t_1 = t_2 = t$. While this is easy to do, we take a different approach while proving the theorem.

Proof of Theorem 12-3: First, we show continuity of $\Gamma(t_1,t_2)$ at $t_1 = t_2 = t$ is sufficient for m.s. continuity of X(t) at time t (*i.e.,* the "if" part). Consider the algebra

$$\begin{aligned} E\Big[\{X(t') - X(t)\}^2\Big] &= E\Big[X(t')^2\Big] - E\big[X(t)X(t')\big] - E\big[X(t')\,X(t)\big] + E\Big[X(t)^2\Big] \\ &= \Gamma(t',t') - \Gamma(t,t') - \Gamma(t',t) + \Gamma(t,t)\,. \end{aligned} \tag{12-26}$$

If $\Gamma(t_1,t_2)$ is continuous at $t_1 = t_2 = t$, the right-hand-side of (12-26) has zero as a limit (as $t' \to t$) so that

$$\underset{t'\to t}{\text{limit}}\, E\Big[\{X(t') - X(t)\}^2\Big] = \underset{t'\to t}{\text{limit}}\big[\Gamma(t',t') - \Gamma(t,t') - \Gamma(t',t) + \Gamma(t,t)\big] = 0\,, \tag{12-27}$$

and the process is m.s. continuous (this establishes the "if" part). Next, we establish necessity (the "only if" part). Assume that X(t) is m.s. continuous at t so that (12-23) is true. Consider the algebra

$$\Gamma(t_1,t_2)-\Gamma(t,t) = E\big[X(t_1)X(t_2)\big]-E\big[X(t)X(t)\big]$$

$$= E\big[\{X(t_1)-X(t)\}\{X(t_2)-X(t)\}\big] \tag{12-28}$$

$$+E\big[\{X(t_1)-X(t)\}X(t)\big]+E\big[X(t)\{X(t_2)-X(t)\}\big]$$

which implies

$$\left|\Gamma(t_1,t_2)-\Gamma(t,t)\right|$$

$$\le \left|\, E\big[\{X(t_1)-X(t)\}\{X(t_2)-X(t)\}\big]\,\right| + \left|\, E\big[\{X(t_1)-X(t)\}X(t)\big]\,\right| \tag{12-29}$$

$$+\left|\, E\big[X(t)\{X(t_2)-X(t)\}\big]\,\right|.$$

Apply the Cauchy-Schwarz inequality to each term on the right-hand-side of (12-29) to obtain

$$\left|\Gamma(t_1,t_2)-\Gamma(t,t)\right| \le \Big(E[\{X(t_1)-X(t)\}^2]\Big)^{1/2}\Big(E[\{X(t_2)-X(t)\}^2]\Big)^{1/2}$$

$$+\Big(E[\{X(t_1)-X(t)\}^2]\Big)^{1/2}\Big(E[X(t)^2]\Big)^{1/2} \tag{12-30}$$

$$+\Big(E[\{X(t_2)-X(t)\}^2]\Big)^{1/2}\Big(E[X(t)^2]\Big)^{1/2}.$$

Since X is m.s. continuous at t, the right-hand-side of (12-30) approaches zero as t_1, t_2 approach t. Hence, if X is mean-square continuous at t then

$$\underset{t_1,t_2\to t}{\text{limit}}\ \Gamma(t_1,t_2) = \Gamma(t,t), \tag{12-31}$$

so that $\Gamma(t_1,t_2)$ is continuous at $t_1 = t_2 = t$.♥

Example 12-1: Consider the Wiener process X(t), $t \geq 0$. In Chapter 6, the Wiener process was shown to be the formal limit of the random walk (a random walk with an infinitely dense collection of infinitesimal steps). As shown in Chapter 7, the correlation function of the Wiener process is given by

$$\Gamma(t_1, t_2) = 2D\min(t_1, t_2),$$

t_1, $t_2 \geq 0$, where D is the diffusion constant. Consider the inequality

$$\left|\Gamma(t_1, t_2) - \Gamma(t, t)\right| = 2D\left|\ \min(t_1, t_2) - t\ \right| \leq 2D\max(\ \left|t_1 - t\right|, \left|t_2 - t\right|\ ). \tag{12-32}$$

The right-hand-side of (12-32) approaches zero as t_1, $t_2 \to t$. As a result, we have $\Gamma(t_1, t_2)$ continuous at $t_1 = t_2 = t$, and the Wiener process is m.s. continuous at every t. Actually, a more advanced analysis would show that the Wiener process is almost surely sample function continuous.

Theorem 12-3 can be specialized to the case of wide sense stationary (W.S.S) processes. Recall that a W.S.S. process is characterized by the fact that its mean is constant and its correlation function depends on the time difference $\tau = t_1 - t_2$.

Corollary 12-3A: Suppose that X(t) is wide sense stationary so that $\Gamma(t_1,t_2) = \Gamma(\tau)$, where $\tau = t_1 - t_2$. Then, X(t) is mean-square continuous at time t if, and only if, $\Gamma(\tau)$ is continuous at $\tau = 0$.

Example 12-2: The *Random Binary Waveform* was introduced in Chapter 7. Recall that this process takes on the values of ±A; every t_a seconds, the process switches state with probability ½. Figure 7-1 depicts a typical sample function of the process. The process has a mean of zero, a correlation function given by

$$\begin{aligned} \Gamma(\tau) &= A^2\left[\frac{t_a - |\tau|}{t_a}\right], \quad |\tau| < t_a \\ &= 0 \qquad\qquad\qquad |\tau| > t_a, \end{aligned} \tag{12-33}$$

(a result depicted by Fig. 7-3), and it is W.S.S. As shown by Fig. 7-1, the sample functions have jump discontinuities (see Fig. 7-1). However, $\Gamma(\tau)$ is continuous at $\tau = 0$, so the process is mean-square continuous. This example illustrates the fact that m.s. continuity is "weaker" than sample-function continuity.

Example 12-3: The *sample and held* random process has m.s. discontinuities at every switching epoch. Consider passing a zero-mean, wide-sense-stationary random process V(t) through a sample-and-hold device that utilizes a T-second cycle time. Figure 12-1 illustrates an example of this; the dotted wave form is the original process V(t), and the piece-wise constant wave form is the output X(t). To generate output X(t), a sample is taken every T seconds, and it is "held" for T seconds until the next sample is taken. Such a process can be expressed as

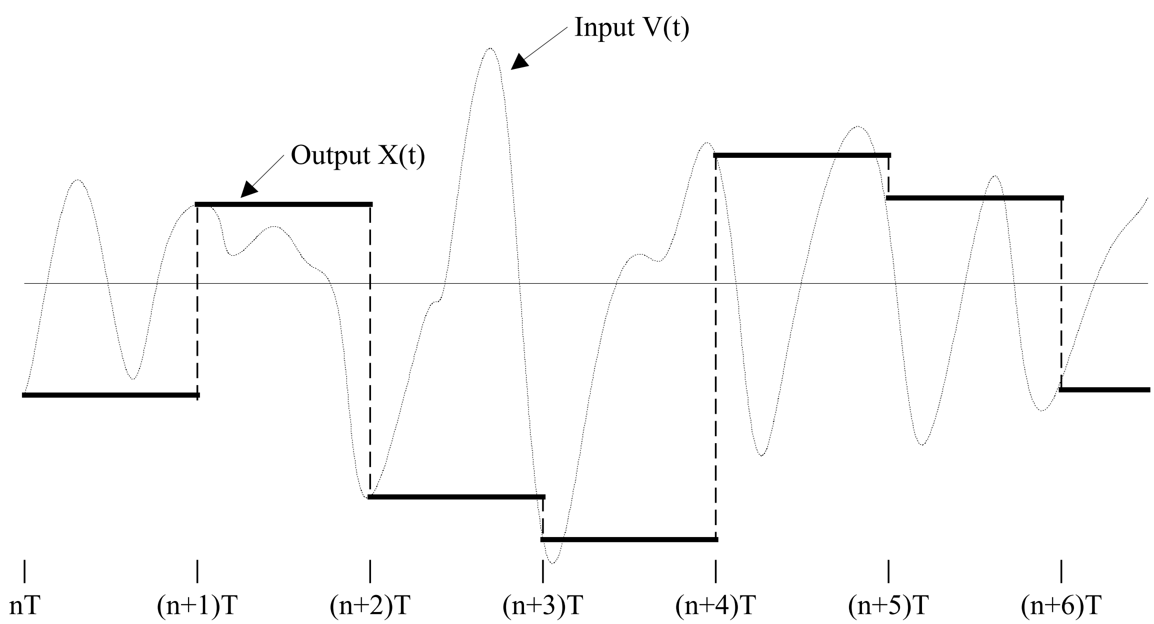

Figure 12-1: Dotted-line process is zero mean and constant variance V(t). Solid line (piece-wise constant) process is called the Sample and Held random process X(t).

$$X(t) = \sum_{n=-\infty}^{\infty} V(nT)q(t-nT)\,, \tag{12-34}$$

where

$$q(t) \equiv \begin{cases} 1, & 0 \le t < T \\ 0, & \text{otherwise} \end{cases} \tag{12-35}$$

Assume that T is large compared to the correlation time of process V(t). So, samples of the process are uncorrelated if they are spaced T seconds (or larger) apart in time. If $\sigma^2 = E[V^2(t)]$ is the constant variance of input waveform V(t), the correlation function for output X(t) is

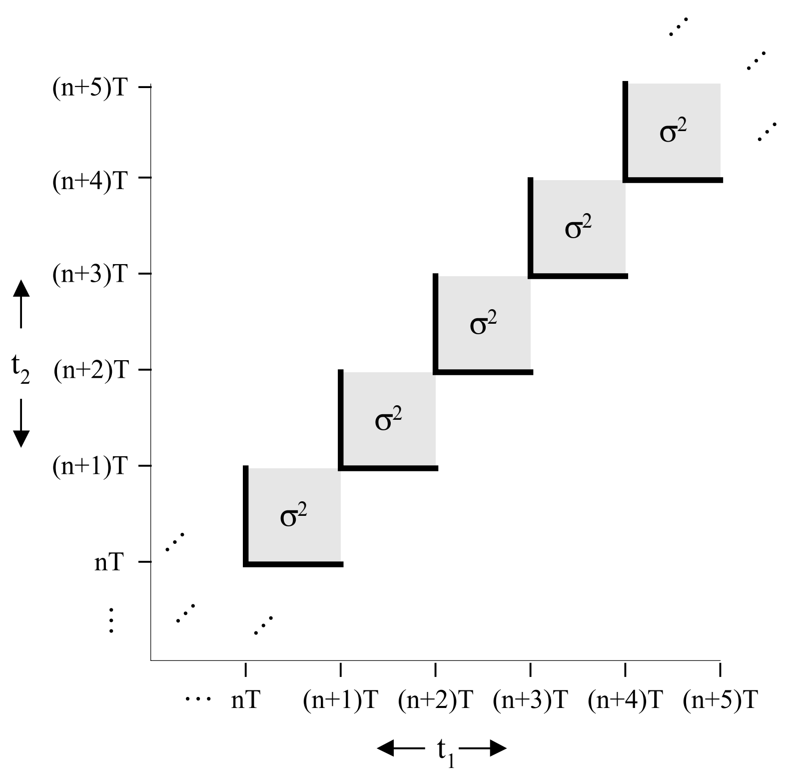

Figure 12-2: Correlation of sample and held random process. Correlation is σ^2 on half-open, T×T squares placed along the diagonal, and it is zero off of these squares.

$$\Gamma(t_1,t_2) = \sum_{n=-\infty}^{\infty} \sigma^2 q(t_1 - nT)q(t_2 - nT)\,, \tag{12-36}$$

a result depicted by Fig. 12-2. The correlation is equal to σ^2 on the half-open, T×T squares that lie along the diagonal, and it is zero off of these squares (in particular, the correlation is σ^2 along the diagonal $t_1 = t_2$). It is obvious from an inspection of Fig. 12-2 that $\Gamma(t_1,t_2)$ is continuous at every diagonal point except $t_1 = t_2 = nT$, n an integer. Hence, by Theorem 12-3, X(t) is m.s. continuous for t not a switching epoch.

Corollary 12-3B: If the correlation function $\Gamma(t_1,t_2)$ is continuous for all $t_1 = t_2 = t$ (*i.e.*, at all points on the line $t_1 = t_2$), then it is continuous at every point $(t_1, t_2) \in \mathcal{R}^2$.

Proof: Suppose $\Gamma(t_1, t_2)$ is continuous for all $t_1 = t_2 = t$. Then Theorem 12-3 tell us that X(t) is m.s. continuous for all t. Hence, for *any* t_1 and *any* t_2 we have

$$\begin{aligned} &\underset{\varepsilon\to 0}{\text{l.i.m}}\, X(t_1+\varepsilon) = X(t_1) \\ &\underset{\varepsilon\to 0}{\text{l.i.m}}\, X(t_2+\varepsilon) = X(t_2) \end{aligned} \tag{12-37}$$

Now, use Equation (12-16), the continuity of the inner product, to write

$$\begin{aligned} \Gamma(t_1,t_2) = E\big[X(t_1)X(t_2)\big] &= E\left[\underset{\varepsilon_1\to 0}{\text{l.i.m}}\, X(t_1+\varepsilon_1)\cdot \underset{\varepsilon_2\to 0}{\text{l.i.m}}\, X(t_2+\varepsilon_2)\right] \\ &= \underset{\varepsilon_1,\varepsilon_2\to 0}{\text{limit}}\; E\big[X(t_1+\varepsilon_1)X(t_2+\varepsilon_2)\big] \\ &= \underset{\varepsilon_1,\varepsilon_2\to 0}{\text{limit}}\; \Gamma(t_1+\varepsilon_1, t_2+\varepsilon_2)\,, \end{aligned} \tag{12-38}$$

so Γ is continuous at $(t_1, t_2) \in \mathcal{R}^2$.

The mean $\eta(t) = E[X(t)]$ of a process is a deterministic, non-random, function of time. It can be time varying. *If* the process is m.s. continuous, then its mean is a continuous function of time.

Theorem 12-4: Let X(t) be a mean-square continuous random process. Under this condition, the mean $\eta(t) = E[X(t)]$ is a continuous function of time.

Proof: Let X(t) be mean-square continuous and examine the non-negative variance of the process increment $X(t')-X(t)$ given by

$$\mathrm{Var}\big[X(t') - X(t)\big] = E\Big[\{X(t') - X(t)\}^2\Big] - \big(E\big[X(t') - X(t)\big]\big)^2 \geq 0 \tag{12-39}$$

From inspection of this result, we can write

$$E\Big[\{X(t') - X(t)\}^2\Big] \geq \big(E\big[X(t') - X(t)\big]\,\big)^2 = \big(\eta(t') - \eta(t)\big)^2 . \tag{12-40}$$

Let t′ approach t in this last equation; due to m.s. continuity, the left-hand-side of (12-40) must approach zero. This implies that

$$\underset{t' \to t}{\text{limit}}\ \eta(t') = \eta(t)\,, \tag{12-41}$$

which is equivalent to saying that the mean η is continuous at time t.♥

Mean-square continuity is " stronger" that p-continuity. That is, mean-square continuous random processes are also p-continuous (the converse is not true). We state this claim with the following theorem.

Theorem 12-5: If a random process is m.s. continuous at t then it is p-continuous at t.

Proof: A simple application of the Chebyshev inequality yields

$$\mathbf{P}\Big[\left|X(t') - X(t)\right| > a\Big] \le \frac{E\Big[\{X(t') - X(t)\}^2\Big]}{a^2} \tag{12-42}$$

for every $a > 0$. Now, let t' approach t, and note that the right-hand-side of (12-42) approaches zero. Hence, we can conclude that

$$\underset{t' \to t}{\text{limit}}\, \mathbf{P}\Big[\left|X(t') - X(t)\right| > a\Big] = 0\,, \tag{12-43}$$

and X is p-continuous at t (see definition (12-21)).♥

Δ Operator

To simplify our work, we introduce some shorthand notation. Let f(t) be any function, and define the *difference operator*

$$\Delta_\varepsilon\, f(t) \equiv f(t+\varepsilon) - f(t)\,. \tag{12-44}$$

On the Δ_ε operator, the subscript ε is the size of the time increment.

We extend this notation to functions of two variables. Let f(t,s) be a function of t (the first variable) and s (the second variable). We define

$$\begin{aligned} \Delta_\varepsilon^{(1)} f(t,s) &\equiv f(t+\varepsilon, s) - f(t,s) \\ \Delta_\varepsilon^{(2)} f(t,s) &\equiv f(t, s+\varepsilon) - f(t,s) \end{aligned}\,. \tag{12-45}$$

On the difference operator, a superscript of (1) (alternatively, a superscript of (2)) denotes that we difference the first variable (alternatively, the second variable).

Mean Square Differentiation

A stochastic process X(t) has a mean square (m.s.) derivative, denoted here as $\dot{X}(t)$, *if* there exists a finite-power random process

$$\dot{X}(t) = \underset{\varepsilon\to 0}{\mathrm{l.i.m}}\left[\frac{X(t+\varepsilon)-X(t)}{\varepsilon}\right] = \underset{\varepsilon\to 0}{\mathrm{l.i.m}}\left[\frac{\Delta_\varepsilon X(t)}{\varepsilon}\right] \tag{12-46}$$

(*i.e.*, if the l.i.m exists). Equation (12-46) is equivalent to

$$\underset{\varepsilon\to 0}{\mathrm{limit}}\ E\left[\left(\frac{X(t+\varepsilon)-X(t)}{\varepsilon} - \dot{X}(t)\right)^2\right] = \underset{\varepsilon\to 0}{\mathrm{limit}}\ E\left[\left(\frac{\Delta_\varepsilon X(t)}{\varepsilon} - \dot{X}(t)\right)^2\right] = 0. \tag{12-47}$$

A necessary and sufficient condition is available for X(t) to be m.s. differentiable. Like the case of m.s. continuity considered previously, the requirement for m.s. differentiability involves a condition on $\Gamma(t_1,t_2)$. The condition for differentiability is based on Theorem 12-1, the Cauchy Convergence Theorem. As $\varepsilon \to 0$, we require existence of the l.i.m of $\{\Delta_\varepsilon X(t)\}/\varepsilon$ for the m.s. derivative to exist. For each arbitrary but fixed t, the quotient $\Delta_\varepsilon X(t)/\varepsilon$ *is a random process that is a function of* ε. So, according to the Cauchy Convergence Theorem, the quantity $\{\Delta_\varepsilon X(t)\}/\varepsilon$ has a m.s. limit (as ε goes to zero) if, and only if,

$$\underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\mathrm{l.i.m}}\left(\frac{\Delta_{\varepsilon_1}X(t)}{\varepsilon_1} - \frac{\Delta_{\varepsilon_2}X(t)}{\varepsilon_2}\right) = 0. \tag{12-48}$$

Note that (12-48) is equivalent to

$$\underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\, E\left[\left(\frac{\Delta_{\varepsilon_1}X(t)}{\varepsilon_1}-\frac{\Delta_{\varepsilon_2}X(t)}{\varepsilon_2}\right)^2\right]$$

$$= \underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\, E\left[\left(\frac{\Delta_{\varepsilon_1}X(t)}{\varepsilon_1}\right)^2 - 2\left(\frac{\Delta_{\varepsilon_1}X(t)}{\varepsilon_1}\right)\left(\frac{\Delta_{\varepsilon_2}X(t)}{\varepsilon_2}\right)+\left(\frac{\Delta_{\varepsilon_2}X(t)}{\varepsilon_2}\right)^2\right] \tag{12-49}$$

$$= 0.$$

In (12-49), there are **two** terms that can be evaluated as

$$\begin{aligned}\underset{\varepsilon\to 0}{\text{limit}}\, E\left[\left(\frac{\Delta_{\varepsilon}X(t)}{\varepsilon}\right)^2\right] &= \underset{\varepsilon\to 0}{\text{limit}}\, E\left[\left(\frac{X(t+\varepsilon)-X(t)}{\varepsilon}\right)^2\right]\\ &= \underset{\varepsilon\to 0}{\text{limit}}\frac{\Gamma(t+\varepsilon,t+\varepsilon)-\Gamma(t+\varepsilon,t)-\Gamma(t,t+\varepsilon)+\Gamma(t,t)}{\varepsilon^2},\end{aligned} \tag{12-50}$$

and a cross term that evaluates to

$$\begin{aligned}\underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\, E\left[\left(\frac{\Delta_{\varepsilon_1}X(t)}{\varepsilon_1}\right)\left(\frac{\Delta_{\varepsilon_2}X(t)}{\varepsilon_2}\right)\right] &= \underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\, E\left[\left(\frac{X(t+\varepsilon_1)-X(t)}{\varepsilon_1}\right)\left(\frac{X(t+\varepsilon_2)-X(t)}{\varepsilon_2}\right)\right]\\ &= \underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\frac{\Gamma(t+\varepsilon_1,t+\varepsilon_2)-\Gamma(t+\varepsilon_1,t)-\Gamma(t,t+\varepsilon_2)+\Gamma(t,t)}{\varepsilon_1\varepsilon_2}\end{aligned} \tag{12-51}$$

Now, substitute (12-50) and (12-51) into (12-49), and observe that (12-48) is equivalent to

$$\underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\, E\left[\left(\frac{\Delta_{\varepsilon_1}X(t)}{\varepsilon_1}-\frac{\Delta_{\varepsilon_2}X(t)}{\varepsilon_2}\right)^2\right]$$

$$= 2\,\underset{\varepsilon\to 0}{\text{limit}}\frac{\Gamma(t+\varepsilon,t+\varepsilon)-\Gamma(t+\varepsilon,t)-\Gamma(t,t+\varepsilon)+\Gamma(t,t)}{\varepsilon^2} \qquad (12\text{-}52)$$

$$-2\,\underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\frac{\Gamma(t+\varepsilon_1,t+\varepsilon_2)-\Gamma(t+\varepsilon_1,t)-\Gamma(t,t+\varepsilon_2)+\Gamma(t,t)}{\varepsilon_1\varepsilon_2}$$

$$= 0$$

As is shown by the next theorem, (12-48), and its equivalent (12-52), can be stated as a differentiability condition on $\Gamma(t_1,t_2)$.

Theorem 12-6: A finite-power stochastic process X(t) is mean-square differentiable at t if, and only if, the double limit

$$\underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\frac{\Delta^{(1)}_{\varepsilon_1}\Delta^{(2)}_{\varepsilon_2}\Gamma(t,t)}{\varepsilon_1\varepsilon_2} = \underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\frac{\Gamma(t+\varepsilon_1,t+\varepsilon_2)-\Gamma(t+\varepsilon_1,t)-\Gamma(t,t+\varepsilon_2)+\Gamma(t,t)}{\varepsilon_1\varepsilon_2}. \qquad (12\text{-}53)$$

exists and is finite (*i.e.,* exists as a real number). Note that (12-53) is the second limit that appears on the right-hand side of (12-52). By some authors, (12-53) is called the 2^{nd} generalized derivative.

Proof: Assume that the process is m.s. differentiable. Then (12-48) and (12-52) hold; independent of the manner in which ε_1 and ε_2 approach zero, the double limit in (12-52) is zero. But this means that the limit (12-53) exists, so m.s. differentiability of X(t) implies the existence of (12-53). Conversely, assume that limit (12-53) exist and has the value R_2 (this must be true independent of how ε_1 and ε_2 approach zero). Then, the first limit on the right-hand-side of (12-52) is also R_2, and this implies that (12-52) (and (12-48)) evaluate to zero because $R_2 - 2R_2 + R_2 = 0$.♥

Note on Theorem 12-6: Many books on random processes include a common error in their statement of this theorem (and a few authors point this out). Instead of requiring the existence of (12-53), many authors state that X(t) is m.s. differentiable at t ***if*** (or ***iff*** –according to some authors) $\partial^2\Gamma(t_1,t_2)/\partial t_1\partial t_2$ exists at $t_1 = t_2 = t$. Strictly speaking, this is incorrect. The existence of (12-53) *implies* the existence of $\partial^2\Gamma(t_1,t_2)/\partial t_1\partial t_2$ at $t_1 = t_2 = t$, the converse is **not** true. Implicit in (12-53) is the requirement of path independence; how ε_1 and ε_2 approach zero should not influence the result produced by (12-53). However, the second-order mixed partial is not defined in such general terms. Using the Δ notation introduced by (12-45), the second-order mixed partial is defined as

$$\frac{\partial^2\Gamma}{\partial t_1\partial t_2} = \frac{\partial}{\partial t_1}\left(\frac{\partial\Gamma}{\partial t_2}\right) = \underset{\varepsilon_1\to 0}{\text{limit}}\,\frac{\Delta_{\varepsilon_1}^{(1)}\left[\underset{\varepsilon_2\to 0}{\text{limit}}\left(\dfrac{\Delta_{\varepsilon_2}^{(2)}\Gamma(t_1,t_2)}{\varepsilon_2}\right)\right]}{\varepsilon_1}. \tag{12-54}$$

Equation (12-54) requires that $\varepsilon_2 \to 0$ **first** to obtain an intermediate, dependent-on-ε_1, result; then, the **second** limit $\varepsilon_1 \to 0$ is taken. The existence of (12-53) at a point implies the existence of (12-54) at the point. The converse is **not** true; the existence of (12-54) **does not** imply the existence of (12-53). The following example shows that existence of the second partial derivative (of the form (12-54)) is not sufficient for the existence of limit (12-53) and the m.s. differentiability of X(t).

Example 12-4 (from T. Soong, *Random Differential Equations in Science and Engineering*, p. 93)**:** Consider the finite-power stochastic process X(t), $-1 \le t \le 1$, defined by

$$\begin{aligned}
X(0) &= 0 \\
X(t) &= \alpha_k, \qquad 1/2^k < t \le 1/2^{k-1},\ k = 1, 2, 3, \ldots \qquad\qquad (12\text{-}55)\\
&= X(-t), \qquad -1 \le t \le 0
\end{aligned}$$

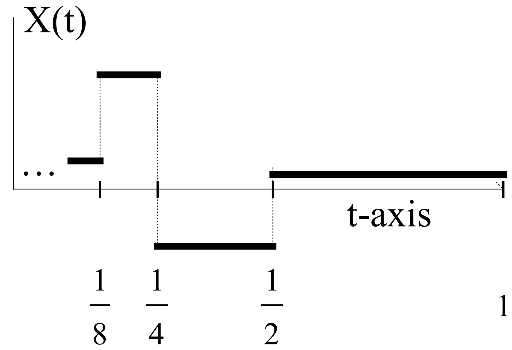

Figure 12-3: For $0 \le t \le 1$, a typical sample function of X(t).

where the α_k are independent, identically distributed random variables each with a mean of zero and a variance of unity. For $t \ge 0$, Figure 12-3 depicts a typical sample function of such a process (fold the graph to get X for negative time).

Process X(t) has, for $t \ge 0$, $s \ge 0$, a correlation function $\Gamma(t,s)$ that is depicted by Figure 12-4 (this plot can be used to obtain the value of Γ for (t,s) in the second, third and fourth quadrants of the (t,s) plane). As depicted on Figure 12-4, in the first quadrant, the correlation

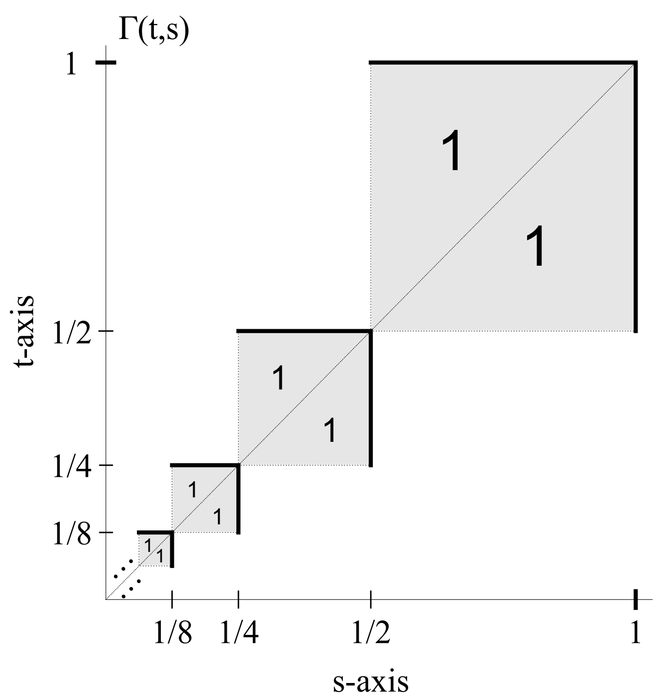

Figure 12-4: Correlation function $\Gamma(t,s)$ is unity on shaded , half-closed rectangles and zero otherwise.

function is unity on the shaded, half-closed squares, and it is zero elsewhere in the first quadrant. Specifically, note that $\Gamma(t,t) = 1$, $0 < t \le 1$. Take the limit along the line $t = \varepsilon$, $s = \varepsilon$ to see that

$$\begin{aligned}\underset{\varepsilon\to 0}{\text{limit}}\ \frac{\Delta_\varepsilon^{(1)}\Delta_\varepsilon^{(2)}\Gamma(t,s)}{\varepsilon^2}\bigg|_{t=s=0} &= \underset{\varepsilon\to 0}{\text{limit}}\frac{\Gamma(\varepsilon,\varepsilon)-\Gamma(\varepsilon,0)-\Gamma(0,\varepsilon)+\Gamma(0,0)}{\varepsilon^2}\\ &= \underset{\varepsilon\to 0}{\text{limit}}\frac{\Gamma(\varepsilon,\varepsilon)-0-0+0}{\varepsilon^2}\\ &= \underset{\varepsilon\to 0}{\text{limit}}\ \frac{1}{\varepsilon^2} = \infty\end{aligned} \qquad . \tag{12-56}$$

By Theorem 12-6, X(t) *does not* have a mean-square derivative at t = 0 since (12-53) does not exist at t = s = 0, a conclusion drawn from inspection of (12-56). But for $-1 \le t \le 1$, it is easily seen that

$$\frac{\partial\Gamma(t,s)}{\partial s}\bigg|_{s=0} = 0, \qquad -1 \le t \le 1\ ,$$

so the second-order partial derivative exists at t = s = 0, and its value is

$$\begin{aligned}\frac{\partial^2\Gamma(t,s)}{\partial t\partial s}\bigg|_{t=s=0} &= \frac{\partial}{\partial t}\left[\frac{\partial\Gamma(t,s)}{\partial s}\bigg|_{s=0}\right]_{t=0}\\ &= 0.\end{aligned}$$

Example (12-4) shows that, at a point, the second partial derivative (*i.e.*, (12-54)) can exist and be finite, but limit (12-53) may not exit. Hence, it serves as a counter example to those authors that claim (incorrectly) that X(t) is m.s. differentiable at t if (or iff –according to some authors) $\partial^2\Gamma(t_1,t_2)/\partial t_1\partial t_2$ exists and is finite at $t_1 = t_2 = t$. However, as discussed next, this second-order partial can be used to state a sufficient condition for the existence of the m.s. derivative of X.

Theorem 12.7 (Sufficient condition for the existence of the m.s. derivative): If $\partial\Gamma/\partial t_1$, $\partial\Gamma/\partial t_2$, and $\partial^2\Gamma/\partial t_1\partial t_2$ exist in a neighborhood of $(t_1,t_2) = (t,t)$, ***and*** $\partial^2\Gamma/\partial t_1\partial t_2$ is continuous at $(t_1,t_2) = (t,t)$, then limit (12-53) exits, and process X is m.s. differentiable at t.

Proof: Review your multivariable calculus. For example, consult Theorem 17.9.1 of L. Leithold, *The Calculus with Analytic geometry*, Second Edition. Also, consult page 79 of E. Wong, B. Hajek, *Stochastic Processes in Engineering Systems*.

One should recall that the mere existence of $\partial^2\Gamma/\partial t_1\partial t_2$ at point (t_1,t_2) does not imply that this second-order partial is continuous at point (t_1,t_2). Note that this behavior in the multidimensional case is in stark contrast to the function-of-one-variable case.

Example 12-5: Consider the Wiener process $X(t)$, $t \ge 0$, that was introduced in Chapter 6. For the case $X(0) = 0$, we saw in Chapter 7 that

$$\Gamma(t_1,t_2) = 2D\min\{t_1,t_2\}$$

for $t_1 \ge 0$, $t_2 \ge 0$. This correlation function does not have a second partial derivative at $t_1 = t_2 = t > 0$. To see this, consider Figure 12-5, a plot of $\Gamma(t_1,t_2)$ as a function of t_1 for fixed $t_2 = t_{20} > 0$. Hence, the Wiener process is not m.s. differentiable at any $t > 0$. This is not unexpected; in the limit, as the step size and time to take a step shrink to zero, the random walk becomes an increasingly dense sequence of smaller and smaller jumps. Heuristically, the Wiener process can be thought of as an infinitely dense sequence of infinitesimal jumps. Now, jumps are not

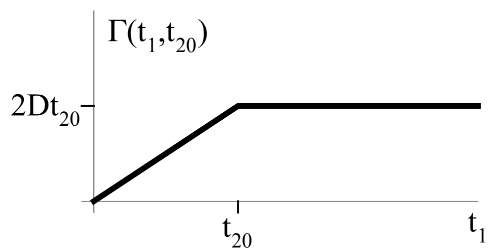

Figure 12-5: $\Gamma(t_1,t_2)$ for fixed $t_2 = t_{20}$.

differentiable, so it is not surprising that the Wiener process is not differentiable.

M.S. Differentiability for the Wide-Sense-Stationary Case

Recall that a W.S.S random process has a correlation function $\Gamma(t_1,t_2)$ that depends only on the time difference $\tau = t_1-t_2$. Hence, it is easily seen that W.S.S process X(t) is m.s. differentiable for *all* t if, and only if, it is m.s. differentiable for *any* t.

In the definition of the 2nd generalized derivative of autocorrelation $\Gamma(t_1, t_2)$ defined by (12-53), the path dependence issue does not arise in the WSS case since Γ only depends on the *single* variable $\tau \equiv t_1 - t_2$. Regardless of how t_1 and t_2 approach zero, the difference $\tau \equiv t_1 - t_2$ always approaches zero in only two ways, from the positive or negative real numbers. In the following development, we use the fact that a function f(x) must exist in a neighborhood of a point x_0 (including x_0) for the derivative df/dx to exist at x_0.

Corollary 12-6

Wide sense stationary, finite power X(t), with autocorrelation $\Gamma(\tau)$, is mean square differentiable at anytime t if and only if the 1st and 2nd derivatives of $\Gamma(\tau)$ exist and are finite at $\tau = 0$ (Since $\Gamma(\tau)$ is even, $\Gamma'(\tau)$ is odd so that $\Gamma'(0) = 0$).

`Proof`: For the WSS case, the second generalized derivative can be written as

$$\underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\frac{\Gamma(\varepsilon_1-\varepsilon_2)-\Gamma(\varepsilon_1)-\Gamma(-\varepsilon_2)+\Gamma(0)}{\varepsilon_1\varepsilon_2} = \underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\frac{\left\{\dfrac{\Gamma(\varepsilon_1-\varepsilon_2)-\Gamma(-\varepsilon_2)}{\varepsilon_1}\right\}-\left\{\dfrac{\Gamma(\varepsilon_1)-\Gamma(0)}{\varepsilon_1}\right\}}{\varepsilon_2}$$

$$= \underset{\substack{\varepsilon_1\to 0\\ \varepsilon_2\to 0}}{\text{limit}}\frac{\left\{\dfrac{\Gamma(\varepsilon_1-\varepsilon_2)-\Gamma(\varepsilon_1)}{\varepsilon_2}\right\}-\left\{\dfrac{\Gamma(-\varepsilon_2)-\Gamma(0)}{\varepsilon_2}\right\}}{\varepsilon_1}.$$

Since $\Gamma(\tau)$ is even, the right-hand side limits are equivalent, and the order in which the limit is taken is immaterial. Since second derivative $\Gamma''(\tau)$ exists at $\tau = 0$, the first derivative $\Gamma'(\tau)$ must exist in a neighborhood of $\tau = 0$. As a result, the above generalized derivative becomes

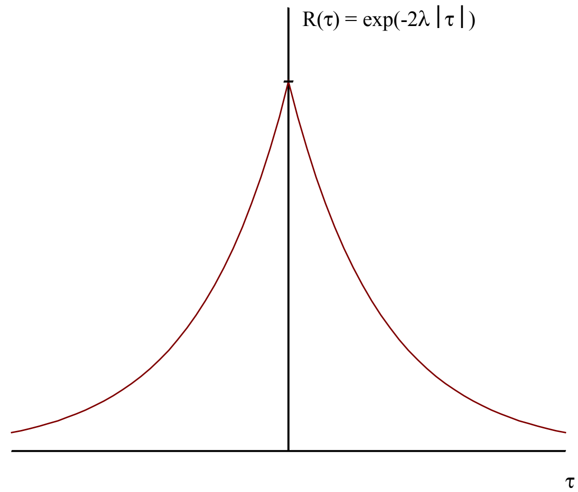

Figure 12-6: Correlation function of the W.S.S. random telegraph signal.

$$\lim_{\substack{\varepsilon_1 \to 0 \\ \varepsilon_2 \to 0}} \frac{\left\{\dfrac{\Gamma(\varepsilon_1 - \varepsilon_2) - \Gamma(-\varepsilon_2)}{\varepsilon_1}\right\} - \left\{\dfrac{\Gamma(\varepsilon_1) - \Gamma(0)}{\varepsilon_1}\right\}}{\varepsilon_2} = \lim_{\varepsilon_2 \to 0} \frac{\Gamma'(-\varepsilon_2) - \Gamma'(0)}{\varepsilon_2} = -\Gamma''(0)$$

Hence, for the WSS case: Γ' and Γ'' exist and are finite at $\tau = 0$

$\Leftrightarrow$ The second generalized derivative of Γ exists at all t

$\Leftrightarrow$ X(t) is m.s. differentiable at all t.

Example 12-6: Consider the random telegraph signal discussed in Chapter 7. Recall that this W.S.S process has the correlation function

$$\Gamma(t_1, t_2) = \Gamma(\tau) = e^{-2\lambda|\tau|}$$

that is depicted by Figure 12-6. Clearly, Γ is not differentiable at $\tau = 0$, so the random telegraph signal is not m.s. differentiable anywhere.

Some Properties of the Mean Square Derivative

Many of the properties of "ordinary" derivatives of deterministic functions have counter parts when it comes to m.s. derivatives of finite-power random processes. We give just a few of these.

Theorem 12-8: For a finite-power random process X(t), mean square differentiability at time t implies mean square continuity at time t.

Proof: Suppose that X(t) is m.s. differentiable at t. Consider

$$\begin{aligned}
\underset{\varepsilon\to 0}{\text{limit}}\, E\Big[\{X(t+\varepsilon)-X(t)\}^2\Big] &= \underset{\varepsilon\to 0}{\text{limit}}\, \varepsilon^2 \cdot E\left[\left(\frac{\{X(t+\varepsilon)-X(t)\}^2}{\varepsilon^2}\right)\right] \\
&= \underset{\varepsilon\to 0}{\text{limit}}\, \varepsilon^2 \left(\underset{\varepsilon\to 0}{\text{limit}}\, E\left[\left(\frac{\{X(t+\varepsilon)-X(t)\}^2}{\varepsilon^2}\right)\right]\right) \\
&= 0 \cdot \underset{\varepsilon\to 0}{\text{limit}} \frac{\Gamma(t+\varepsilon,t+\varepsilon)-\Gamma(t+\varepsilon,t)-\Gamma(t,t+\varepsilon)+\Gamma(t,t)}{\varepsilon^2} \\
&= 0,
\end{aligned} \tag{12-57}$$

since the limit involving Γ is finite (as given by Theorem 12-6). Hence, a m.s. differentiable function is also m.s. continuous.♥

Theorem 12-9: If $X_1(t)$ and $X_2(t)$ are finite-power random processes that are m.s. differentiable, then $\alpha X_1(t) + \beta X_2(t)$ is a finite power process that is m.s. differentiable for any real constants α and β. Furthermore, m.s. differentiation is a linear operation so that

$$\frac{d}{dt}\big[\alpha X_1(t)+\beta X_2(t)\big] = \alpha\frac{dX_1(t)}{dt}+\beta\frac{dX_2(t)}{dt}. \tag{12-58}$$

Mean and Correlation Function of dX/dt

In Theorem 11-7, we established that the operations of l.i.m and expectation were interchangeable for sequences of random variables. An identical result is available for finite-power random processes. Recall that if we have

$$X(t) = \underset{t' \to t}{\text{l.i.m}}\, X(t'),$$

then

$$E\big[X(t)\big] = E\Big[\underset{t' \to t}{\text{l.i.m}}\, X(t')\Big] = \underset{t' \to t}{\text{limit}}\; E\big[X(t')\big] \tag{12-59}$$

for finite-power random process X(t). This result can be used to obtain an expression for E[$\dot{X}$] in terms of E[X].

Theorem 12-10: Let X(t) be a finite-power, m.s. differentiable random process. The mean of the derivative is given by

$$E\big[\dot{X}(t)\big] = \frac{d}{dt} E[X(t)]. \tag{12-60}$$

In words, you can interchange the operations of expectation and differentiation for finite power, m.s. differentiable random processes.

Proof: Observe the simple steps

$$\begin{aligned} E\big[\dot{X}(t)\big] &= E\left[\underset{\varepsilon \to 0}{\text{l.i.m}}\, \frac{X(t+\varepsilon) - X(t)}{\varepsilon}\right] = \underset{\varepsilon \to 0}{\text{limit}}\; \frac{E[X(t+\varepsilon)] - E[X(t)]}{\varepsilon} \\ &= \frac{d}{dt} E[X(t)] \end{aligned}. \tag{12-61}$$

In this result, the interchange of l.i.m and expectation is justified by Theorem 11-7, as outlined above.♥

Theorem 12-11: Let finite-power process X(t) be m.s. differentiable for all t. Then for all t and s, the quantities $E[\dot{X}(t)X(s)]$, $E[X(t)\dot{X}(s)]$ and $E[\dot{X}(t)\dot{X}(s)]$ are finite and expressible in terms of Γ(t,s) = E[X(t)X(s)]. The relevant formulas are

$$\Gamma_{\dot{X}X}(t,s) \equiv E[\dot{X}(t)X(s)] = \frac{\partial\Gamma(t,s)}{\partial t}$$

$$\Gamma_{X\dot{X}}(t,s) \equiv E[X(t)\dot{X}(s)] = \frac{\partial\Gamma(t,s)}{\partial s} \tag{12-62}$$

$$\Gamma_{\dot{X}\dot{X}}(t,s) \equiv E[\dot{X}(t)\dot{X}(s)] = \frac{\partial^2\Gamma(t,s)}{\partial t\partial s}.$$

Proof: Use the Cauchy-Schwarz inequality to see that the quantities exist and are finite. Now, the first of (12-62) follows from

$$\begin{aligned}\Gamma_{\dot{X}X}(t,s) &\equiv E[\dot{X}(t)X(s)] = E\Big[\Big(\underset{\varepsilon\to 0}{\text{l.i.m}}\frac{X(t+\varepsilon)-X(t)}{\varepsilon}\Big)X(s)\Big] \\ &= \underset{\varepsilon\to 0}{\text{limit}}\ E\left[\frac{X(t+\varepsilon)X(s)-X(t)X(s)}{\varepsilon}\right] = \underset{\varepsilon\to 0}{\text{limit}}\ \frac{\Gamma(t+\varepsilon,s)-\Gamma(t,s)}{\varepsilon} \\ &= \frac{\partial\Gamma(t,s)}{\partial t}.\end{aligned} \tag{12-63}$$

The remaining three correlation functions are obtained in a similar manner.♥

For the wide sense stationary case, Theorem 12-11 can be simplified. If X is W.S.S., then (12-62) simplifies to

$$\Gamma_{\dot{X}X}(\tau) \equiv E[\dot{X}(t)X(t+\tau)] = \frac{\partial\Gamma(\tau)}{\partial\tau}$$

$$\Gamma_{X\dot{X}}(\tau) \equiv E[X(t)\dot{X}(t+\tau)] = -\frac{\partial\Gamma(\tau)}{\partial\tau} \quad . \tag{12-64}$$

$$\Gamma_{\dot{X}\dot{X}}(\tau) \equiv E[\dot{X}(t)\dot{X}(t+\tau)] = -\frac{\partial^2\Gamma(\tau)}{\partial\tau^2}.$$

Generalized Derivatives, Generalized Random Processes and White Gaussian Noise

There are many applications where a broadband, zero-mean Gaussian noise process drives a dynamic system described by a differential equation. Often, the Gaussian forcing function has a bandwidth that is large compared to the bandwidth of the system, and the spectrum of the Gaussian noise looks flat over the system bandwidth. In most cases, the analysis of such a problem is simplified by assuming that the noise spectrum is flat over all frequencies. In deference to the ideal that white light is composed of all colors, a random process with a flat spectrum is said to be *white*; if also Gaussian, it is said to be *white Gaussian noise*.

Clearly, white Gaussian noise cannot exist as a *physically realizable* random process. The reason for this is simple: a white process contains infinite power, and it is delta correlated. These requirements cannot be met by any physically realizable process.

Now, lets remember our original motivation. We want to analyze a system driven by a wide-band Gaussian process. To *simplify* our work, we chose an *approximate* analysis based on the use of Gaussian white noise as a system input. We are not concerned by the fact that such an input noise process does not exist physically. We know that Gaussian white noise *does exist* mathematically (*i.e.,* it can be described using rigorous mathematics) as a member of a class of processes known as *generalized random processes* (books have been written on this topic). After all, we use delta functions although we know that they are not physically realizable (the delta function exists mathematically as an example of what is known as a *generalized function*). We will be happy to use (as a mathematical concept) white Gaussian noise if doing so simplifies our system analysis and provides final results that agree well with physical reality. Now that Gaussian white noise has been accepted as a generalized random process, we seek its relationship to other processes that we know about.

Consider the Wiener process $X(t)$ described in Chapter 6. In Example 12-5, we saw that X is not m.s. differentiable at any time t. That is, the quotient

$$Y(\mathrm{t};\varepsilon) \equiv \frac{X(\mathrm{t}+\varepsilon)-X(\mathrm{t})}{\varepsilon} \tag{12-65}$$

does not have a m.s. limit as ε approaches zero. Also, almost surely, Wiener process sample functions are not differentiable (in the "ordinary" Calculus sense). That is, there exists A, $\mathbf{P}(A) = 1$, such that for each $\omega \in A$, $X(\mathrm{t},\omega)$ is not differentiable at any time t. However, when interpreted as a generalized random process (see discussion above), the Wiener process has a generalized derivative (as defined in the literature) that is white Gaussian noise (a generalized random process).

For fixed $\varepsilon > 0$, let us determine the correlation function $\Gamma_Y(\tau;\varepsilon)$ of $Y(\mathrm{t},\varepsilon)$. Use the fact that the Wiener process has independent increments to compute

$$\Gamma_Y(\tau;\varepsilon) = \begin{cases} 2D\left[\dfrac{\varepsilon-|\tau|}{\varepsilon^2}\right], & |\tau| \le \varepsilon \\ 0, & |\tau| > \varepsilon \end{cases}, \tag{12-66}$$

a result depicted by Figure 12-7 (can you show (12-66)?). Note that the area under Γ_Y is 2D, independent of ε. As ε approaches zero, the base width shrinks to zero, the height goes to infinity, and

$$\underset{\varepsilon\to 0^+}{\operatorname{limit}}\ \Gamma_Y(\tau;\varepsilon) = 2D\,\delta(\tau). \tag{12-67}$$

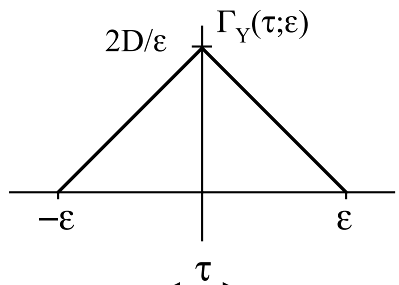

Figure 12-7: Correlation of $Y(\mathrm{t},\varepsilon)$. It approaches $2D\delta(\tau)$ as ε approaches zero.

That is, as ε becomes "small", $Y(t;\varepsilon)$ approaches a delta-correlated, white Gaussian noise process!

To summarize a complicated discussion, as $\varepsilon \to 0$, Equation (12-65) has no m.s. limit, so the Wiener process is not differentiable in the m.s. sense (or, as it is possible to show, in the sample function sense). However, when considered as a *generalized random process*, the Wiener process has a *generalized derivative* that turns out to be Gaussian white noise, another generalized random process. See Chapter 3 of *Stochastic Differential Equations: Theory and Applications*, by Ludwig Arnold, for a very readable discussion of the Wiener process and its generalized derivative, Gaussian white noise. The theory of generalized random processes tends to parallel the theory of generalized functions and their derivatives.

Example 12-7: We know that the Wiener process is not m.s. differentiable; in the sense of classical Calculus, the third equation in (12-62) cannot be applied to the autocorrelation function $\Gamma(t_1,t_2) = 2D\min\{t_1,t_2\}$ of the Wiener process. However, $\Gamma(t_1,t_2)$ is twice differentiable if we formally interpret the derivative of a jump function with a delta function. First, note that

$$\begin{aligned}\frac{\partial}{\partial t_2}\Gamma(t_1,t_2) \;=\; 2D\frac{\partial}{\partial t_2}\min(t_1,t_2) \;&=\; 0, \quad t_1 < t_2 \\ &=\; 2D, \quad t_1 > t_2\end{aligned}, \tag{12-68}$$

a step function in t_1. Now, differentiate (12-68) with respect to t_1 and obtain

$$\frac{\partial^2}{\partial t_1 \partial t_2}\Gamma(t_1,t_2) = 2D\,\delta(t_1 - t_2)\,. \tag{12-69}$$

The Wiener process can be thought of as an infinitely dense sequence of infinitesimal jumps (*i.e.*, the limit of the random walk as both the step size and time to take a step approach zero). As mentioned previously, Gaussian white noise is the generalized derivative of the Wiener process. Hence, it seems plausible that wide-band Gaussian noise might be constructed by using a very dense sequence of very narrow pulses, the areas of which are zero mean Gaussian with a very small variance.

Example 12-8 (Construct a Wide-Band Gaussian Process): How can we construct a Gaussian process with a large-as-desired bandwidth? In light of the discussion in this section, we should try to construct a sequence of "delta-like" pulse functions that are assigned Gaussian amplitudes. As our need for bandwidth grows (*i.e.*, as the process bandwidth increases), the "delta-like" pulse functions should (1) become increasingly dense in time and (2) have areas with increasingly smaller variance. These observations result from the fact that the random walk becomes an increasingly dense sequence of increasingly smaller jumps as it approaches the Wiener process, the generalized derivative of which is Gaussian white noise. We start with a discrete sequence of independent Gaussian random variables x_k, $k \ge 0$,

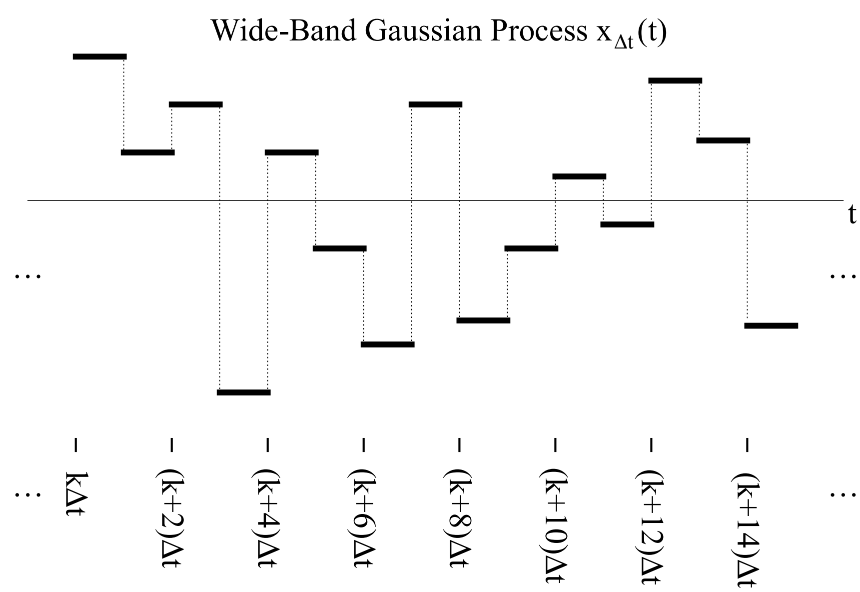

Figure 12-8: Wide-band Gaussian random process composed of "delta-like" pulses with height $x_k / \sqrt{\Delta t}$ and weight (area) $x_k \sqrt{\Delta t}$. The area of each pulse has a variance that is proportional to Δt.

$$
\begin{aligned}
& E[x_k] = 0, \ k \ge 0, \\
& E[x_k x_j] \ = \ 2D, \quad k = j \\
& \qquad\qquad = \ 0, \quad k \ne j
\end{aligned}
\tag{12-70}
$$

For $\Delta t > 0$ and $k \ge 0$, we define the random process

$$
x_{\Delta t}(t) \equiv \frac{x_k}{\sqrt{\Delta t}}\ , \ k\Delta t \le t < (k+1)\Delta t\ , \tag{12-71}
$$

a sample function of which is illustrated by Figure 12-8. As Δt approaches zero, our process becomes an increasingly dense sequence of "delta-like rectangles" (rectangle amplitudes grow like $1/\sqrt{\Delta t}$) that have increasingly smaller weights (rectangle areas diminish like $\sqrt{\Delta t}$). Clearly, from (12-70), we have $E[x_{\Delta t}(t)] = 0$. Also, the autocorrelation of our process approaches a delta function of weight 2D since

$$
R_{x_{\Delta t}}(t_1, t_2) = \begin{cases} E[x_{\Delta t}(t_1)x_{\Delta t}(t_2)] \ = \ \dfrac{2D}{\Delta t}, & k\Delta t \le t_1, t_2 < (k+1)\Delta t \text{ for some integer k} \\ \qquad = \ 0, & \text{otherwise}, \end{cases} \tag{12-72}
$$

and

$$
\lim_{\Delta t \to 0} R_{x_{\Delta t}}(t_1, t_2) = 2D\delta(t_1 - t_2)\,. \tag{12-73}
$$

Hence, by taking Δt sufficiently small, we can, at least in theory, create a Gaussian process of any desired bandwidth.

Mean Square Riemann Integral

Integrals of random processes crop up in many applications. For example, a slowly varying signal may be corrupted by additive high-frequency noise. Sometimes, the rapid fluctuation can be "averaged" or "filtered out" by an operation involving an integration. As a second example, the integration of random processes is important in applications that utilize integral operators such as convolution.

A *partition* of the finite interval [a, b] is a set of subdivision points t_k, k = 0, 1 , ... , n, such that

$$a = t_0 < t_1 < \cdots < t_n = b. \tag{12-74}$$

Also, we define the *time increments*

$$\Delta t_i = t_i - t_{i-1},$$

$1 \le i \le n$. We denote such a partition as $\mathcal{P}_n$, where n+1 is the number of points in the partition. Let $\varDelta_n$ denote the *upper bound on the mesh size*; that is, define

$$\varDelta_n = \max_k \Delta t_k, \tag{12-75}$$

a value that decreases as the partition becomes *finer* and n becomes larger.

For $1 \le k \le n$, let t'_k be an arbitrary point in the interval $[t_{k-1}, t_k)$. For a finite-power random process X(t), we define the *Riemann sum*

$$\sum_{k=1}^{n} X(t'_k)\Delta t_k . \tag{12-76}$$

Now, the *mean square Riemann integral* over the interval [a, b] is defined as

$$\int_a^b X(t)dt \equiv \underset{\Delta_n \to 0}{\text{l.i.m}} \sum_{k=1}^{n} X(t'_k)\Delta t_k \,. \tag{12-77}$$

As $\Delta_n \to 0$ (the upper bound on the mesh size approaches zero) integer n approaches infinity, and the Riemann sum converges (in mean square) to the mean-square Riemann integral, if all goes well. As is the case for m.s. continuity and m.s. differentiability, a necessary and sufficient condition, involving Γ, is available for the existence of the m.s. Riemann integral.

Theorem 12-12: The mean-square Riemann integral (12-77) exists if, and only if, the "ordinary" double integral

$$\int_a^b \int_a^b \Gamma(\alpha,\beta)\, d\alpha d\beta \tag{12-78}$$

exists as a finite quantity.

Proof: Again, the Cauchy Convergence Criteria serves as the basis of this result. Let P_n and P_m denote two distinct partitions of the [a, b] interval. We define these partitions as

$$P_n : \begin{cases} a = t_0 < t_1 < \cdots < t_n = b \\ \Delta t_i = t_i - t_{i-1} \\ \Delta_n \equiv \max\limits_i \Delta t_i \end{cases}$$

$$P_m : \begin{cases} a = \hat{t}_0 < \hat{t}_1 < \cdots < \hat{t}_m = b \\ \Delta \hat{t}_i = \hat{t}_i - \hat{t}_{i-1} \\ \Delta_m \equiv \max\limits_i \Delta \hat{t}_i \end{cases} \tag{12-79}$$

Partion P_n has time increments denoted by $\Delta t_i = t_i - t_{i-1}$, $1 \le i \le n$, and it has an upper bound on mesh size of Δ_n . Likewise, partition P_m has time increments of $\Delta\hat{t}_k = \hat{t}_k - \hat{t}_{k-1}$, $1 \le k \le m$, and it has an upper bound on mesh size of Δ_m. According to the Cauchy Convergence Criteria, the mean-square Riemann integral (12-77) exists if, and only if,

$$\underset{\substack{\Delta_n \to 0 \\ \Delta_m \to 0}}{\text{limit}} \; E\left[\left(\sum_{k=1}^{n} X(t'_k)\Delta t_k - \sum_{j=1}^{m} X(\hat{t}'_j)\Delta\hat{t}_j\right)^2\right] = 0\,, \tag{12-80}$$

where t'_k and $\hat{t}'_j$ are arbitrary points with $t_{k-1} \le t'_k < t_k$ for $1 \le k \le n$, and $\hat{t}_{j-1} \le \hat{t}'_j < \hat{t}_j$ for $1 \le j \le m$. Now, expand out the square, and take expectations to see that

$$\begin{aligned} &E\left[\left(\sum_{k=1}^{n} X(t'_k)\Delta t_k - \sum_{j=1}^{m} X(\hat{t}'_j)\Delta\hat{t}_j\right)^2\right] \\ &\quad = \sum_{k=1}^{n}\sum_{i=1}^{n}\Gamma(t'_k, t'_i)\Delta t_k \Delta t_i - 2\sum_{k=1}^{n}\sum_{j=1}^{m}\Gamma(t'_k, \hat{t}'_j)\Delta t_k \Delta\hat{t}_j + \sum_{j=1}^{m}\sum_{i=1}^{m}\Gamma(\hat{t}'_j, \hat{t}'_i)\Delta\hat{t}_j \Delta\hat{t}_i \end{aligned} \,. \tag{12-81}$$

As Δ_n and Δ_m approach zero (the upper bounds on the mesh sizes approach zero), Equation (12-80) is true if, and only if,

$$\underset{\substack{\Delta_n \to 0 \\ \Delta_m \to 0}}{\text{limit}} \; \sum_{k=1}^{n}\sum_{j=1}^{m}\Gamma(t'_k, \hat{t}'_j)\Delta t_k \Delta\hat{t}_j = \int_a^b\int_a^b \Gamma(\alpha,\beta)\, d\alpha d\beta\,. \tag{12-82}$$

Note that cross-term (12-82) must converge independent of the paths that n and m take as $n \to \infty$ and $m \to \infty$. If this happens, the first and third sums on the right-hand side of (12-81) also converge to the same double integral, and (12-80) converges to zero.

Example 12-7: Let $X(t)$, $t \ge 0$, be the Wiener process, and consider the m.s. integral

$$Y(t) = \int_0^t X(\tau)d\tau\,. \tag{12-83}$$

Recall that $\Gamma(t_1,t_2) = 2D\min(t_1,t_2)$, for some constant D. This can be integrated to produce

$$\begin{aligned}\int_0^t\int_0^t \Gamma(\tau_1,\tau_2)d\tau_1 d\tau_2 &= \int_0^t\int_0^t D\min(\tau_1,\tau_2)d\tau_1 d\tau_2 \\ &= D\int_0^t\Big[\int_0^{\tau_2}\tau_1\,d\tau_1 + \int_{\tau_2}^t \tau_2\,d\tau_1\Big]d\tau_2\,. \qquad (12\text{-}84) \\ &= Dt^3/3\,.\end{aligned}$$

The Wiener process X(t) is m.s. Riemann integrable (*i.e.*, (12-83) exists for all finite t) since the double integral (12-84) exists for all finite t.

Example 12-8: Let Z be a random variable with $E[Z^2] < \infty$. Let c_n, $n \ge 0$, be a sequence of real numbers converging to real number c. Then c_nZ, $n \ge 0$, is a sequence of random variables. Example 11-9 shows that

$$\underset{n\to\infty}{\text{l.i.m}}\; c_n Z = cZ\,.$$

We can use this result to evaluate the m.s. Riemann integral

$$\begin{aligned}\int_0^t 2Z\tau\,d\tau &= \underset{\substack{n\to\infty\\ \Delta_n\to 0}}{\text{l.i.m}}\; 2Z\sum_{k=0}^{n-1}\tau'_{i_k}(\tau_{k-1}-\tau_k) \\ &= 2Z\underset{\substack{n\to\infty\\ \Delta_n\to 0}}{\text{limit}}\;\sum_{k=0}^{n-1}\tau'_{i_k}(\tau_{k-1}-\tau_k) \\ &= 2Z\int_0^t \tau\,d\tau = 2Z\frac{\tau^2}{2}\Big|_0^t \\ &= Zt^2\end{aligned}$$

Properties of the Mean-Square Riemann Integral

As stated in the introduction of this chapter, many concepts from the mean-square calculus have analogs in the "ordinary" calculus, and vice versa. We point out a few of these parallels in this section.

Theorem 12-13: If finite-power random process X(t) is mean-square continuous on [a, b] then it is mean-square Riemann integrable on [a, b].

Proof: Suppose that X(t) is m.s. continuous for all t in [a, b]. From Theorem 12-3, $\Gamma(t_1,t_2)$ is continuous for all $a \le t_1 = t_2 \le b$. From Corollary 12-3B, $\Gamma(t_1,t_2)$ is continuous at all t_1 and t_2, where $a \le t_1,\ t_2 \le b$, (not restricted to $t_1 = t_2$). But this is sufficient for the existence of the integral (12-78), so X(t) is m.s. Riemann integrable on [a, b].

Theorem 12-14: Suppose that X(t) is m.s. continuous on [a, b]. Then the function

$$Y(t) \equiv \int_a^t X(\tau)\,d\tau\,,\ a \le t \le b, \tag{12-85}$$

is m.s. continuous and differentiable on [a, b]. Furthermore, we have

$$\dot{Y}(t) = X(t)\,. \tag{12-86}$$

Mean and Correlation of Mean Square Riemann Integrals

Suppose f(t,u) is a deterministic function, and X(t) is a random process. If the integral

$$Y(u) = \int_a^b f(t,u)X(t)dt \tag{12-87}$$

exists, then

$$E[Y(u)] = E\left[\int_a^b f(t,u)X(t)\,dt\right]$$

$$= E\left[\underset{\Delta_n \to 0}{\text{l.i.m}} \sum_{k=1}^{n} f(t'_k,u)X(t'_k)\Delta_{t_k}\right] \tag{12-88}$$

$$= \underset{\Delta_n \to 0}{\text{limit}}\; E\left[\sum_{k=1}^{n} f(t'_k,u)\, X(t'_k)\Delta_{t_k}\right],$$

where $\Delta t_k \equiv t_k - t_{k-1}$. For every finite n, the expectation and sum can be interchanged so that (12-88) becomes

$$E[Y(u)] = \underset{\Delta_n \to 0}{\text{limit}} \sum_{k=1}^{n} f(t'_k,u)\; E[X(t'_k)]\Delta_{t_k} = \int_a^b f(t,u)E[X(t)]dt\,. \tag{12-89}$$

In a similar manner, the autocorrelation of Y can be computed as

$$\Gamma_Y(u,v) = E\left[\int_a^b f(t,u)X(t)dt \int_a^b f(\hat{t},v)X(\hat{t})d\hat{t}\right]$$

$$= E\left[\underset{\Delta_n \to 0}{\text{l.i.m}} \sum_{k=1}^{n} f(t'_k,u)X(t'_k)\Delta_{t_k} \;\; \underset{\Delta_m \to 0}{\text{l.i.m}} \sum_{j=1}^{m} f(\hat{t}'_j,u)X(\hat{t}'_j)\Delta_{\hat{t}_j}\right] \tag{12-90}$$

$$= \underset{\substack{\Delta_n \to 0 \\ \Delta_m \to 0}}{\text{limit}}\; E\left[\sum_{k=1}^{n} f(t'_k,u)X(t'_k)\Delta_{t_k} \;\; \sum_{j=1}^{m} f(\hat{t}'_j,u)X(\hat{t}'_j)\Delta_{\hat{t}_j}\right]$$

But, for all finite n and m, the expectation and double sum can be interchanged to obtain

$$\Gamma_Y(u,v) = \underset{\substack{\Delta_n \to 0 \\ \Delta_m \to 0}}{\text{limit}} \sum_{k=1}^{n}\sum_{j=1}^{m} f(t'_k, u)\Gamma_x(t'_k, \hat{t}'_j) f(\hat{t}'_j, u)\Delta_{\hat{t}_j}\Delta_{t_k}$$
$$= \int_a^b \int_a^b f(t,u) f(s,v)\Gamma_x(t,s)\, dtds \tag{12-91}$$

where Γ_X is the correlation function for process X.

By now, the reader should have realized what mean square calculus has to offer. Mean square calculus offers, in a word, *simplicity*. To an uncanny extent, the theory of mean square calculus parallels that of "ordinary" calculus, and it is easy to apply. Based on only the correlation function, simple criterion are available for determining if a process is m.s. continuous, m.s. differentiable, and m.s. integrable.