Chapter 5

Brownian Motion and the Wiener Process

In continuous time, stochastic systems described by recursions of the form

\[ x_{k+1} = f(x_k, u_k, w_k), \quad x \in \mathbb{R}^n, u \in \mathbb{R}^m, w \in \mathbb{R}^p, k \in \mathbb{Z}_+, \]

\[ x_0 \sim w, w, k \in \mathbb{Z}_+, \text{ i.i.d. } N(0, \Sigma), \]

are replaced by systems described by (Itô) stochastic differential equations of the form

\[ dx_t = f(x_t, u_t)dt + g(x_t, u_t)dw_t, \quad t \in \mathbb{R}_+, \]

where \( x_0 \) is a random initial condition, \( x_0 \sim w \) and \( dw_t \) denotes the random measure arising from the what may be viewed intuitively as the infinitesimal increments of a Wiener process. The increment \( dw_t \) will be seen to be of order \( \sqrt{dt} \) in the sense that

\[ E(w_{t+dt} - w_t)(w_{t+dt} - w_t)^T = Idt. \]

The Physical Basis of Brownian Motion

The motivation for Wiener’s construction of the Wiener process came from the work of Brown (1926) and Einstein and Shmoluchovski (1906) on what is called the physical Brownian Motion process (see Einstein [1956], Chandrasekhar (in Wax [1959]), and Shuss [1980], who we are partly following in the discussion below). Brown was the first to notice that small particles and grains maintained perpetual irregular motions when suspended in a fluid. This is due to molecular collisions: colloidal particles of radius 50µ (\( \mu = 10^{-4} \text{ cm} \)) will undergo approximately \( 10^{21} \) collisions per second with the molecules of an immersing fluid. (For comparison we note that in ideal gases at normal temperatures and pressures there are \( 6.02 \times 10^{23} \times (22.4 \times 10^3)^{-1} = 2.7 \times 10^{19} \) molecules per cc. By lattice packing the average distance between molecules is 30 Angst. = \( 30 \times 10^{-4} \mu = 30 \times 10^{-8} \text{ cm.} \) and each molecule makes approximately \( 10^9 \) collisions per second.)

The following is a sketch of the mathematical theory of physical Brownian motion which is based upon the references cited above, and whose modelling argument displays an interesting
three time scale (macro, mezzo and micro) structure. Since a particle is much more massive than the molecules hitting it, we may consider the motion to be the result of 2 forces:

(i) Friction: By Stokes’ Law, the drag force per unit mass acting on a spherical particle of radius $a$ is

$$ f_d = -\beta v \in \mathbb{R}^3, \quad \text{where} \quad \beta = \frac{6\pi a \eta}{m} > 0, $$

where $\eta$ is the coefficient of viscosity.

(ii) Collision forces: we assume there is a randomly fluctuating force $f(t) \in \mathbb{R}^3, t \in \mathbb{R}_+$, per unit mass where:

(i) $f(t)$ is statistically independent of $v(t)$;

(ii) $f(t)$ variations are much more frequent than those of $v(t)$;

(iii) the average of $f(t)$ is zero for all $t$.

Consequently Newton’s equations of motion for the particle are

$$ \frac{dv(t)}{dt} = -\beta I v(t) + f(t), \quad (1) $$

where this is a differential equation with random right hand side. Since such equations and their solution have not yet been constructed we proceed formally. First consider

$$ P(v(t) \in A|v(0) = v_0) = \int_A p(v, t|v_0)dv, $$

where $p(v, t, v_0) \rightarrow \delta(v - v_0)$ as $t \rightarrow 0$.

From statistical physics it is know that, as $t \rightarrow \infty$, $p(v(t, v_0)$ must approach the Maxwellian density (parametrized by the temperature $T$) of the surrounding medium (independently of $v_0$); hence

$$ \lim_{t \rightarrow \infty} p(v, t, v_0) = \frac{1}{(2\pi)^{3/2}(\frac{kT}{m})^{3/2}} \exp \left\{ -\frac{1}{2}\begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} \begin{bmatrix} \frac{kT}{m} & 0 & 0 \\ 0 & \frac{kT}{m} & 0 \\ 0 & 0 & \frac{kT}{m} \end{bmatrix}^{-1} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \right\} = \frac{1}{(2\pi)^{3/2}(\frac{kT}{m})^{3/2}} \exp \left\{ -\frac{1}{2\frac{kT}{m}} ||v||^2 \right\} $$

(2)
Since the formal solution to (??) is

\[ v(t) = e^{-\beta(t-t_0)}v_0 + \int_{t_0}^{t} e^{-\beta(t-s)} f(s) \, ds, \]

we obtain

\[ v(t) - \int_{t_0}^{t} e^{-\beta(t-s)} f(s) \, ds \to 0, \quad \text{as} \quad t_0 \to -\infty \]

and in the limit

\[ v(t) = \int_{-\infty}^{t} e^{-\beta(t-s)} f(s) \, ds \approx \int_{0}^{t} e^{-\beta(t-s)} f(s) \, ds, \quad t \gg 0. \]

For \( t \gg 0 \) consider approximating Riemann sums for the integral. Then

\[ \int_{0}^{t} e^{-\beta(t-s)} f(s) \, ds \approx e^{-\beta t} \sum_{n=1}^{\lfloor \frac{t}{\Delta t} \rfloor} e^{\beta n \Delta t} f(n \Delta t) \Delta t \]

\[ = \sum_{n=1}^{N} e^{\beta(n \Delta t - t)} \Delta b_n \]

where \( \Delta b_n \) denotes the random acceleration of the particle over \([n \Delta t, (n + 1) \Delta t]\), where these events are assumed to be independent on distinct intervals.

Next assume \( \Delta t \) is large with respect to the average period of fluctuation of \( f \) due to successive molecular collisions. (For instance, if \( \Delta t = 10^{-6} \) seconds, \( \Delta b_n \) will be the sum of \(0(10^{21}) \times 10^{-6} = 0(10^{15}) \) collisions.) Then we may write

\[ \Delta b_n = \sqrt{M} \left( \frac{1}{\sqrt{M}} \sum_{k=1}^{M} (s_k^n) \right), \]

where \( M \gg 1 \) and \( s_k^n \) denotes the change in force (per unit mass) on the particle over the associated sub-interval \( M^{-1} \Delta t \) due to molecular collisions.

From the theory of statistical thermodynamics, it is known that the distribution of these changes in each dimension is such that \( \{s_k^n; 1 \leq k \leq M\} \) are independent for distinct \( k \), have zero mean and have variance \( E\|s_k^n\|^2 = 2qM^{-1} \Delta t \) for some \( q > 0 \). The fact that the this variance is linear in the time interval \( M^{-1} \Delta t \) is the crucial step in the construction of Brownian motion (and in the derivation of the properties of the Wiener process and consequently the Što calculus).
Consequently, from the Central Limit Theorem, we have
\[ \Delta b_n \sim N(0, \text{diag}^3(2q \Delta t)), \]
where \( \text{diag}^3(\alpha) \) denotes a 3 \( \times \) 3 diagonal matrix with diagonal entry \( \alpha \).

Finally we obtain
\[
E\|v(t)\|^2 = \sum_{n=1}^{N} (2q(\Delta t))e^{2\beta n (\Delta t-t)} \\
- 2q \int_0^t e^{2\beta (s-t)} ds = 2q \frac{1}{2\beta} [1 - e^{2\beta (t-t)}],
\]
as \( N \to \infty \). But from the Maxwellian density we have \( E\|v(t)\|^2 \to \frac{kT}{m} \) in each dimension as \( t \to \infty \), and hence \( q = \frac{\beta kT}{m} \).

**Position Distribution Calculation via State Space Solution in \( \mathbb{R}^6 \)**

Solving the six dimensional state space system
\[
d\begin{bmatrix} x \\ v \end{bmatrix} = \begin{bmatrix} 0 & I \\ 0 & -\beta I \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix} dt + \begin{bmatrix} 0 \\ I \end{bmatrix} f(t) dt,	 t \in \mathbb{R}_+,
\]
with initial condition \( \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} \) for \( x \) gives
\[
x_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left\{ e^{0 - \beta t} \right\} t \begin{bmatrix} x_0 \\ v_0 \end{bmatrix} + \int_0^t e^{0 - \beta (t-s)} f(s) ds \right\} \\
= x_0 + \left( 1 - e^{-\beta t} \right) v_0 + \int_0^t \left( 1 - e^{-\beta (t-s)} \right) f(s) ds,
\]
since \( e^{0 - \beta t} = \begin{bmatrix} 1 \frac{1-e^{-\beta t}}{\beta} \frac{1-e^{-\beta t}}{\beta} \frac{1-e^{-\beta t}}{\beta} \end{bmatrix} \).

Hence \( x_t - (x_0 + (1-e^{-\beta t}) v_0) \in \mathbb{R}^3 \) is a zero mean Gaussian random variable with variance in each dimension given by
\[
\sigma^2 = 2q \int_0^t \frac{(1-e^{-\beta (t-s)})^2}{\beta^2} ds = \frac{q}{\beta^3} (2\beta t - 3 + 4e^{-\beta t} - e^{-2\beta t}) = \frac{q}{\beta^3} \gamma(t) = \frac{\beta kT}{\beta^3 m} \gamma(t).
\]

Hence \( \text{Cov}_t = [\text{diag}(\frac{kT \gamma(t)}{m \beta^2})] \in \mathbb{R}^{3 \times 3} \) and
\[
p(x, t|x_0, v_0) = \left\{ \frac{m \beta^2}{(2\pi kT \gamma(t))^{3/2}} \right\}^{3/2} \exp \left\{ -\frac{m \beta^2 \|x - (x_0 + (1-e^{-\beta t}) v_0)\|^2}{2 kT \gamma(t)} \right\}.
\]
Take $\gamma(t) \Delta 2\beta t$ for $t \gg 0$; then averaging over a zero mean symmetric distribution $p(v_0)$ for $v_0$ yields

$$p(x, t| x_0) = \int_{\mathbb{R}^3} p(x, t| x_0, v_0)p(v_0)dv_0 = \frac{1}{(2\pi 2Dt)^{3/2}} \exp \left\{ - \frac{1}{2} \frac{\|x - x_0\|^2}{2Dt} \right\},$$

where

$$4Dt \Delta 2 \frac{kT\gamma}{m\beta^2} = \frac{4kTt}{m\beta}$$

and so

$$D = \frac{kT}{m\beta} = \frac{kT}{6\pi a\eta}.$$  

We conclude that

$$\frac{\partial p}{\partial t} = D \Delta p \equiv D \left( \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \right) p,$$

with $p(x, t| x_0) \to \delta(x - x_0)$ as $t \to 0$.

Hence we may say that an idealization of physical Brownian Motion is a stochastic process in $\mathbb{R}^3$ such that

$$P(x(t) \in A| x_0) = (4\pi Dt)^{-3/2} \int_A e^{-\frac{\|x-x_0\|^2}{4Dt}} dx, \quad A \in B(\mathbb{R}^3),$$

and such that

(i) the increments $x(t + v) - x(t), x(t) - x(t - u)$ are independent for all $t \geq 0, v \geq 0, u \geq 0, t - u \geq 0$, and, further, are functionally independent of $t \geq 0$;

(ii) the paths $\{x(t); t \geq 0\}$ are continuous;

(iii) the joint distribution of $x(t_1), \cdots, x(t_n)$ is Gaussian for all $n$ and all $t_1 < \cdots < t_n$.

(Notice that the $x$ is a Markov process; but it has been constructed as an approximation to a part of a state $(x, v)$ of the Markov process in (??).) This physical analysis leads to the following definition of a Wiener process as a standardized version of Brownian Motion.
The Mathematical Theory of the Wiener Process

Rigorous mathematical analyses which establish the Wiener process nature of the limiting behaviour of the particle models (as the number of particles and the collision rates go to infinity) are to be found in [F. Spitzer, *Journal of Math. and Mech.* Vol 18, No. 19, 1969, pp. 973-989] and [R. Holley, *Z. Wahrscheinlichkeitsth. verw. Geb*, 17, pp 181-219, 1971]; in this context the classical text by Khinchin [*Mathematical Principles of Statistical Mechanics*, Dover, 1969] should also be cited.

**Definition 5.1** A Wiener process \( w \) is an \( \mathbb{R}^n \) valued stochastic process such that

Axiom W1. \( w_0 = 0 \), w.p.1,

Axiom W2. \( p(w_t|w_s) = \frac{1}{(2\pi)^{n/2}} \frac{1}{(t-s)^{n/2}} \exp - \frac{1}{2(t-s)} \|w_t - w_s\|^2 \), \( 0 \leq s < t \),

Axiom W3. The process increments satisfy the independent increment property, \( \prod \{w_{p_i} - w_{q_i}\}_1^N \), for all \( N \in \mathbb{Z}_1 \) and all \( q_1^N, p_1^N \in \mathbb{R} \), such that \( 0 \leq q_1 < p_1 \leq q_2 < \cdots \leq q_N < p_N \).

We observe that Axiom W2 is equivalent to

\[
p_y(\alpha|w_s) = \frac{1}{(2\pi)^{n/2}} \frac{1}{|t-s|^{n/2}} \exp - \frac{1}{2(t-s)} \|\alpha\|^2,
\]

when \( \alpha \Delta w_t - w_s \); and that Axiom W3 may be written

\[
p(w_{p_N} - w_{q_N}, \ldots, w_{p_1} - w_{q_1}) = \prod_{i=1}^N p(w_{p_i} - w_{q_i}).
\]

The transition distribution of a stochastic process \( x \) is defined as \( P(x_t \in A|x_s) \) for all \( A \in B(\mathbb{R}^n) \) and all \( s < t, s, t \in \mathbb{R} \), and, by the definition of a density, we have the representation

\[
P(x_t \in A|x_s) = \int_A p_{t,s}(x|x_s)dx, \quad s < t, A \in B(\mathbb{R}^n),
\]

when a transition density exists.
For a Wiener process, Axiom W2 gives the representation

\[
P(w_t \in A|w_s) = \int_A \frac{1}{(2\pi)^{n/2}} \frac{1}{t - s} \exp - \frac{1}{2} \left\{ \frac{||w - w_s||^2}{t - s} \right\} dw_t, \quad A \in B(\mathbb{R}^n),
\]
and, in case \( A = \prod_{i=1}^n [a_i, b_i] \),

\[
P(w_t \in \prod_{i=1}^n [a_i, b_i]|w_s) = \int_{[a_1, b_1]} \cdots \int_{[a_n, b_n]} \frac{1}{(2\pi)^{n/2}} \frac{1}{t - s} \exp - \frac{1}{2} \left\{ \frac{||w - w_s||^2}{t - s} \right\} dw.
\]

The finite dimensional distributions of a Wiener process are Gaussian of the form given in (??) below. This is the case since, in general, assuming the indicated conditional densities exist, we have that for any \( m \in \mathbb{Z}_+ \), any collection of sets \( A_1^m \in B(\mathbb{R}^n) \), and any collection of instants \( t_i^m \) with \( 0 < t_i - t_{i-1}, i = 1, \cdots, m \),

\[
P(w_{t_1} \in A_1, \ldots, w_{t_m} \in A_m)
= \int_{A_1} \cdots \int_{A_m} p_{m, m-1, \ldots, 1}(w_m, w_{m-1}, \cdots, w_1) dw_1 \cdots dw_m
= \int_{A_1} \cdots \int_{A_m} p(w_m|w_1^{m-1})p(w_{m-1}|w_1^{m-2}) \cdots p(w_1) dw_m \cdots dw_1
= \int_{A_1} \cdots \int_{A_{m-w_{m-1}}} p((w_m - w_{m-1})(w_i - w_{i-1})_1^{m-1}) \cdots p(w_1 - w_0) d\bar{w}_m \cdots d\bar{w}_1,
\]

where for simplicity we have written \( w_j \) for \( w_{t_j} \), \( 1 \leq k \leq m \), \( \bar{w}_m \Delta w_m - w_{m-1}, \cdots, \bar{w}_1 \Delta w_1 - w_0 \), and where we have invoked Axiom W1 and the Change of Variables of Formula (henceforth denoted CVF). However, in the case under consideration, Axioms W2 and W3 together imply that the conditional densities above exist and take the Gaussian form in the third expression below:

\[
P(w_{t_1} \in A_1, \ldots, w_{t_m} \in A_m)
= \int_{A_1} \cdots \int_{A_{m-w_{m-1}}} p(w_m - w_{m-1}) \cdots p(w_1 - w_0) d\bar{w}_m \cdots d\bar{w}_1 \quad \text{(by Axiom W3)}
= \int_{A_1} \cdots \int_{A_{m-w_{m-1}}} \frac{1}{(2\pi)^{n/2}} \frac{1}{|t_m - t_{m-1}|^{n/2}} \exp - \frac{1}{2} \left\{ \frac{1}{(t_m - t_{m-1})} \right\} ||w_m - w_{m-1}||^2 d\bar{w}_m
\times \frac{1}{(2\pi)^{n/2}} \frac{1}{|t_{m-1} - t_{m-2}|^{n/2}} \exp - \frac{1}{2} \left\{ \frac{1}{(t_{m-1} - t_{m-2})} \right\} ||w_{m-1} - w_{m-2}||^2 d\bar{w}_{m-1}
\cdots \times \frac{1}{(2\pi)^{n/2}} \frac{1}{|t_1|^{n/2}} \exp - \frac{1}{2} \left\{ \frac{1}{t_1} \right\} ||w_1||^2 d\bar{w}_1. \quad \text{(by Axiom W2)}
\]

\(7\)
From the axioms W1 - W3, or from the expression above, it follows that the finite dimensional densities of $x$ are Gaussian and are given explicitly by:

$$p(w_t, \cdots, w_{t_m}) = \frac{1}{(2\pi)^{mn/2}} \prod_{i=1}^{m} (t_i - t_{i-1})^{n/2} \times \exp\left(-\frac{1}{2}[w_{t_m}^T - w_{t_m-1}^T, w_{t_m-1}^T - w_{t_m-2}^T, \cdots, w_{t_1}^T]^T \begin{bmatrix} (t_m - t_{m-1})I & 0 \\ 0 & (t_{m-1} - t_{m-2})I \\ \vdots & \vdots \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} w_{t_m} - w_{t_m-1} \\ w_{t_m-1} - w_{t_m-2} \\ \vdots \\ w_{t_1} \end{bmatrix}\right).$$

But when

$$L_m \triangleq \begin{bmatrix} I & -I & \cdots & 0 \\ 0 & I & \cdots & 0 \\ \vdots & \vdots & \ddots & I - I \\ 0 & 0 & \cdots & I \end{bmatrix} \quad (mn \times mn)$$

and

$$V_m \triangleq \begin{bmatrix} t_m & t_{m-1} & \cdots & t_1 \\ t_{m-1} & t_{m-1} & \cdots & t_1 \\ \vdots & \vdots & \ddots & \vdots \\ t_1 & t_1 & \cdots & t_1 \end{bmatrix} \otimes I_n, \quad (mn \times mn).$$
we have the equality

\[
[w_{t_m}^T, \ldots, w_{t_1}^T][V_m]^{-1} \begin{bmatrix}
  w_{t_m} \\
  : \\
  w_{t_1}
\end{bmatrix}
\]

\[
= [w_{t_m}^T, \ldots, w_{t_1}^T]L_m^T[V_m L_m^{-1}L_m]^{-1}L_m \begin{bmatrix}
  w_{t_m} \\
  : \\
  w_{t_1}
\end{bmatrix}
\]

\[
= [w_{t_m}^T - w_{t_{m-1}}^T, w_{t_{m-1}}^T - w_{t_{m-2}}^T, \ldots, w_{t_1}^T] \begin{bmatrix}
  (t_m - t_{m-1})I & 0 \\
  0 & (t_{m-1} - t_{m-2})I \\
  \vdots & \ddots & t_1I
\end{bmatrix}^{-1} \begin{bmatrix}
  w_{t_m} - w_{t_{m-1}} \\
  w_{t_{m-1}} - w_{t_{m-2}} \\
  \vdots
\end{bmatrix}
\]

Hence

\[
p(w_{t_1}, \ldots, w_{t_m}) = \frac{1}{(2\pi)^{mn/2}} |V_m|^{1/2} \exp\left(-\frac{1}{2} \|w_{t_m}, \ldots, w_{t_1}\|_{V_m^{-1}}^2\right).
\]  \hspace{1cm} (8)

We observe that

\[
V_m \equiv V_{t_i}^m = \begin{bmatrix}
  \vdots \\
  \cdots [Ew_{t_i}w_{t_j}^T] \cdots \\
  \vdots
\end{bmatrix},
\]

since, taking \( t_i < t_j \) for definiteness,

\[
Ew_{t_i}w_{t_j}^T = Ew_{t_i}[(w_{t_j} - w_{t_i}) + w_{t_i}]^T
\]

\[
= E(w_{t_i} - w_0)(w_{t_j} - w_{t_i})^T + E(w_{t_i} - w_0)(w_{t_i} - w_0)^T
\]

\[
= E(w_{t_i} - 0)E(w_{t_j} - w_{t_i}) + E(w_{t_i} - w_0)(w_{t_i} - w_0)^T \quad \text{(by Axiom W1)}
\]

\[
= 0 + I(t_i - 0) \quad \text{(by Axioms W2, W3)}
\]

\[
= I \min(t_i, t_j).
\]

We note that the family of finite dimensional distributions constructed above necessarily satisfies the Compatibility Condition and hence, if it is taken as the starting point, the
Daniell-Kolmogorov Theorem implies the existence of the Wiener process with the properties stated in Definition 4.1.

**Lemma 5.1**

The following three sets of axioms for a Wiener process \( w \) are equivalent:

(i) W1. \( w_0 = 0 \), w.p.1,

W2. \( p(w_t|w_s) = \frac{1}{(2\pi)^{n/2}|t-s|^{n/2}} \exp -\frac{1}{2} \left\{ \frac{1}{|t-s|} \|w_t - w_s\|^2 \right\} \), 0 < s < t,

W3. \( w_t - w_s \prod \{w_{p_i} - w_{q_i}\}_{1}^{N} \), for all \( N \in \mathbb{Z}_1 \) and all \( s,t, q_1^{N}, p_1^{N} \in \mathbb{R} \), such that 0 ≤ q_1 < p_1 ≤ q_2 < ⋯ ≤ q_N < p_N ≤ s < t.

(ii) B1. \( w_t, t \geq 0 \), is a Gaussian process,

B2. \( Ew_t = 0 \), for all \( t \in \mathbb{R}^{+} \),

B3. \( Ew_tw_s^{T} = \min(t,s)I, \quad t, x \in \mathbb{R}^{+} \).

(iii) C1. \( w_0 = 0 \), w.p.1,

C2. W2 holds,

C3. \( w \) is a Markov process.

**Proof** We show \((i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)\)

(1) Assume (W1) (W2) (W3) holds for a stochastic process \( w \). Then it was proved earlier that \( w \) is a Gaussian process, hence B1 holds. Since

\[
Ew_t = EE|_{w_0}(w_t - 0) = E|_{w_0}(w_t - w_0) \\
= \int_{\mathbb{R}^{n}} w_t p(w_t|0)dw_t = 0, \quad \text{(by A2)}
\]

B(2) holds. And next,

\[
Ew_tw_s^{T} = E((w_t - w_s) + w_s)(w_s)^{T} \quad s < t \\
= E(w_t - w_s)E(w_s - 0)^{T} + E(w_s - 0)(w_s - 0)^{T} \\
= \min(t,s)I, \quad \text{(by W3, W2)}
\]

gives B3.
Assume B1 - B3 hold; then B3 gives $Ew_0^2 = 0$ and hence $w_0 = 0$ w.p.1, which is C1.

Next,

$$p(w_t | w_s) = p(w_t, w_s)(p(w_s))^{-1} \begin{cases} 	ext{with } Ew_t = Ew_s = 0 & (B2) \\ \text{and } Ew_p w_q^T = \min(p, q) I & (B3) \end{cases}$$

$$= \frac{1}{(2\pi)^{n/2}} s^{n/2} \exp -\frac{1}{2} \frac{s(t-s)}{s(t-s)} \left\{ sw_t^2 - 2sw_s w_t + tw_s^2 - (t-s)w_s^2 \right\},$$

i.e. $w_t - w_s \sim N(0, (t-s))$,

which is C2, as required.

Finally $w$ is a Markov process since by B1 and B3, $w$ is a Gaussian process with orthogonal, and hence independent, increments. Hence

$$p(w_{t_n} | w_{t_{n-1}}) = p(w_{t_n} - w_{t_{n-1}} | (w_{t_i} - w_{t_{i-1}})_{i=1}^{n-1})$$

But $p(w_{t_n} - w_{t_{n-1}}) = p(w_{t_n} - w_{t_{n-1}}) = p(w_{t_n} | w_{t_{n-1}})$ by C(2), and so C3 holds.

C1-C3 ⇒ W1 - W2 is obvious. The equalities

$$p(w_t - w_s | \{w_{p_i} - w_{q_i}\}) = p(w_t | w_s, \{w_{p_i} - w_{q_i}\}) \quad \text{(by CVF)}$$

$$= p(w_t | w_s) \quad \text{(by C3)}$$

$$= p(w_t - w_s), \quad \text{(by C2)}$$

show that C2-C3 give W3.

The Wiener process has almost surely continuous sample paths (see Doob (1953, p 393)) and is almost surely of unbounded variation on any finite interval (see Doob (1953, p 395)).

**Lemma 5.2**

For the Weiner process $w$ define

(i) $w_1^s(t) \triangleq w_{t+s} - w_s, t, s \in \mathbb{R}_+$,

(ii) $w_2(t) \triangleq tw(\frac{1}{t}), \quad t > 0, \quad w_2(0) \triangleq 0 \ a.s.$

Then $w_1^s$ is a Wiener process with respect to the standard Wiener measure conditioned on the value of $w_s$, and $w_2$ is a Wiener process.
Proof.

(i) $w_s^x(t)$ is a Wiener process, with respect to the probability distribution conditional on $w_s$ since the axioms for a WP are satisfied as follows:

Axiom W1 is satisfied since $w_s^x(0) = (w_{t+s} - w_s)|_{t=0} = 0$. Axiom W2 holds because, for $p > q$,

$$p(w_s^x(p)|w_s^x(q)) = p(w_{p+s} - w_s|w_{q+s} - w_s, w_s)$$

$$= p((w_{p+s} - w_{q+s}) + (w_{q+s} - w_s)|w_{q+s}, w_s)$$

$$= p(w_{p+s} - w_{q+s}|w_{q+s}, w_s)\cdot 1 \quad \text{(by CVF)}$$

$$= p(w_{p+s} - w_{q+s}|w_{q+s} - w_s, w_s - w_0) \quad \text{(by Axiom W1)}$$

$$= \frac{p(w_{p+s} - w_{q+s}|w_{q+s})}{p(w_{q+s} - w_0)} \quad \text{(by definition cond. density and CVF)}$$

$$= \frac{1}{(2\pi)^{n/2}} \frac{1}{|p - q|^{n/2}} \exp\left(-\frac{1}{2(p-q)}\|w_s^x(p) - w_s^x(q)\|^2\right) \quad \text{(by Axiom W2)}$$

where we have used the abbreviated notation that $p(x|y)$ denotes $p_{x|y}(\alpha|\beta)|_{x,y}$.

Finally Axiom W3 (for $w$) holds for $w_1$ since

$$w_1^x(t_2) - w_1^x(t_1) = w_{t_2+s} - w_s - w_{t_1+s} + w_s = w_{t_2+s} - w_{t_1+s},$$

and this dif, by Axiom W3 (for $w$), is independent of $w_{t+s} - w_{t+s} = w_1^x(t_i) - w_1^x(t_j)$, whenever the intervals have disjoint interiors, as required.

(ii) Axiom W1 holds for the process $w_2$ since (treating the scalar case for convenience and without loss of generality)

$$\lim_{t \to 0} Et^2w^2(\frac{1}{t}) = \lim_{t \to 0} t^2\frac{1}{t} = 0$$

implies that

$$\lim_{t \to 0} w_2(t) = \lim_{t \to 0} tw(\frac{1}{t}) = 0 \text{ a.s.}$$
Axiom W3 clearly holds (with \( t_1 > t_2 \geq t_3 > t_4 \geq \cdots \) implying
\[
\frac{1}{t_1} < \frac{1}{t_2} \leq \frac{1}{t_3} < \frac{1}{t_4} \leq \cdots
\].
Axiom W2 is the case since \( tw(\frac{1}{t}) \) and \( sw(\frac{1}{s}) \) are jointly Gaussian random variables, with mean
\[
Etw(\frac{1}{t}) = 0 = Esw(\frac{1}{s}),
\]
and with covariance
\[
E \left( \begin{array}{c}
tw(\frac{1}{t}) \\
sw(\frac{1}{s})
\end{array} \right) \left( \begin{array}{c}
tw(\frac{1}{t}) \\
sw(\frac{1}{s})
\end{array} \right) = \begin{bmatrix}
\frac{t^2}{t} & ts \min(\frac{1}{t}, \frac{1}{s}) \\
ts \min(\frac{1}{t}, \frac{1}{s}) & s^2
\end{bmatrix} = \begin{bmatrix}
t & s \\
s & s
\end{bmatrix}.
\]
Hence
\[
p(tw(\frac{1}{t})|sw(\frac{1}{s})) = \frac{p(tw(\frac{1}{t}), sw(\frac{1}{s}))}{p(sw(\frac{1}{s}))}
= \frac{1}{(2\pi)^{1/2} |t-s|^{1/2}} \exp \left\{ -\frac{1}{2} \left( w_2(t), w_2(s) \right) \left[ \begin{array}{c}
t \\
s
\end{array} \right]^{-1} \left( \begin{array}{c}
w_2(t) \\
w_2(s)
\end{array} \right) + \frac{1}{2s} w_2^2(s) \right\}
= \frac{1}{(2\pi)^{1/2} |t-s|^{1/2}} \exp \left\{ -\frac{1}{2} \left( sw_2^2(t) - 2sw_2(t)w_2(s) + sw_2^2(s) \right) s(t-s) \right\}
= \frac{1}{(2\pi)^{1/2} |t-s|^{1/2}} \exp \left\{ -\frac{1}{2} \|w_2(t) - w_2(s)\|^2 \right\},
\]
as required.

An application of the Law of the Iterated Logarithm (see e.g. Caines(1988)) shows that the process \( w_2 \) is continuous at 0 almost surely.
Self similarity of Wiener processes

Consider the magnified or miniaturized, i.e. scaled, version of the Wiener process given by

\[ w_c(t) \triangleq c w \left( \frac{t}{c^2} \right), \quad c > 0, \quad t \in \mathbb{R}_+. \]

Then

(i) \( w_c(t)|_{t=0} = cw \left( \frac{t}{c^2} \right) = 0 \) w.p.1,

(ii) \( p_{w_c}(w_c(t)|w_c(s)) = p_{w_c}(c.w \left( \frac{t}{c^2} \right)|c.w \left( \frac{s}{c^2} \right)) \)

\[ = p_w \left( w \left( \frac{t}{c^2} \right) \left| \frac{w(s)}{c^2} \right\} \frac{\partial(w)}{\partial(cw)} \right| \]

\[ = \frac{1}{(2\pi)^{n/2}} \frac{1}{\left| \frac{t}{c^2} - \frac{s}{c^2} \right|^{n/2}} \exp \left\{ -\frac{1}{2} \left\{ \frac{1}{\left| \frac{t}{c^2} - \frac{s}{c^2} \right| c^2} \right\} \right\} \left| \frac{\partial(w)}{\partial(cw)} \right| \]

\[ = \frac{1}{(2\pi)^{n/2}} \frac{1}{\left| \frac{t}{c^2} - \frac{s}{c^2} \right|^{n/2}} |Ic|^{-1} \exp \left\{ -\frac{1}{2} \left\{ \frac{1}{|t-s|} \right\} \right\} \left| w_c(t) - w_c(s) \right|^2, \]

\[ = p_w(w_c(t)|w_c(s)). \]

So the measure governing \( w_c(t) \) given \( w_c(s) \) is identical to that governing \( w(t) \) given \( w(s) \).

(iii) \( w \) a Wiener process implies

\[ \left( w \left( \frac{t}{c^2} \right) - w \left( \frac{s}{c^2} \right) \right) \prod_{i=1}^{N} \left\{ w \left( \frac{p_i}{c^2} \right) - w \left( \frac{q_i}{c^2} \right) \right\} \]

when \( 0 \leq \frac{q_i}{c^2} \leq \frac{p_i}{c^2} \leq \cdots \leq \frac{q_N}{c^2} \leq \frac{p_N}{c^2} \leq s < t \), and so multiplying by \( c \) it is clear that \( w_t - w_s \prod \{w_{p_i} - w_{q_i}\}_{i=1}^{N} \). Hence \( w_c \) is a Brownian motion for any \( c > 0 \).
Stochastic Integration and Linear Stochastic Differential Equations

Stochastic integrals with respect to Wiener process increments are defined via the mean square limits of approximating finite sums. Here we present an introduction to the basic ideas.

By definition, for an \( \mathbb{R}^n \) valued Wiener process,
\[
\int_{t_0}^{t_1} dw_t = w_{t_1} - w_{t_0}, \quad t_0 < t_1, \quad t_0, t_1 \in \mathbb{R}_+,
\]
(9)
which is the increment of the Wiener process on \([t_0, t_1]\).

Next consider a sequence of \((m \times n)\) matrices \(\{A_i; 0 \leq i \leq N\}\), instants \(t_0 < t_1 < \cdots < t_{N+1}\), \(t_0 = T_0, t_{N+1} = T_1\) and the associated matricial step function \(A^N(t) \triangleq \sum_{i=0}^{N} A_i (s_i(t) - s_{i+1}(t)), t \in \mathbb{R}_+\), where \(s_\tau(t) = \{0\text{ if } t < \tau, 1\text{ if } \tau \leq t; t \in \mathbb{R}\}\). We give the definition of the integral of a matricial step function against the increments of the standard vector Weiner process \(w\). It is seen to specialize to (9) when \(A\) is the constant identity function,

We set
\[
\int_{t_0}^{t_{N+1}} A^N(t)dw_t = \int_{t_0}^{t_{N+1}} \sum_{i=0}^{N} A_i (s_i(t) - s_{i+1}(t))dw_t
\]
\[
\Delta \sum_{i=0}^{N} A_i \int_{t_i}^{t_{i+1}} dw_t \quad \text{(where } (s_i(t) - s_{i+1}(t) = 0, t \notin [t_i, t_{i+1}])\text{)}
\]
\[
= \sum_{i=0}^{N} A_i(w_{t_{i+1}} - w_{t_i})
\]
We see that the definition automatically gives the linearity of integration with respect to the integrand. Furthermore, since \(E(w_{t_{i+1}} - w_{t_i})(w_{t_{j+1}} - w_{t_j})^T = (t_{i+1} - t_i)I\), if \(i = j\), and is 0 otherwise,
\[
E(\int_{t_0}^{t_{N+1}} A^N(t)dw_t)(\int_{t_0}^{t_{N+1}} A^N(t)dw_t)^T = \sum_{i=0}^{N} (t_{i+1} - t_i)A_iA_i^T
\]
(10)

Now let us assume that the sequence of deterministic matrix step functions of time \(\{A^N(\cdot); N \in \mathbb{Z}_1\}\) converges to a (measurable in \(t \in \mathbb{R}\)) matrix function \(A(\cdot) \in L^2\) on \([t_0 = T_0, t_{(N+1)} = T_1]\) in such a way that
\[
\int_{T_0}^{T_1} (A^N(t) - A(t))(A^N(t) - A(t))^T dt \to 0,
\]
15
as \( N \to \infty \); in other words, let \( \{ A^N(\cdot); N \in \mathbb{Z}_1 \} \) converge in \( L^2 \) to the \( L^2 \) matrix function \( A(\cdot) \). But then the sequence \( \{ A^N(\cdot); N \in \mathbb{Z}_1 \} \) is a Cauchy sequence in \( L^2 \) and hence (reader check) \( \{ \int_{T_0}^{T_1} A^N(t)dw_t; N \in \mathbb{Z}_1 \} \) constitutes a Cauchy sequence in the Hilbert space \( H^w \) spanned by the process \( w \). So by the completeness of the Hilbert space \( H^w \) we conclude that \( \int_{T_0}^{T_1} A^N(t)dw_t \) converges in mean square to a matrix random variable. We denote the limit by \( \int_{T_0}^{T_1} A(t)dw_t \), i.e.

\[
E\| \int_{T_0}^{T_1} A(t)dw_t - \int_{T_0}^{T_1} A(t)dw_t \|^2 \to 0,
\]

as \( N \to \infty \) which is readily verified to be unique for all sequences of functions \( \{ A^N(\cdot); N \in \mathbb{Z}_1 \} \) converging to \( A(\cdot) \) in \( L^2 \). It is consistent and convenient at this point to make the definition

\[
d(\int_0^t A(s)dw_s) \triangleq A(t)dw_t.
\]

We note that any \( L^2 \) matrix function of time on a finite interval is the \( L^2 \) limit of a sequence of \( L^2 \) matrix step functions.

**Lemma 5.3**

Let the sequences of vector random variables \( \{ I_N; N \in \mathbb{Z}_+ \} \) and \( \{ J_N; N \in \mathbb{Z}_+ \} \) form a Cauchy sequences in the Hilbert space \( H^w \) generated by the process \( w \). Then their mean square limits \( I \) and \( J \) exist and satisfy

(i) \( \lim_{N \to \infty} EI_N = EI < \infty \),

(ii) \( \lim_{N,M \to \infty} EI_N J_M^T = EIJ^T < \infty \).

**Proof**

(i) Since \( \{ I_N; N \in \mathbb{Z}_+ \} \) forms a Cauchy sequence in the necessarily complete Hilbert space \( H^w \), there exists a limiting vector random variable \( I \) such that

\( \lim_{N \to \infty} (E\| I_N - I \|^2)^{1/2} = 0 \) and similarly for the sequence \( \{ J_N; N \in \mathbb{Z}_+ \} \). Next we observe that \( \{ EI_N; N \in \mathbb{Z}_+ \} \) forms a Cauchy sequence in \( \mathbb{R}^n \); this is the case since the inequalities

\[
\| EI_N - EI_M \| \leq E\| I_N - I_M \| \leq (E\| I_N - I_M \|^2)^{1/2},
\]

16
show the leftmost term goes to zero as \( N, M \to \infty \) because \( \{I_N; N \in \mathbb{Z}_+\} \) is a Cauchy sequence in \( \mathbf{H}^w \). Hence a unique (non-random) limit \( \lim_{N \to \infty} EI_N \) of the sequence \( \{EI_N; N \in \mathbb{Z}_+\} \) exists in \( \mathbf{R}^n \). Finally, the inequalities

\[
\|EI - \lim_{N \to \infty} EI_N\| \leq \|EI - EI_M\| + \|EI_M - \lim_{N \to \infty} EI_N\|
\]

\[
\leq (E\|I - I_M\|^2)^{1/2} + \|EI_M - \lim_{N \to \infty} EI_N\|
\]

for all \( M \), show that \( EI = \lim_{N \to \infty} EI_N < \infty \).

(ii) For \( I_N, I_M \) and \( I \) as above, the inequality

\[
\|EIJ^T - EI_M J_N^T\| = \|E(I - I_M)(J - J_N)^T + EI_M (J - J_N)^T + EI_N (J - J_M)^T\|
\]

\[
\leq (E\|I - I_M\|^2)^{1/2}(E\|J - J_N\|^2)^{1/2} + (E\|I_M\|^2)^{1/2}(E\|J - J_N\|^2)^{1/2}
\]

\[
+ (E\|I_N\|^2)^{1/2}(E\|J - J_N\|^2)^{1/2}
\]

shows that \( \lim_{N,M \to \infty} EI_N I_M^T = EII^T < \infty \), as required.

Since \( E \int_{T_0}^{T_1} A^N(t)dw_t = 0 \) for all \( N \), Lemma 4.3 gives

\[
E \int_{T_0}^{T_1} A(t)dw_t = E \lim_{N \to \infty} \int_{T_0}^{T_1} A^N(t)dw_t = \lim_{N \to \infty} E \int_{T_0}^{T_1} A^N(t)dw_t = 0.
\]

When the sequence of step functions \( \{A^N(\cdot); N \in \mathbb{Z}_+\} \) converges to \( A(\cdot) \in L^2 \), and \( T_1 \leq T_2 \) (for definiteness), that

\[
E(\int_{T_0}^{T_1} A(t)dw_t)(\int_{T_0}^{T_2} A(t)dw_t) = E \lim_{N \to \infty} \left[ (\int_{T_0}^{T_2} A^N(t)dw_t)(\int_{T_0}^{T_1} A^N(t)dw_t)^T \right]
\]

\[
= \lim_{N \to \infty} E \left[ \int_{T_0}^{T_1} A^N(t)dw_t + \int_{T_1}^{T_2} A^N(t)dw_t)(\int_{T_0}^{T_1} A^N(t)dw_t)^T \right]
\]

(by Lemma 4.3)

\[
= \lim_{N \to \infty} \int_{T_0}^{T_1} A^N(t)A^{NT}(t)dt
\]

\[
= \int_{T_0}^{T_1} A(t)A(t)dt.
\]
We have now constructed stochastic integrals of deterministic functions with respect to the Wiener process and may consider linear SDEs of the form

\[ dx_t = A_t x_t dt + B_t u_t dt + C_t dw_t, \quad t \geq t_0, \]  

(12)

where \( x_{t_0} \) is a random variable such that \( E\|x_{t_0}\|^2 < \infty, x_{t_0} \parallel w \) and \( A, B, C \) and \( u \) are piecewise continuous functions on the finite interval \([t_0, t]\). (That is to say, functions which are continuous on the interiors of a finite set of intervals whose union is \([t_0, t]\) and which are everywhere continuous from the right and everywhere have finite limits from the left). Then (12) is interpreted to be equivalent to the stochastic integral equation

SIE \[ x_t = x_{t_0} + \int_{t_0}^{t} A_s x_s ds + \int_{t_0}^{t} B_s u_s ds + \int_{t_0}^{t} C_s dw_s, \quad t \geq t_0, \]  

with the stated conditions on \( x_{t_0}, w \). A solution to the SDE is then defined to be any solution to the SIE above which is w.p.1 continuous and satisfies the given hypotheses on \( x_{t_0} \). Then it may be verified that

\[ x_t = \Phi(t, t_0) x_{t_0} + \int_{t_0}^{t} \Phi(t, s) B_s u_s ds + \int_{t_0}^{t} \Phi(t, s) C_s dw_s, \quad t \geq t_0, \]  

(13)

with the stated conditions on \( x_{t_0}, w \) gives a well defined stochastic process \( x \) when

\[ \frac{d}{dt} \Phi(t, s) = A(t) \Phi(t, s), \quad \Phi(s, s) = I, \]

for all \( t \geq s, \quad s, t \in \mathbb{R} \), defines the fundamental matrix \( \Phi(\cdot, \cdot) \) of \( A(\cdot) \).

Furthermore, it may be verified that when \( E x_{t_0} = 0, \) and \( E x_{t_0} x_{t_0}^T = \Sigma, \) (12) gives the unique solution to (12) and the covariance function of \( \{ x_t - \int_{t_0}^{t} \Phi(t, s) u_s ds; t \geq t_0 \} \) is given by

\[ E(x_t - \int_{t_0}^{t} \Phi(t, s) u_s ds)(x_{t+\tau} - \int_{t_0}^{t+\tau} \Phi(t+\tau, s) u_s ds)^T = E(\Phi(t; t_0) x_{t_0})(\Phi(t + \tau, t_0) x_{t_0})^T + \int_{t_0}^{t} \Phi(t, s) C_s C_s^T \Phi(t + \tau, s) ds, \quad \tau > 0, \]  

18
with a corresponding expression in case $\tau \leq 0$. Evaluating the first covariance on the right hand side of the equation above, and setting $u \equiv 0$ for simplicity, we obtain

$$Ex_t x_{t+\tau}^T = \Phi(t, t_0)\Sigma \Phi^T(t + \tau, t_0) + \int_{t_0}^t \Phi(t, s)C_s C_s^T \Phi^T(t + \tau, s)ds.$$  

In general (see e.g. [Ikeda-Watanabe, 1981]) it may be shown that under reasonable smoothness and growth conditions an $f(\cdot, \cdot)$ and $g(\cdot, \cdot)$, a solution process exists for the general non-linear SDE

$$dx_t = f(x_t, u_t)dt + g(x_t, u_t)dw_t, \quad t \geq t_0,$$

when $x_{t_0}$ is a random initial condition satisfying $E\|x_{t_0}\|^2 < \infty$ and $x_0 \sim \mathcal{N} w$, and $u$ is a continuous function.

**Continuous Time Parameter Linear Filtering: The Kalman-Bucy Filter**

In analogy with the discrete time state estimation problem, the continuous time Kalman-Bucy filtering problem is formulated as follows: the system equation is given by

$$dx_t = Fx_t dt + Q^{1/2} dw_t, \quad t \geq t_0, \quad x_t \in \mathbb{R}^n, w_t \in \mathbb{R}^w, \quad (14)$$

and the observation equation is

$$dy_t = Hx_t dt + R^{1/2} dv_t, \quad t \geq t_0, \quad y_t \in \mathbb{R}^p, v_t \in \mathbb{R}^v. \quad (15)$$

Here $w$ and $v$ are Wiener processes such that for all $t, s \in \mathbb{R},$

$$E(w_t - w_0)(v_s - v_0) = 0,$$

$$E(w_t - w_0)(w_s - w_0) = I_w \min(t, s),$$

$$E(v_t - v_0)(v_s - v_0) = I_v \min(t, s),$$

and such that
\[ Ex_{t_{0}}w_{t}^{T} = 0, Ex_{t_{0}}v_{s}^{T} = 0 \quad s \geq t_{0}, \quad t \geq t_{0}. \quad (16) \]

where \( x_{t_{0}} \) has a joint Gaussian distribution with \( w_{t}, v_{t} \) for all \( t \geq t_{0} \).

In this case the corresponding continuous time parameter equations generating \( \hat{x}_{t} = E_{t_{0}}x_{t} \) have the well known Kalman-Bucy filter form

\[
d\hat{x}_{t} = F\hat{x}_{t}dt + \Delta_{t}[dy_{t} - H\hat{x}_{t}dt], \quad t \geq t_{0}, \quad (17)\]

with

\[
\Delta_{t} = V_{t}H^{T}R^{-1}, \quad t \geq t_{0},
\]

where the associated Riccati equation generating the state estimation error covariance process \( \{V_{t}; t \geq t_{0}\} \) is given by the deterministic equation

\[
\frac{dV_{t}}{dt} = FV_{t} + V_{t}F^{T} - V_{t}H^{T}R^{-1}HV_{t} + Q, \quad t \geq t_{0}, \quad (18)\]

\[
V_{t_{0}} = E(x_{t_{0}} - \hat{x}_{t_{0}})(x_{t_{0}} - \hat{x}_{t_{0}})^{T}.
\]

For a clear derivation of these equations see e.g. [Davis]. In the continuous time case, as in the discrete time parameter case, the filter equations for a system and observation model with deterministic time varying matrices is obtained by merely subscripting the matrices in (17), (18) by the time variable \( t \).
Stationary Continuous Time Processes Generated by LSSS

We recall the following definition of a strictly stationary stochastic process.

**Definition 5.2.** A **strictly stationary stochastic process** is a stochastic process for which all finite dimensional distributions are **shift invariant** with respect to the time parameter, i.e. when $T$ is an interval in $\mathbb{R}^1$ or $\mathbb{Z}$, for all $N \in \mathbb{Z}^+$, all $t_1, \cdots, t_N \in T$, and all $\tau \in \mathbb{R}^1$ such that $t_i + \tau \in T$, $1 \leq i \leq N$,

$$
P(x_{t_1} \in A_1, x_{t_2} \in A_2, \cdots, x_{t_N} \in A_N) = P(x_{t_1+\tau} \in A_1, x_{t_2+\tau} \in A_2, \cdots, x_{t_N+\tau} \in A_N)
$$

And we also recall that for any sequence of vector valued random variables is said to converge **in mean square**, or in **quadratic mean**, to a random variable $x$ (denoted $\text{m.s.lim}_{n \to \infty} x_n$) defined on the same underlying probability space if

$$
\lim_{n \to \infty} E\|x_n - x\|^2 = 0.
$$

For simplicity, we begin by considering the scalar process $x$ generated by the SDE

$$
dx_{t,-T} = -\alpha x_{t,-T} dt + dw_t, \quad \alpha > 0, x_{-T,-T} \sim N(0, \Sigma),
$$

with $x_{IC} \prod w_{-T}$ for each $T \in \mathbb{R}$. In order to have a well defined sample path with respect to $T, T \to \infty$, we take $w_t, t < 0$, to be $\tilde{w}_{-t}$ for a Wiener process $\tilde{w}$ on $[0, \infty)$, and for $t \geq 0$, take $w_t = v_t, v$ a Wiener process independent of $\tilde{w}$. There is no loss of generality in taking $w_0 = 0$ since it is only the increments of $w$ which enter (??). Its solution

$$
x_{t,-T} = e^{-\alpha(t+T)} x_{IC} + \int_{-T}^{t} e^{-\alpha(t-\tau)} dw_{\tau}
$$

satisfies

$$
x_{t,-T} \to \int_{-\infty}^{t} e^{-\alpha(t-\tau)} dw_{\tau} \prod x_{t}^{\infty}, \quad t \in \mathbb{R}.
$$
in mean square as $T \to \infty$. In order to find the mean and covariance function of the limiting stochastic process $x_{t}^{\infty}$, we proceed as follows. Applying Lemma 5.3, and using the fact the mean square (i.e. $L^2$) limit of $e^{-\alpha(t+T)} x_{IC}$ is 0, we obtain
\[ E_{t}^{\infty} = \text{Em.s.} \lim_{T \to \infty} \left[ e^{-\alpha(t+T)}x_{IC} + \int_{-T}^{t} e^{-\alpha(t-\tau)}d\tau \right] = \lim_{t \to \infty} E \left[ \int_{-T}^{t} e^{-\alpha(t-\tau)}d\tau \right] = 0. \]

Applying Lemma 5.3 again yields

\[ E_{t_1}^{\infty}x_{t_2}^{\infty} = \]

\[ E \left[ \text{m.s.} \lim_{T \to \infty} \left( e^{-\alpha(t_1+T)}x_{IC} + \int_{-T}^{t_1} e^{-\alpha(t_1+\tau)}d\tau \right) \right] \left[ \text{m.s.} \lim_{T' \to \infty} \left( e^{-\alpha(t_2+T')}x_{IC} + \int_{-T'}^{t} e^{-\alpha(t_2-\tau)}d\tau \right) \right] \]

\[ = \lim_{T,T' \to \infty} E \int_{-T}^{t_1} \int_{-T'}^{t_2} e^{-\alpha(t_1-\tau)}e^{-\alpha(t_2-\rho)}d\tau d\rho + 0, \]

since \( \int_{-T}^{t} e^{-\alpha(t-\tau)}d\tau \) and \( e^{-\alpha(t+Tn)}x_{IC} \) are \( L^2 \) Cauchy sequences (reader check).

Hence

\[ E_{t_1}^{\infty}x_{t_2}^{\infty} = \lim_{T,T' \to \infty} \left\{ E \int_{-T}^{t_1 \wedge t_2} \int_{-T}^{t_1 \wedge t_2} e^{-\alpha(t_1-\tau)-\alpha(t_2-\rho)}d\tau d\rho \right. \]

\[ + E \int_{-T}^{t_1 \wedge t_2} \int_{-T'}^{t_1 \wedge t_2} e^{-\alpha(t_1-\tau)-\alpha(t_1-\rho)}d\tau d\rho \]

\[ + E \int_{t_1 \wedge t_2}^{t_2 \wedge T} \int_{-T'}^{t_1 \wedge t_2} e^{-\alpha(t_1-\tau)-\alpha(t_2-\rho)}d\tau d\rho \left\} \right. \]

\[ = \lim_{T \to \infty} \int_{-T}^{t_1 \wedge t_2} e^{-\alpha(t_1+t_2)+2\alpha \tau}d\tau + 0, \]

where the reader is recommended to check the decomposition of the double integral via the appropriate diagram, and where we have taken \( -T > -T' \) for definiteness, the converse case being similar. Hence

\[ E_{t_1}^{\infty}x_{t_2}^{\infty} = \int_{-\infty}^{t_1 \wedge t_2} e^{-\alpha(t_1+t_2)+2\alpha \tau}d\tau = \frac{1}{2\alpha} e^{-\alpha(t_1+t_2)} e^{2\alpha(t_1 \wedge t_2)}. \]

Let \( t_1 < t_2 \) (the case \( t_2 < t_1 \) is symmetric to this), then

\[ E_{t_1}^{\infty}x_{t_2}^{\infty} = \frac{1}{2\alpha} e^{-\alpha(t_2-t_1)} = \frac{1}{2\alpha} e^{-\alpha((t_2-\tau)-(t_1-\tau))} = E_{t_1-\tau}^{\infty}x_{t_2-\tau}^{\infty} \]
and hence $R_{t_1,t_2}$ is shift invariant. In particular, $E x_t^2 = \frac{1}{2\alpha}$.

Because $L^2$ convergence implies convergence in probability, the FDDs of a process $x^\infty$ are the limits of the FDDs of any sequence of processes converging in mean square to $x^\infty$; for instance, for any $t_1, t_2 \in \mathbb{R}, A_1, A_2 \in \mathcal{B}(\mathbb{R}^n)$, $P(x_{t_1}^\infty \in A_1, x_{t_2}^\infty \in A_2) = \lim_{n \to \infty} P(x_{t_1}^n \in A_1, x_{t_2}^n \in A_2)$ where $x^n \to x^\infty$ in m.s. as $n \to \infty$. Hence the process $x^\infty = \{x_t^\infty; t \in \mathbb{R}\}$ has a family of finite dimensional distributions which can (in principle) all be computed by taking the limit in the calculation of all possible moments of the corresponding Gaussian process $\{x_{t,-T}; t \geq -T\}$. (Equivalently, one may consider the limit of the characteristic function

$$E\{\exp i\omega \left( \sum_{j=1}^M \alpha_j x_{t_j,-T} \right) \}$$

for all $M \in \mathbb{Z}_1$, and all $(\alpha_1, \cdots, \alpha_M) \in \mathbb{R}^M, \omega \in \mathbb{R}$.)

But it may be verified that these limits equal the corresponding moments (or equivalently characteristic function values) for a Gaussian process with mean 0 and covariance

$$R_{t_1,t_2} = R_{t_2,t_3} = \frac{1}{2\alpha} e^{-\alpha |t_2-t_1|}, \quad t_1, t_2 \in \mathbb{R}.$$  

Hence the limiting process $x_t^\infty$ given as the mean square limit of $x_{t,-T}, t \in \mathbb{R}$ as $T \to \infty$, is a Gaussian process which is strictly stationary.
Stationary Continuous Time LS\(^3\) Processes: The Lyapunov Equation

Continuous time results for (6.1), which are analogous to the discrete time results in the previous subsection, have a somewhat more elegant appearance. Consider the continuous time system with \(n\) dimensional state process and \(m\) dimensional Brownian motion input given by

\[\begin{align*}
\text{LS}^3 \quad dx_t &= Ax_t dt + C dw_t, \quad t \in \mathbb{R}_+, \\
\end{align*}\]  

(20)

with \(x_0 \sim N(0, \Sigma), \Sigma_0 < \infty\) and \(x_0 \prod w\). Assume that \(A\) is asymptotically stable, i.e. \(\max_{1 \leq i \leq n}(\text{Re} \lambda_i(A)) < 0\).

From (??) with \(u = 0\) we obtain

\[
R_{t_0}^{t_0} = \Phi(t, t_0) \Sigma_0 \Phi^T(t, t_0) + \int_{t_0}^t \Phi(t + \tau, u)CC^T \Phi^T(t, u) du.
\]

Let \(\Pi(t, t_0) \triangleq R_{t_0}^{t_0}\), then differentiation yields the (time varying) Lyapunov equation

\[
\frac{d}{dt} \Pi(t, t_0) = A \Pi(t, t_0) + \Pi(t, t_0) A^T + CC^T, \quad t \geq t_0,
\]

\[
\Pi(t_0, t_0) = \Sigma_0.
\]

(21)

Next consider \(\Pi(t) - \Pi_\infty \triangleq \Pi(t, t_0) - \Pi_\infty\), where \(\Pi_\infty\) is assumed to satisfy the (time invariant) Lyapunov equation

\[\Pi_\infty A^T + A \Pi_\infty + CC^T = 0.\]  

(22)

Then

\[
\frac{d}{dt}(\Pi_t - \Pi_\infty) = \Pi_t A^T + A \Pi_t + CC^T - 0
\]

\[
= (\Pi_t - \Pi_\infty) A^T + A(\Pi_t - \Pi_\infty),
\]

with initial condition \(\Pi_{t_0} - \Pi_\infty = \Sigma - \Pi_\infty\). Hence

\[\Pi_t - \Pi_\infty = e^{A(t-t_0)}(\Sigma - \Pi_\infty)e^{A^T(t-t_0)}, \quad t \geq t_0,\]

and the asymptotic stability of \(A\) implies

\[\Pi_t - \Pi_\infty \to 0\]
as \( t \to \infty \). To show (??) has a solution, we observe that \( P_T(X) \Delta \int_0^T e^{As} C C^T e^{At} ds \) converges to a finite limit for any matrix \( X \), because \( A \) is asymptotically stable. But
\[
\int_0^T \frac{d}{ds} \left[ e^{As} C C^T e^{At} \right] ds = e^{AT} C C^T e^{AT} - C C^T,
\]
and also
\[
\int_0^T \frac{d}{ds} \left[ e^{As} C C^T e^{At} \right] ds = A \int_0^T e^{As} C C^T e^{At} ds + \int_0^T e^{As} C C^T e^{At} ds A^T
\]
\[
\nabla A P_T(I) + P_T(I) A^T.
\]
Then since \( P_T(I) \to \int_0^\infty e^{As} C C^T e^{At} ds \) and \( e^{AT} C C^T e^{AT} \to 0 \) as \( T \to \infty \) we obtain
\[
A P_\infty(I) + P_\infty(I) A^T = -C C^T,
\]
which shows (??) has a solution in the given integral form. Next, since the asymptotic stability of \( A \) implies the solution to (??) is unique (reader check), we may identify \( P_\infty \) and \( P_\infty(I) \) to obtain
\[
\Pi_t \to \Pi_\infty = P_\infty(I) = \int_0^\infty e^{As} C C^T e^{At} ds,
\]
as \( t \to \infty \).

Finally we see that the mean square limiting process \( x_\infty \) satisfies
\[
E x_{t+\tau}^\infty x_t^\infty = E m.s.lim_{T \to \infty} (e^{At} x_{t-T} + \int_t^{t+\tau} e^{A(t+\tau-s)} dw_s) m.s.lim_{T \to \infty} x_{t-T}^T \quad \tau \geq 0
\]
\[
= e^{At} \lim_{T \to \infty} E x_{t-T} x_{t-T}^T + 0
\]
\[
= e^{At} \left[ \int_{-\infty}^t e^{As} C C^T A^T ds \right] \quad \text{(by Lemma 5.3)}
\]
\[
= e^{At} \Pi_\infty,
\]
and similarly
\[
E x_{t+\tau}^\infty x_t^\infty = E x_\infty x_{t-\tau}^\infty = \Pi_\infty e^{-A T}, \quad \tau < 0.
\]
This last property may also be seen to hold by taking the initial state of the linear system to be distributed according to the invariant distribution \( N(O, \Pi_\infty) \). Then at any later instant
\[
E x_{t+\tau}^\infty x_t^\infty = e^{At} \Pi_t = e^{At} \Pi_\infty.
\]
These results are summarized in the theorem statement below.
Theorem 5.2

For the asymptotically stable continuous time LS3 system (??) the properties of a continuous time process $x^\infty$ corresponding to those of (i)-(iv) of Theorem 5.1 hold, with (??) replaced by

$$0 = A\Pi_\infty + \Pi_\infty A^T + CC^T$$

and with the covariance functional relation replaced by

$$E_{x_t^\infty, x_t^\infty^T} = \begin{cases} e^{A\tau}\Pi_\infty, & \tau \geq 0, \\ \Pi_\infty e^{-A^T\tau}, & \tau < 0. \end{cases}$$
Second Order Processes

Just as for a discrete time process, a continuous time wide sense stationary $\mathbb{R}^n$ valued process is defined as a second order process for which

$$Ex_t = \mu, \quad \forall t \in \mathbb{R}^1,$$

$$R(t, s) = E(x_t - \mu)(x_s - \mu) = E(x_{t-s} - \mu)(x_0 - \mu)$$

$$\nabla R(t - s), \quad \forall t, s \in \mathbb{R}^1.$$

Example 5.6

As in Theorem 5.2, the scalar $LS^3$ process $x$ generated by

$$dx_t = \alpha x_t dt + q dw_t, \quad \alpha < 0, \quad t \in \mathbb{R}_+,$$

with $x_0 \sim N(0, \pi), x_0 \prod w$ and $\pi$ satisfying $0 = \alpha \pi + \pi \alpha + q^2$, is a wss process on $\mathbb{R}_+$. From the explicit calculation for the system (?), $Ex_t x_s = R(t - s) = \frac{-q^2}{2\alpha} e^{-\alpha|t-s|}$, and for any distribution $N(0, \sigma_0^2)$ for the initial condition $x_0$,

$$E x_{T+t} x_{T+s} \to R(t - s),$$

as $T \to \infty$.

Definition 5.6

An $\mathbb{R}^n$ valued stochastic process $\{x_t; t \in \mathbb{R}\}$ is quadratic mean (q.m.) continuous if

$$E\|x_{t+h} - x_t\|^2 \to 0 \text{ as } h \to 0 \text{ for all } t \in \mathbb{R}.$$

When it exists, the q.m. limit

$$\frac{d}{dt} x_t \triangleq q.m.lim_{h \to 0} \frac{1}{h} (x_{t+h} - x_t)$$

(24)

of a second order process $x$ is called the quadratic mean (q.m.) derivative of $x$ at $t \in \mathbb{R}$. If $\frac{d}{dt} x_t$ exists for all $t \in \mathbb{R}$ then $x$ is called a q.m. differentiable process.
It may be verified that a zero mean second order process for which the covariance \( R(t, s) \)
is jointly continuous in \((t, s)\) at \((u, u)\) for any \(u \in \mathbb{R}\) is q.m. continuous for all \(t \in \mathbb{R}\). Conversely, a second order process \(x\) which is q.m. continuous for all \(t, s \in \mathbb{R}\) has a covariance function which is continuous in \((t, s)\).

**Theorem 5.6**

(a) A second order process \(x\) is q.m. differentiable if \(\frac{\partial^2 R(t, s)}{\partial t \partial s}\) exists and is continuous at \((t, t)\) for all \(t \in \mathbb{R}\).

(b) If a second order process \(x\) is q.m. differentiable, then the second partial differentials taken in either order exist.

**Proof**

For notational convenience we shall only consider scalar processes in this proof.

(a) To begin, observe that the limit (25) exists if and only if the Cauchy condition

\[
\lim_{h, h' \to 0} E\left( \frac{1}{h} (x_{t+h} - x_t) - \frac{1}{h'} (x_{t+h'} - x_t) \right)^2 = 0,
\]

holds.

Now the existence and continuity at \((t, t)\) of the mixed partial derivative \(\frac{\partial^2 R(t, s)}{\partial t \partial s}\) of \(R(t, s)\) implies that its value is independent of the order of evaluation of the partial derivatives. This implies that a unique limit

\[
\lim_{(h, h') \to (0, 0)} h'^{-1}\{h^{-1}[R(t+h', t+h) - R(t+h', t)] - h^{-1}[R(t+h, t) - R(t, t)]\}
\]

implies that its value is independent of the order of evaluation of the partial derivatives. This implies that a unique limit

\[
\lim_{(h, h') \to (0, 0)} (hh')^{-1}\{(Ex_{t+h'} x_{t+h} - Ex_{t+h', t}) - (Ex_{t+h} x_{t+h'} - Ex_{t+h'} x_t)\} = \lim_{(h, h') \to 0} E \Delta (h, h')
\]
exists for all sequences \((h, h')\) converging to \((0, 0)\), where

\[
\Delta (h, h') \triangleq (hh')^{-1}(x_{t+h} - x_t)(x_{t+h'} - x_t).
\]

Hence

\[
\lim_{(h, h') \to (0, 0)} (E[\Delta (h, h) + \Delta (h', h') - 2 \Delta (h, h')]) = \lim_{h \to 0} E \Delta (h, h) + \lim_{h \to 0} E \Delta (h', h') - \lim_{(h, h') \to (0, 0)} 2E \Delta (h, h') = 0.
\]

However, the left-most expression above is equal to \(E(\frac{1}{h}(x_{t+h} - x_t)) - (\frac{1}{h'}(x_{t+h'} - x_t))^2\) (reader check), and hence the proven convergence to zero establishes that the differences \(\{h^{-1}(x_{t+h} - x_t); h \in \mathbb{R}\}\) form an \(L^2\) Cauchy sequence as \(h \to 0\). This establishes the existence of the q.m. differential of the process \(x\) for \(t \in \mathbb{R}_+\).

(b) If the process \(x\) is q.m. differentiable, the approximating differences form an \(L^2\) Cauchy sequence as \(h \to 0\). Hence the approximating differences for the second partial derivatives in either order exist, as required.

For processes which are q.m. differentiable we evidently have

\[
E(\frac{d}{dt}x)^2 = \frac{\partial^2 R(t, t)}{\partial s \partial t} \bigg|_{t=s} \quad \forall t \in \mathbb{R}_+.
\]

**Example 5.6**

A Wiener process is q.m. continuous at any \(t \in \mathbb{R}_+\), and this fact is implied by the joint continuity of the covariance function \(\min(t, s)I; t, s \in \mathbb{R}_+\). However, the second partial derivatives do not exist for the covariance function of a Wiener process, and this implies that it is not q.m. differentiable. This, of course, is evident by a direct calculation involving approximating differences of the Wiener process.
Spectral Theory for WSS Continuous Time Stochastic Process

Let $x$ be an $\mathbb{R}^n$ valued zero mean q.m. continuous wss stochastic processes with matrix covariance function $R(\tau); \tau \in \mathbb{R}^1$. Assume $\int_{-\infty}^{\infty} \|R(\tau)\|d\tau \leq \infty$.

Since $x$ is q.m. continuous, $R(.)$ is continuous and so $R(.) \in L_1 \cap C_0$, hence we may define the spectral density matrix of $x$ via

$$\Phi(\omega) \triangleq \int_{-\infty}^{\infty} e^{-2\pi i \omega t} R(t)dt, \quad \omega \in \mathbb{R}.$$

When $i\omega$ replaces $e^{i\theta}$ in the matrix function $\{\Phi(\omega); \omega \in \mathbb{R}\}$ it can be seen that the three conjugate transpose, conjugate and positivity properties of a spectral density matrix hold as given in Definition 5.3 for (discrete time) spectral density matrices; namely $\Phi(.)$ is a positive complex Hermitian matrix.

Since $R(.) \in L_1 \cap C_0$, it may be verified that $\Phi(.) \in L_1 \cap C_0$, and so the inversion formula

$$Ex_{t+\tau}x_t = R(\tau) = \int_{-\infty}^{\infty} e^{2\pi i \omega \tau} \Phi(\omega)d\omega, \quad \omega \in \mathbb{R},$$

holds. Next, in analogy with the discrete time case, we may consider the action of a (not necessarily non-anticipative) linear system with $L^2$ matrix impulse response $\{H(\tau); \tau \in \mathbb{R}\}$ on a wss stochastic process satisfying the hypotheses of this section.

In the time domain we may define the output process $y$ by

$$y_t = \int_{-\infty}^{\infty} H(\tau)x(t-\tau)d\tau, \quad \tau \in \mathbb{R}.$$

Since $H(.) \in L^2$ the transfer function

$$H(\omega) \triangleq \int_{-\infty}^{\infty} e^{-2\pi i \omega \tau} H(t)d\tau, \quad \omega \in \mathbb{R},$$

also lies in $L^2$. Finally this gives the fundamental relation
\[ R_y(\tau) = E y_{t+\tau} y_t \]
\[ = E \left( \int_{-\infty}^{\infty} H(p) x(t+\tau-p) dp \right) \left( \int_{-\infty}^{\infty} H(q) x(t-q) dq \right) \]
\[ = \int_{-\infty}^{\infty} e^{2\pi i \omega \tau} H(\omega) \Phi_x(\omega) \overline{H(\omega)}^T d\omega, \ \omega \in \mathbb{R}. \]

The formula above immediately yields the continuous time version of the Wiener-Khinchin formula for the class of processes under consideration; namely

\[ \Phi_y(\omega) = H(\omega) \Phi_x(\omega) \overline{H(\omega)}^T, \ \omega \in \mathbb{R}. \]

**Orthogonal Representations: The Karhunen-Loeve Expansion**

Again in this section we assume all processes to have zero mean. Let \( x \) be a q.m. continuous second order (not necessarily stationary) stochastic process taking values in the Hilbert space \( H \). For convenience we take \( x \) to be a scalar process. We seek a set of random second order time independent basis functions for \( H \), say \( \{ \psi_n(\omega); n \in \mathbb{Z}_+ \} \), such that a representation of the form

\[ x_t(\omega) = \sum_{n=0}^{\infty} \alpha_n,t \psi_n(\omega) \]

\[ t \in \mathbb{R}_+, \quad (26) \]

holds for some set of time dependent coefficients \( \{ \alpha_{n,t}; n \in \mathbb{Z}_+ \} \) which are non-random. Moreover, we seek an orthonormal (o.n.) complex valued family \( \{ Z_n(\omega); n \in \mathbb{Z}_+ \} \) such that

\[ E|Z_n(\omega)|^2 = 1, n \in \mathbb{Z}_+, \]

and

\[ \langle Z_n, Z_m \rangle \triangleq E Z_n(\omega) \overline{Z_m(\omega)} = 0 \quad n \neq m, \quad n, m \in \mathbb{Z}_+. \]

Consider the non-random set of time functions

\[ \sigma_n(t) \triangleq \langle x_t, Z_n \rangle = E x_t(\omega) Z_n(\omega), \quad n \in \mathbb{Z}_+ \]
where it is assumed that \( \{Z_n; n \in \mathbb{Z}_+\} \) is an o.n. family.

Then
\[
0 \leq E|x_t - \sum_{n=0}^{\infty} \langle x_t, Z_n \rangle Z_n|^2 \\
= E|x_t|^2 - 2 \sum_{n=0}^{\infty} \langle x_t, Z_n \rangle \langle x_t, Z_n \rangle + \sum_{n=0}^{\infty} \langle x_t, Z_n \rangle^2 \\
= E|x_t|^2 - \sum_{n=0}^{\infty} |\langle x_t, Z_n \rangle|^2.
\]

Consider the Hilbert space \( H \) and assume that for all \( x_t \in H \)
\[
\|x_t\|^2 = \langle x_t, x_t \rangle = \sum_{n=0}^{\infty} |\langle x_t, Z_n \rangle|^2;
\]
then we say \( \{Z_n; n \in \mathbb{Z}_+\} \) is a complete o.n. (c.o.n.) family for the space \( H \).

In this case we necessarily have
\[
x_t(\omega) = \text{q.m.lim}_{N \to \infty} \sum_{n=0}^{N} \sigma_n(t)Z_n(\omega)
\]
\[
\nabla \sum_{n=0}^{\infty} \sigma_n(t)Z_n(\omega).
\]
The functions \( \{\sigma_n(t); t \in \mathbb{R}\} \) may be verified to be continuous because \( x \) is q.m. continuous. (Reader check.)

Assume \( x_t \) is not contained in any proper subspace of \( \{Z_n\}_{n=0}^{\infty} \). Then the functions \( \sigma_n(\cdot) \) are linearly independent; otherwise
\[
\sum_{n=0}^{N} a_n \sigma_n(t) = 0, \quad \text{for some} \quad \{a_n\}_{0}^{N},
\]
and hence
\[
0 = \sum_{n=0}^{N} a_n \langle x_t, Z_n \rangle = \langle x_t, \sum_{n=0}^{N} a_n Z_n \rangle;
\]
which implies
\[
x_t \perp \left\{ \sum_{n=0}^{N} a_n Z_n \right\} \quad \text{while} \quad x_t = \sum_{n=0}^{\infty} \alpha_n(t)Z_n.
\]
But this contradicts the proper subspace condition.

**Separable Covariance Function**

Suppose \( x_t = \sum_{n=0}^{\infty} \sigma_n(t)Z_n, \ t \in \mathbb{R}, \) where \( \{Z_n; n \in \mathbb{Z}_+\} \) is a c.o.n. family. Then

\[
R(t, s) = E x_t x_s = E[\sum_{n=0}^{\infty} \sigma_n(t)Z_n][\sum_{m=0}^{\infty} \sigma_m(s)\overline{Z}_m] = \sum_{n=0}^{\infty} \sigma_n(t)\overline{\sigma}_n(s), \quad \forall t, s \in \mathbb{R}. \tag{27}
\]

A covariance function of the form (27) is called *separable*.

Mercer’s Theorem states that continuous positive functions \( R(s, t); s, t \in \mathbb{R} \) on \( L^2 \times L^2 \) possess complete orthonormal families of eigenfunctions with respect to which they have expansions of the form (27) which converge uniformly over compact sets of the form \( \{-M \leq s, t \leq M; s, t \in \mathbb{R}\} \). Hence, in particular, continuous covariance functions are separable.

Clearly, if

\[
R(s, t) = \sum_{n=-\infty}^{\infty} \sigma_n(s)\overline{\sigma}_n(t), \quad s, t \in \mathbb{R}, \tag{28}
\]

with \( \int_{-\infty}^{\infty} \sigma_n(s)\sigma_m(s)ds = \lambda_n \delta_{nm}, n, m \in \mathbb{Z}, \) then

\[
\int_{-\infty}^{\infty} R(s, t)\sigma_n(t)dt = \int_{-\infty}^{\infty} \left( \sum_{k=0}^{\infty} \sigma_k(s)\overline{\sigma}_k(t) \right) \sigma_n(t)dt
\]

\[
= \sum_{k=0}^{\infty} \sigma_k(s)\lambda_n \delta_{k,n}
\]

\[
= \lambda_n \sigma_n(s), \quad \forall s \in \mathbb{R}.
\]

**Theorem 5.7 Karhunen-Loeve**

Let \( x \) be a q.m. continous second order process with covariance function \( R(t, s), t, s \in \mathbb{R}. \)

(a) Let \( \{\psi_n; n \in \mathbb{Z}_+\} \) be the set of orthonormal eigenfunctions of \( R(\cdot, \cdot) \) such that

\[
\int_{-\infty}^{\infty} R(t, s)\psi_n(s)ds = \lambda_n \psi_n(t), \quad \forall t \in \mathbb{R},
\]
i.e. for which \( \{ \lambda_n; n \in \mathbb{Z}_+ \} \) is the set of eigenvalues. Further let
\[
b_n(\omega) = \sqrt{\lambda_n} \int_{-\infty}^{\infty} x(\omega, t) \overline{\psi}_n(t) \, dt
\]
for which, necessarily,
\[
E b_n b_m = \delta_{m,n}, \quad m, n \in \mathbb{Z}_+.
\]
Then
\[
x(\omega, t) = \text{q.m.lim}_{N \to \infty} \sum_{n=0}^{N} \sqrt{\lambda_n} \psi_n(t) b_n(\omega)
\]
uniformly on compact intervals.

(b) Conversely, if \( x(\omega, t) \) has an expansion of the form
\[
\int_{-\infty}^{\infty} \psi_m(t) \overline{\psi}_n(t) \, dt = \delta_{m,n} = E b_m \overline{b_n} \quad m, n \in \mathbb{Z}_+,
\]
then \( \{ \psi_n; n \in \mathbb{Z}_+ \} \) and the associated \( \{ \lambda_n; n \in \mathbb{Z}_+ \} \) are the eigenfunctions and eigenvalues respectively of \( R(\cdot, \cdot) \).

Note that for real processes \( x, \psi \) and \( \lambda \) are real.

**Proof**

(a)

From the hypotheses of the theorem and the definition of the terms, we see that \( \{ b_n; n \in \mathbb{Z}_+ \} \) is a c.o.n. for the Hilbert space \( \mathbf{H} \) spanned by \( x \). Consequently, the result follows directly from
\[
E|x_t - \sum_{n=0}^{N} \sqrt{\lambda_n} \psi_n(t) b_n|^2
\]
\[
= R(t, t) - \sum_{n=0}^{N} \lambda_n |\psi_n(t)|^2 \to 0
\]
as \( N \to \infty \), where the latter convergence follows from Mercer’s Theorem.
(b) If \( x_t = \sum_{n=0}^{\infty} \sqrt{\lambda_n} \psi_n(t) b_n(\omega) \), then

\[
R(t, s) = \text{Ex}_t \mathbb{E}_s = \sum_{n=0}^{\infty} \lambda_n \psi_n(t) \overline{\psi_n(s)}.
\]

Hence

\[
\int_{-\infty}^{\infty} R(t, s) \psi_m(s) ds = \int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \lambda_n \psi_n(t) \overline{\psi_n(s)} \psi_m(s) ds = \lambda_m \psi_m(t), \quad \forall m \in \mathbb{Z}_+, \forall t \in \mathbb{R}.
\]

\[\blacksquare\]

**Example 5.7 (Wong [1971, p.87])**

Let \( R(t, s) = \min(t, s), t, s \in \mathbb{R} \); in other words \( R(\cdot, \cdot) \) is the covariance of the Wiener process. Consider

\[
\int_0^T \min(t, s) \psi(s) ds = \lambda \psi(t), \quad 0 \leq t \leq T,
\]

or, equivalently,

\[
\int_0^t s \psi(s) ds + t \int_t^T \psi(s) ds = \lambda \psi(t), \quad 0 \leq s, t \leq T.
\]

This gives

\[
t \psi(t) - t \psi(t) + \int_t^T \psi(s) ds = \lambda \frac{d}{dt} \psi(t),
\]

and so

\[
\ddot{\psi}(t) = -\frac{1}{\lambda} \psi(t), \quad \lambda \neq 0,
\]

with \( \psi(0) = 0, \frac{d}{dt} \psi(T) = 0 \). This gives

\[
\psi(t) = A \sin \frac{1}{\sqrt{\lambda}} t, \quad \text{with} \quad \cos \frac{T}{\sqrt{\lambda}} = 0, \quad \text{implying} \quad \sqrt{\lambda} = \frac{2T}{(2n + 1)\pi}, n \in \mathbb{Z}_+,
\]

where here, since \( \sin(-x) = -\sin x \), we do not need to consider the negative integers \( n \in \mathbb{Z} \) in order to find additional eigenfunctions.

So, on \( \mathbb{Z}_+ \), the set of normalized eigenfunctions are
\[ \psi_n(t) = \left( \frac{2}{T} \right)^{1/2} \sin \left( \frac{2n + 1}{2} \right) \left( \frac{\pi t}{T} \right), \quad n \in \mathbb{Z}_+, \]

and hence the process \( x \) with the specified covariance \( R(t, s) \) satisfies

\[ x_t = \text{q.m.lim}_{N \to \infty} \sum_{n=0}^{N} \left( \frac{2T}{(2n + 1)\pi} \right) \left( \frac{2}{T} \right)^{1/2} \left( \sin \left( \frac{2n + 1}{T} \left( \frac{\pi t}{2} \right) \right) \right) b_n(\omega), \]

where

\[ b_n(\omega) = \left( \frac{2T}{(2n + 1)\pi} \right)^{-1} \int_0^T \left( \frac{2}{T} \right)^{1/2} \left( \sin \left( \frac{2n + 1}{2} \left( \frac{\pi t}{T} \right) \right) \right) x(\omega, t) dt \]

in q.m. Incidentally, as observed by Wong, Mercer’s Theorem in this case gives the expansion

\[ \min(t, s) = \frac{2}{T} \sum_{n=0}^{\infty} \frac{T^2}{\pi^2(n + \frac{1}{2})^2} \left( \sin\left(n + \frac{1}{2}\right) \pi t \right) \left( \sin\left(n + \frac{1}{2}\right) \pi s \right), \]

where the convergence above is uniform on \([0, T]^2\).

\[ \blacksquare \]

Weiner-Kolmogorov Filtering
Figure 1: Self-similarity of the Wiener process