4

RANDOM PROCESSES

A completely satisfactory definition of random sequence is yet to be discovered.
—G. James and R. C. James, Mathematics Dictionary, Van Nostrand, Princeton NJ, 1959

4.1 CHAPTER FOCUS

Chapter 2 was about models for dynamic systems with manageable numbers of moving parts. These are models for *deterministic mechanics*, in which the state of every component of the system is represented and propagated explicitly.

Chapter 3 was about probability distributions and their statistical parameters, how the parameters evolve under transformations of the variates, and properties of the parameters that do not depend on details of the underlying probability distributions.

In this chapter, some of the basic notions and mathematical models of statistical and deterministic mechanics are combined into a *stochastic system model*, which represents the evolution over time of key statistical parameters in systems with uncertain dynamics.

These stochastic system models are used to define random processes (RPs) in continuous time and in discrete time (also called *random sequences*). They represent the *state of knowledge* about a dynamic system—including its state of uncertainty. They represent *what we know* about a dynamic system, including a quantitative model for what we do not know.

Kalman Filtering: Theory and Practice Using MATLAB®, Fourth Edition.
Mohinder S. Grewal and Angus P. Andrews.

In the next chapter, methods will be derived for modifying the state of knowledge based on noisy measurements (i.e., sensor outputs) related to the state of the dynamic system.

4.1.1 Main Points to be Covered

The theory of RPs and stochastic systems represents the evolution over time of the uncertainty of our knowledge about physical systems. This representation includes the effects of any *measurements* (or *observations*) that we make of the physical process and the effects of uncertainties about the measurement processes and dynamic processes involved. The uncertainties in the measurement and dynamic processes are modeled by RPs and stochastic systems.

Properties of uncertain dynamic systems are characterized by statistical parameters such as *means, correlations,* and *covariances.* By using only these numerical parameters, one can obtain finite representations of some probability distributions, which is important for implementing the solution on digital computers. This representation depends upon statistical properties such as orthogonality, stationarity, ergodicity, and Markovianness of the RPs involved and the Gaussianity of probability distributions. Gaussian, Markov, and uncorrelated (white-noise) processes will be used extensively in the following chapters. The autocorrelation functions and power spectral densities (PSDs) of such processes are also used. These are important in the development of frequency-domain and time-domain models. The time-domain models may be either continuous or discrete.

Shaping filters (continuous and discrete) are developed as models for many applications encountered in practice. These include random constants, random walks and ramps, sinusoidally correlated processes, and exponentially correlated processes. We derive the linear covariance equations for continuous and discrete systems to be used in Chapter 5. The *orthogonality principle* is developed and explained with scalar examples. This principle will be used in Chapter 5 to derive the Kalman filter equations.

4.1.2 Topics Not Covered

The stochastic calculus for dynamic systems with white process noise is not defined, although its results are used. The interested reader is referred to books on the mathematics of stochastic differential equations (e.g., those by Allen [1], Arnold [2], Baras and Mirelli [3], Itô and McKean [4], Oksendal [5], Sobczyk [6], or Stratonovich [7]).

4.2 RANDOM VARIABLES, PROCESSES, AND SEQUENCES

4.2.1 Historical Background

4.2.1.1 Random Processes We have mentioned in Chapter 1 the early history of analysis and modeling of unpredictable events in gambling. Financial

markets—which some prefer to ordinary gambling—came under similar analysis in the nineteenth century, when Carl Friedrich Gauss (1777–1855) did quite well in managing his own investments as well as those for widows of professors at the University of Göttingen. Danish astronomer-turned-actuary Thorvald Nicolai Thiele (1838–1910) did some seminal work on modeling of RPs and sequences, and French mathematician Louis Bachelier (1870–1946) developed models for prices on the Paris Bourse, a stock exchange. These developments in stochastic economics received scant attention among other scientists and engineers until quite recently, but the underappreciated synergy between mathematical economics and mathematical engineering has continued to this day. There is no Nobel Prize in either mathematics or engineering, but ever since the Bank of Sweden established the Nobel Prize in Economics in 1969 many of its recipients have been mathematical economists.

Mathematicians and mathematical physicists have also been interested in modeling RPs in nature.

An early impetus for the development of a mathematical theory of stochastic systems was the 1828 publication of "A brief account of microscopical observations made on the particles contained in the pollen of plants and on the general existence of active molecules in organic and inorganic bodies" by the British botanist Robert Brown. In it, Brown described a phenomenon he had observed while studying pollen grains of the herb *Clarkia pulchella* suspended in water, and similar observations by earlier investigators. The particles appeared to move about erratically, as though propelled by some unknown force. This phenomenon came to be called *Brownian movement* or *Brownian motion*. It has been studied extensively—both empirically and theoretically—by many eminent scientists (including Albert Einstein [8]) for much of the twentieth century. Empirical studies demonstrated that no biological forces were involved and eventually established that individual collisions with molecules of the surrounding fluid were causing the motion observed. The empirical results quantified how some statistical properties of the random motion were influenced by such physical properties as the size and mass of the particles and the temperature and viscosity of the surrounding fluid.

Mathematical models for RPs were derived in terms of what has come to be called *stochastic differential equations*. French scientist Paul Langevin[1] (1872–1946) modeled the velocity v of a particle in Brownian motion in terms of a differential equation [9] of the form

$$\frac{dv}{dt} = -\beta v + a(t), \tag{4.1}$$

where v is the velocity of a particle, β is a damping coefficient (due to the viscosity of the suspending medium), and $a(t)$ is called a *random force*. Equation 4.1 is now called the *Langevin equation.*

What Brown saw was a particle zig-zagging under the influence of velocity changes imparted by collisions with water molecules in an aqueous solution. The accelerations applied could be ascribed to van der Waals forces, which are sufficiently smooth that the velocity changes would not be instantaneous.

[1]Langevin is also famous for pioneering the development of sonar in World War I.

However, the random forcing function $a(t)$ of the Langevin equation (Equation 4.1) has been idealized in three ways from the physically motivated example of Brownian motion:

1. The velocity changes imparted to the particle have been assumed to be statistically independent from one collision to another.
2. The mean velocity change from independent collisions is assumed to be zero.
3. The effective time between collisions has been allowed to shrink to zero, with the magnitude of the imparted velocity change shrinking accordingly.

This new process model for $a(t)$ transcends the ordinary (Riemann) calculus, because the resulting "white-noise" process $a(t)$ is not integrable in the ordinary calculus of Georg Friedrich Bernhard Riemann (1826–1866). At about the time that the Kalman filter was introduced, however, a special calculus was developed by Kiyosi Itô (called the *Itô calculus* or the *stochastic calculus*) for this model. Equivalent derivations were done by Russian mathematician Ruslan L. Stratonovich (1930–1997) [7] and others, and the stochastic calculus is now commonly used in modeling of RPs.

Another mathematical characterization of white noise was provided by Norbert Weiner, using his generalized harmonic analysis. Wiener preferred to focus on the mathematical properties of $v(t)$ in Equation 4.1 with $\beta = 0$, a process now called a *Wiener process.*

4.2.2 Definitions

A *random variable* (RV) X was defined in Chapter 3 in terms of a probability measure. For the purposes of this book, the values of RVs will almost always be n-dimensional real vectors.

An RP assigns a RV $X(t)$ to every time t on some interval. (When a variate appears in differential equation, the lowercase letter might be substituted, however.)

A **random sequence** (RS) assigns a RV X_k to every integer k in some range of integers. A random sequence may also be denoted by curly brackets as $\{x_k\}$, in which case the lowercase letter might be used.

In Kalman filtering, the statistical properties of interest for RPs and sequences include how their statistics and joint statistics are related across time or between components of vectors.

4.3 STATISTICAL PROPERTIES

4.3.1 Independent Identically Distributed (i.i.d.) Processes

For *identically distributed* (i.d.) processes and sequences, the RV distribution is identical for all values of time (t) or index (k). Figure 4.1(a) shows an example of

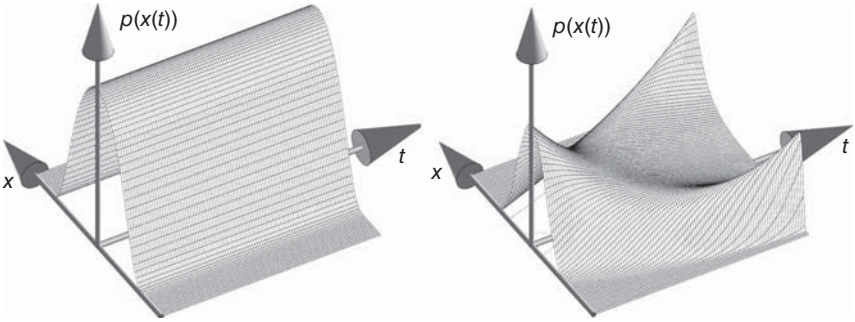


Figure 4.1 Probability densities of (a) identically distributed (i.d.) and (b) non-identically distributed processes.

a one-dimensional i.d. process, and Figure 4.1(b) shows an example of a non-i.d. process for which the shape of the probability density function evolves over time.

A process $X(t)$ is considered *time independent* if for any choice of distinct times $t_1, t_2, \ldots t_n$, the RVs $X(t_1), X(t_2), \ldots, X(t_n)$ are independent RVs. That is, their joint probability density is equal to the product of their individual probability densities:

$$p[x(t_1),\ x(t_2),\ \ldots,\ x(t_n)] = \prod_{i=1}^{n} p[x(t_i)]. \tag{4.2}$$

A similar definition for an independent identically distributed (i.i.d.) RS has k_i in place of t_i.

An *i.i.d. RP* is both i.d. and *independent*. Discrete and continuous i.i.d. processes play a significant role in Kalman filtering.

4.3.2 Process Means

For non-i.d. processes and sequences, the *mean* μ_x of an RP $X(t)$ or random sequence $\{X_k\}$ may be a function of time. For n-vector-valued RPs or RSs, the means are defined by the individual expected value taken at each time (continuous or discrete):

$$\mu_x(t) \stackrel{\text{def}}{=} \mathop{\mathrm{E}}_{x(t)\in X(t)} \langle x(t) \rangle \tag{4.3}$$

$$\mu_x(k) \stackrel{\text{def}}{=} \mathop{\mathrm{E}}_{x(k)\in X(k)} \langle x(k) \rangle, \tag{4.4}$$

respectively.

If the associated probability distributions $X(t)$ or $\{X_k\}$ can be defined by integrable probability density functions $p(\cdot)$, then the means can be defined in terms of

probability integrals as

$$\mu_x(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} x(t) p\ (x(t))\ dx(t), \tag{4.5}$$

$$\mu_x(k) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} x(k) p\ (x(k))\ dx(k), \tag{4.6}$$

which can be defined element-wise as

$$\mathop{\mathrm{E}}_{x(t)\in X(t)} \langle x_i(t)\rangle = \int_{-\infty}^{\infty} x_i(t)\ p(x_i(t))\ dx_i(t) \quad i = 1,\ 2,\ \ldots,n \tag{4.7}$$

$$\mathop{\mathrm{E}}_{x(k)\in -} \langle x_i(k)\rangle = \int_{-\infty}^{\infty} x_i(k)\ p(x_i(k))\ dx_i(k) \quad i = 1,\ 2,\ \ldots,n. \tag{4.8}$$

4.3.2.1 Zero-Mean Processes in Kalman Filtering Dynamic models for Kalman filtering separate zero-mean inputs from other inputs—usually labeled as (known) "control inputs" or (unknown) "slow variables." Any nonzero-mean RP would then be separated into its mean component and its (unknown) zero-mean component. If the mean of an RP were to change unpredictably and slowly over time, then its mean would be modeled as a separate "slow variable," not as part of a short-term RP. Models for such slow variables are an integral part of Kalman filtering.

Therefore, the fundamental RP models used in Kalman filtering will be *zero-mean process models*. Models for all other slowly varying parameters or variables can be built up from zero-mean RP models—which are usually i.i.d. processes, as well.

4.3.3 Time Correlation and Covariance

The *time correlation* of the n-vector-valued process $X(t)$ between any two times t_1 and t_2 is defined as the $n \times n$ matrix

$$\mathop{\mathrm{E}}_{\substack{x(t_1)\in X(t_1)\\ x(t_2)\in X(t_2)}} \langle x(t_1)x^T(t_2)\rangle = \mathop{\mathrm{E}}_{\substack{x(t_1)\in X(t_1)\\ x(t_2)\in X(t_2)}} \left\langle \begin{bmatrix} x_1(t_1)x_1(t_2) & \cdots & x_1(t_1)x_n(t_2) \\ \vdots & \ddots & \vdots \\ x_n(t_1)x_1(t_2) & \cdots & x_n(t_1)x_n(t_2) \end{bmatrix} \right\rangle \tag{4.9}$$

$$= \begin{bmatrix} \mathrm{E}\,\langle x_1(t_1)x_1(t_2)\rangle & \cdots & \mathrm{E}\,\langle x_1(t_1)x_n(t_2)\rangle \\ \vdots & \ddots & \vdots \\ \mathrm{E}\,\langle x_n(t_1)x_1(t_2)\rangle & \cdots & \mathrm{E}\,\langle x_n(t_1)x_n(t_2)\rangle \end{bmatrix}, \tag{4.10}$$

where, if the distributions can be defined in terms of probability density functions,

$$\mathrm{E}\,\langle x_i(t_1)x_j(t_2)\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i(t_1)x_j(t_2)p[x_i(t_1), x_j(t_2)]dx_i(t_1)dx_j(t_2). \tag{4.11}$$

4.3.3.1 Covariance Across Time If the RP involved is time independent but not zero-mean, time correlation can still be nonzero—even for an i.i.d. process. That is, if

$$p[x(t_1),\ x(t_2)] = p[x(t_1)] \times p[x(t_2)], \tag{4.12}$$

then for any i,j such that $1 \le i \le n$ and $1 \le j \le n$ and $t_1 \ne t_2$, the cross-correlation in time

$$\mathrm{E}\,\langle x_i(t_1)x_j(t_2)\rangle = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x_i(t_1)x_j(t_2)p\,[x_i(t_1),\ x_j(t_2)]\,dx_i(t_1)dx_j(t_2) \tag{4.13}$$

$$= \left\{\int_{-\infty}^{+\infty} x_i(t_1)p\left[x_i\left(t_1\right)\right]dx_i(t_1)\right\}$$

$$\times\left\{\int_{-\infty}^{+\infty} x_j(t_2)p\left[x_j\left(t_2\right)\right]dx_j(t_2)\right\} \tag{4.14}$$

$$= \underset{x(t_1)\in X(t_1)}{\mathrm{E}}\langle x_i\rangle \times \underset{x(t_2)\in X(t_2)}{\mathrm{E}}\langle x_j\rangle, \tag{4.15}$$

which is the product of the means.

This issue can be avoided by defining the *covariance* across time by subtracting the means before taking expected values:

$$\underset{\substack{x(t_1)\in X(t_1)\\ x(t_2)\in X(t_2)}}{\mathrm{E}}\langle[x(t_1)-\mathrm{E}x(t_1)][x(t_2)-\mathrm{E}x(t_2)]^{\mathrm{T}}\rangle. \tag{4.16}$$

When the process $X(t)$ has zero mean (i.e., $\mathrm{E}x(t) = 0$ for all t), its correlation and covariance are equal.

4.3.3.2 Cross-time-correlation and Covariance between RPs The time-cross-correlation matrix of two RPs $X(t)$, an n-vector, and $Y(t)$, an m-vector, is given by an $n \times m$ matrix

$$\mathrm{E}\langle x(t_1)y^{\mathrm{T}}(t_2)\rangle, \tag{4.17}$$

where

$$\mathrm{E}x_i(t_1)y_j(t_2) = \int_{-\infty}^{\infty}\int x_i(t_1)y_j(t_2)p[x_i(t_1),y_j(t_2)]dx_i(t_1)dy_j(t_2). \tag{4.18}$$

Similarly, the cross-covariance $n \times m$ matrix is

$$\mathrm{E}\langle[x(t_1)-\mathrm{E}x(t_1)][y(t_2)-\mathrm{E}y(t_2)]^{\mathrm{T}}\rangle. \tag{4.19}$$

4.3.4 Uncorrelated and Orthogonal Random Processes

4.3.4.1 Uncorrelated Random Processes A *random process* $X(t)$ is called *uncorrelated* if its *time covariance*

$$\mathrm{E}\,\langle [x(t_1) - \mathrm{E}\,\langle x(t_1)\rangle][x(t_2) - \mathrm{E}\,\langle x(t_2)\rangle]^{\mathrm{T}}\rangle = Q(t_1, t_2)\delta(t_1 - t_2), \tag{4.20}$$

where $\delta(t)$ is the Dirac delta "function," defined by

$$\int_a^b \delta(t)\, dt = \begin{cases} 1 & \text{if } a \le 0 \le b, \\ 0 & \text{otherwise.} \end{cases} \tag{4.21}$$

Similarly, a *random sequence* x_k is called *uncorrelated* if

$$\mathrm{E}\,\langle [x_k - \mathrm{E}\,\langle x_k\rangle][x_j - \mathrm{E}\,\langle x_j\rangle]^{\mathrm{T}}\rangle = Q(k, j)\Delta(k - j), \tag{4.22}$$

where $\Delta(\cdot)$ is the Kronecker delta function,[2] defined by

$$\Delta(k) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4.23}$$

Two RPs $X(t)$ and $Y(t)$ are called *uncorrelated* if their cross-covariance matrix is identically zero for all t_1 and t_2:

$$\mathrm{E}\,\langle [x(t_1) - \mathrm{E}\,\langle x(t_1)\rangle][y(t_2) - \mathrm{E}\,\langle y(t_2)\rangle]^{\mathrm{T}}\rangle = 0. \tag{4.24}$$

White Noise. A *white-noise* process or sequence is an example of an uncorrelated process or sequence.

4.3.4.2 Orthogonal Random Processes The processes $X(t)$ and $Y(t)$ are called *orthogonal* if their *correlation matrix* is identically zero:

$$\mathrm{E}\,\langle x(t_1)y^{\mathrm{T}}(t_2)\rangle = 0. \tag{4.25}$$

A process $X(t)$ is considered independent if for any choice of distinct times $t_1, t_2, \ldots, t_n$, the RVs $x(t_1), x(t_2), \ldots, x(t_n)$ are independent. That is,

$$p[x(t_1), \ldots, x(t_n)] = \prod_{i=1}^{n} p[x(t_i)]. \tag{4.26}$$

Independence (all of the moments) also implies no correlation (which restricts attention to the second moments), but the opposite implication is not true, except in such special cases as Gaussian processes. Note that *whiteness* means *uncorrelated* in time rather than *independent* in time (i.e., including all moments), although this distinction disappears for the important case of white Gaussian processes (see Chapter 5).

[2]Named after the German mathematician Leopold Kronecker (1823–1891).

4.3.5 Strict-Sense and Wide-Sense Stationarity

The RP $X(t)$ (or random sequence x_k) is called *strict-sense stationary* if all its statistics (meaning $p[x(t_1), x(t_2), \ldots]$) are invariant with respect to shifts of the time origin:

$$p(x_1, x_2, \ldots, x_n, t_1, \ldots, t_n) = p(x_1, x_2, \ldots, x_n, t_1 + \varepsilon, t_2 + \varepsilon, \ldots, t_n + \varepsilon). \tag{4.27}$$

The RP $X(t)$ (or x_k) is called *wide-sense stationary* (WSS) (or "weak-sense" stationary) if

$$\mathrm{E}\,\langle x(t)\rangle = c \text{ (a constant)} \tag{4.28}$$

and

$$\mathrm{E}\,\langle x(t_1)x^{\mathrm{T}}(t_2)\rangle = Q(t_2 - t_1) = Q(\tau), \tag{4.29}$$

where Q is a matrix with each element depending only on the difference $t_2 - t_1 = \tau$. Therefore, when $X(t)$ is stationary in the weak sense, it implies that its first- and second-order statistics are independent of time origin, while strict stationarity by definition implies that statistics of all orders are independent of the time origin.

4.3.6 Ergodic Random Processes

4.3.6.1 Historical Note The term *ergodic* came originally from the development of statistical mechanics for thermodynamic systems. It is taken from the Greek words for *energy* and *path*. The term was applied by the American physicist Josiah Willard Gibbs (1839–1903) to the time history (or path) of the state of a thermodynamic system of constant energy. Gibbs had assumed that a thermodynamic system would eventually take on all possible states consistent with its energy. It was shown to be impossible from function-theoretic considerations in the nineteenth century. The so-called ergodic hypothesis of James Clerk Maxwell (1831–1879) is that the temporal means of a stochastic system are equivalent to the ensemble means. The concept was given firmer mathematical foundations by George David Birkhoff and John von Neumann around 1930 and by Norbert Wiener in the 1940s.

Following Maxwell's hypothesis, an RP is considered **ergodic** if all of its statistical parameters (mean, variance, and so on) can be determined from arbitrarily chosen member functions. A sampled function $X(t)$ is ergodic if its time-averaged statistics equal its ensemble averages.

4.3.7 Markov Processes and Sequences

An RP $X(t)$ is called a *Markov process*[3] if its future state distribution, conditioned on knowledge of its present state, is not improved by knowledge of previous

[3]Named after the Russian mathematician Andrei Andreyevich Markov (1856–1922), who first developed many of the concepts and the related theory.

states:

$$p[x(t_i)|x(\tau);\tau < t_{i-1}] = p[x(t_i)|x(t_{i-1})], \tag{4.30}$$

where the times $t_1 < t_2 < t_3 < \cdots < t_i$.

Similarly, an RS x_k is called a *Markov sequence* if

$$p[x_i|x_k; k \leq i-1] = p[x_i|x_{i-1}]. \tag{4.31}$$

The solution to a general first-order differential or difference equation with an independent process (uncorrelated normal RP) as a forcing function is a Markov process. That is, if $x(t)$ and x_k are n-vectors satisfying

$$\dot{x}(t) = F(t)x(t) + G(t)w(t) \tag{4.32}$$

or

$$x_k = \Phi_{k-1}x_{k-1} + G_{k-1}w_{k-1}, \tag{4.33}$$

where $w(t)$ and w_{k-1} are realizations of r-dimensional independent RPs and RSs, the solutions $X(t)$ and x_k are then vector Markov processes and sequences, respectively.

4.3.8 Gaussian Random Processes

An n-dimensional RP $X(t)$ is called *Gaussian* (or normal) if its probability density function is Gaussian, as given by the formulas of Example 3.1, with covariance matrix

$$P = \mathrm{E}\left\langle |x(t) - \mathrm{E}\left\langle x(t)\right\rangle [x(t) - \mathrm{E}\left\langle x(t)\right\rangle]^{\mathrm{T}}\right\rangle \tag{4.34}$$

for the RV $x \in X$.

Gaussian RPs have some useful properties:

1. A Gaussian RP $X(t)$ is WSS—and stationary in the strict sense.
2. Orthogonal Gaussian RPs are independent.
3. Any linear function of jointly Gaussian RP results in another Gaussian RP.
4. All statistics of a Gaussian RP are completely determined by its first- and second-order statistics.

4.3.9 Simulating Multivariate Gaussian Processes

By using the Cholesky decomposition algorithm, Gaussian n-vector RPs with specified means and covariances can be simulated using scalar Gaussian pseudorandom number generators, such as randn in MATLAB software©.

Cholesky decomposition methods are discussed in Chapters 7 and 8, and in Appendix B (on the Wiley website). We show here how these methods can be used

to generate uncorrelated pseudorandom vector sequences with zero mean (or any specified mean) and a specified covariance Q.

There are many programs that will generate pseudorandom sequences of uncorrelated Gaussian scalars $\{s_i \mid i = 1, 2, 3, \dots\}$ with zero mean and unit variance:

$$s_i \in \mathcal{N}(0, 1) \text{ for all } i, \tag{4.35}$$

$$\mathrm{E}\,\langle s_i s_j \rangle = \begin{cases} 0 & \text{if} \quad i \neq j, \\ 1 & \text{if} \quad i = j. \end{cases} \tag{4.36}$$

These can be used to generate sequences of Gaussian n-vectors x_k with mean zero and covariance I_m:

$$u_k = [s_{nk+1},\ s_{nk+2},\ s_{nk+3},\ \dots,\ s_{n(k+1)}]^{\mathrm{T}}, \tag{4.37}$$

$$\mathrm{E}\,\langle u_k \rangle = 0, \tag{4.38}$$

$$\mathrm{E}\,\langle u_k u_k^{\mathrm{T}} \rangle = I_n. \tag{4.39}$$

These vectors, in turn, can be used to generate a sequence of n-vectors w_k with zero mean and covariance Q. For that purpose, let

$$CC^{\mathrm{T}} = Q \tag{4.40}$$

be the Cholesky decomposition of Q, and let the sequence of n-vectors w_k be generated according to the rule

$$w_k = Cu_k. \tag{4.41}$$

Then the sequence of vectors $\{w_0, w_1, w_2, \dots\}$ will have mean

$$\mathrm{E}\,\langle w_k \rangle = C\,\mathrm{E}\,\langle u_k \rangle \tag{4.42}$$

$$= 0 \tag{4.43}$$

(an n-vector of zeros) and covariance

$$\mathrm{E}\,\langle w_k w_k^{\mathrm{T}} \rangle = \mathrm{E}\,\langle Cu_k (Cu_k)^{\mathrm{T}} \rangle \tag{4.44}$$

$$= CI_n C^{\mathrm{T}} \tag{4.45}$$

$$= Q. \tag{4.46}$$

The same technique can be used to obtain pseudorandom Gaussian vectors with a given mean v by adding v to each w_k. These techniques are used in simulation and Monte Carlo analysis of stochastic systems.

4.3.10 Power Spectral Densities

Let $X(t)$ be a zero-mean scalar stationary RP with autocorrelation function

$$\psi_x(\tau) = \underset{t}{\mathrm{E}}\langle x(t)x(t+\tau)\rangle. \tag{4.47}$$

The PSD is defined as the Fourier transform $\Psi_x(\omega)$ of $\psi_x(\tau)$,

$$\Psi_x(\omega) = \int_{-\infty}^{\infty} \psi_x(\tau) e^{-j\omega\tau}\, d\tau, \tag{4.48}$$

where the inverse transform as

$$\psi_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi_x(\omega) e^{j\omega\tau}\, d\omega. \tag{4.49}$$

The following are useful properties of autocorrelation functions:

1. Autocorrelation functions are symmetrical ("even" functions).
2. An autocorrelation function attains its maximum value at the origin.
3. Its Fourier transform is nonnegative (greater than or equal to zero).

These properties are satisfied by all valid autocorrelation functions.

Setting $\tau = 0$ in Equation 4.49 gives

$$\underset{t}{\mathrm{E}}\langle x^2(t)\rangle = \psi_x(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi_x(\omega)\, d\omega. \tag{4.50}$$

Because of property 1 of the autocorrelation function,

$$\Psi_x(\omega) = \Psi_x(-\omega); \tag{4.51}$$

that is, the PSD is a symmetric function of frequency.

Example 4.1 (PSD for an Exponentially Correlated Process) The exponentially correlated process diagrammed in Figure 4.2(a) has an autocovariance function of the general form

$$\psi_x(t) = \sigma^2 e^{-|t|/\tau},$$

where σ^2 is the mean-squared (MS) process amplitude and τ is the autocorrelation time constant. Its PSD is the Fourier transform of its autocovariance function,

$$\begin{aligned}\Psi_x(\omega) &= \int_{-\infty}^{0} \sigma^2 e^{t/\tau} e^{-j\omega t}\, dt + \int_{0}^{\infty} \sigma^2 e^{-t/\tau} e^{-j\omega t}\, dt \\ &= \sigma^2 \left(\frac{1}{1/\tau - j\omega} + \frac{1}{1/\tau + j\omega} \right) = \frac{2\sigma^2\tau}{1/\tau^2 + \omega^2},\end{aligned}$$

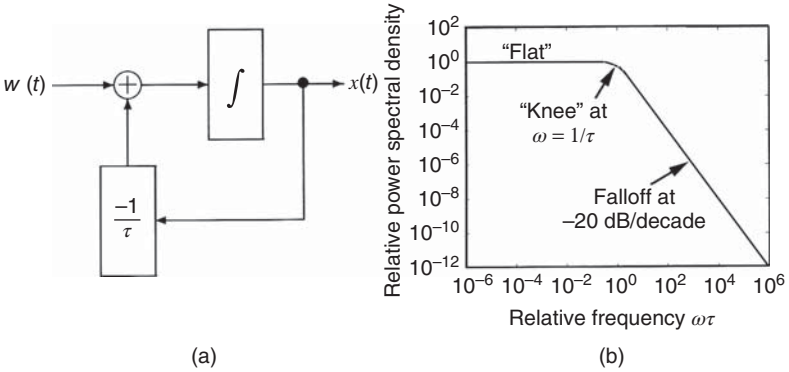


Figure 4.2 Exponentially correlated random process. (a) Block diagram and (b) PSD.

the shape of which is illustrated in Figure 4.2(b) as a log–log plot. White noise passed through a low pass resistor–capacitor (RC) filter with resistance R (in ohms) and capacitance C (in farads) would have a filter output spectrum of this shape, with the "knee" at $\omega = 1/\tau = 1/(R \times C)$.

Example 4.2 (Underdamped Harmonic Oscillator as a Shaping Filter) This is an example of a second-order Markov process generated by passing WSS white noise with zero mean and unit variance through a second-order "shaping filter" with the dynamic model of a harmonic resonator. The dynamic model for a damped harmonic oscillator was introduced in Examples 2.2, 2.3, 2.6, and 2.7 and will be used again in Chapters 5 and 7.

The transfer function of the dynamic system is

$$H(s) = \frac{as + b}{s^2 + 2\zeta w_n s + w_n^2}.$$

Definitions of ζ, w_n, and s are the same as in Example 2.7. The state-space model of $H(s)$ is given as

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -w_n^2 & -2\zeta w_n \end{bmatrix} \begin{bmatrix} x_{1(t)} \\ x_2(t) \end{bmatrix} + \begin{bmatrix} a \\ b - 2a\zeta w_n \end{bmatrix} w(t),$$
$$z(t) = x_1(t) = x(t).$$

The general form of the autocorrelation is

$$\psi_x(\tau) = \frac{\sigma^2}{\cos\theta} e^{-\zeta w_n |\tau|} \cos\left(\sqrt{1-\zeta^2} \quad w_n|\tau| - \theta\right).$$

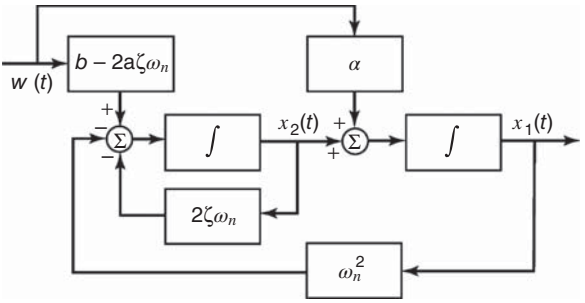


Figure 4.3 Diagram of a second-order Markov process

In practice, σ^2, θ, ζ, and w_n are chosen to fit empirical data (see Problem 4.4). The PSD corresponding to the $\psi_x(\tau)$ will have the form

$$\Psi_x(\omega) = \frac{a^2\omega^2 + b^2}{\omega^4 + 2\ \omega_n(2\zeta^2 - 1)\ \omega^2 + \omega_n}.$$

(The peak of this PSD will not be at the "natural" (undamped) frequency ω_n, but at the "resonant" frequency defined in Example 2.6.)

The block diagram corresponding to the state-space model is shown in Figure 4.3.

The *mean power* of a scalar RP is given by the equations

$$\mathop{\mathrm{E}}_{t}\langle x^2(t)\rangle = \lim_{T\to\infty}\int_{-T}^{\mathrm{T}} x^2(t)\ dt \tag{4.52}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\Psi_x(\omega)\ d\omega \tag{4.53}$$

$$= \sigma^2. \tag{4.54}$$

The *cross power spectral density* between an RP $X(t)$ and an RP $Y(t)$ is given by the formula

$$\Psi_{xy}(\omega) = \int_{-\infty}^{\infty}\psi_{xy}(\tau)e^{-j\omega\tau}\ d\tau. \tag{4.55}$$

4.4 LINEAR RANDOM PROCESS MODELS

Linear system models of the type illustrated in Figure 4.4 are defined by the equation

$$y(t) = \int_{-\infty}^{\infty} x(\tau)h(t,\tau)\ d\tau, \tag{4.56}$$

where $X(t)$ is input and $h(t, \tau)$ is the linear system *weighting function* (see Figure 4.4). If the system is time invariant (i.e., h does not depend on t), then Equation 4.56 can

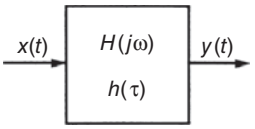


Figure 4.4 Block diagram representation of a linear system.

be written as

$$y(t) = \int_0^\infty h(\tau)x(t-\tau)\,d\tau. \tag{4.57}$$

This type of integral is called a *convolution integral,* and $h(\tau)$ is called its *kernel function.* Manipulation of Equation 4.57 leads to relationships between autocorrelation functions of $X(t)$ and $Y(t)$,

$$\psi_y(\tau) = \int_0^\infty d\tau_1 h(\tau_1) \int_0^\infty d\tau_2 h(\tau_2)\psi_x(\tau + \tau_1 - \tau_2)\,\text{(autocorrlation,)} \tag{4.58}$$

$$\psi_{xy}(\tau) = \int_0^\infty h(\tau_1)\psi_x(\tau - \tau_1)\,d\tau_1\,\text{(cross-correlation,)} \tag{4.59}$$

and spectrum relationships

$$\Psi_{xy}(\omega) = H(j\omega)\Psi_x(\omega)\,\text{(cross-spectrum)}, \tag{4.60}$$

$$\Psi_y(\omega) = |H(j\omega)|^2\Psi_x(\omega)\,\text{(PSD)}, \tag{4.61}$$

where H is the *system transfer function* (also shown in Figure 4.4), defined by the Laplace transform of $h(\tau)$ as

$$\int_0^\infty h(\tau)e^{s\tau}\,d\tau = H(s) = H(j\omega), \tag{4.62}$$

where $s = j\omega$ and $j = \sqrt{-1}$.

4.4.1 Stochastic Differential Equations for RPs

4.4.1.1 The Calculus of Stochastic Differential Equations Differential equations involving RPs are called *stochastic differential equations*. Introducing RPs as inhomogeneous terms in ordinary differential equations has ramifications beyond the level of rigor that will be followed here, and the reader should be aware of them. The problem is that RPs are not integrable functions in the conventional (Riemann) calculus. The resolution of this problem requires foundational modifications of the calculus to obtain many of the results presented. The Riemann integral of the "ordinary" calculus must be modified to what is called the *Itô calculus*. The interested reader will find these issues treated more rigorously in the books by Bucy and Joseph [10] and Itô [11].

4.4.1.2 Linear Stochastic Differential Equations A linear stochastic differential equation as a model of an RP with initial conditions has the general form

$$\dot{x}(t) = F(t)x(t) + G(t)w(t) + C(t)u(t), \tag{4.63}$$

$$z(t) = H(t)x(t) + v(t) + D(t)u(t), \tag{4.64}$$

where the variables are defined as

$x(t)$ is an $n \times 1$ state vector,
$z(t)$ is an $l \times 1$ measurement vector,
$u(t)$ is an $r \times 1$ deterministic input vector,
$F(t)$ is an $n \times n$ time-varying dynamic coefficient matrix,
$C(t)$ is an $n \times r$ time-varying input coupling matrix,
$H(t)$ is an $l \times n$ time-varying measurement sensitivity matrix,
$D(t)$ is an $l \times r$ time-varying output coupling matrix,
$G(t)$ is an $n \times r$ time-varying process noise coupling matrix,
$w(t)$ is an $r \times 1$ zero-mean uncorrelated "plant noise" process,
$v(t)$ is an $l \times 1$ zero-mean uncorrelated "measurement noise" process,

and the expected values will be

$$\begin{aligned}
\mathrm{E}\,\langle w(t)\rangle &= 0,\\
\mathrm{E}\,\langle v(t)\rangle &= 0,\\
\mathrm{E}\,\langle w(t_1)w^{\mathrm{T}}(t_2)\rangle &= Q(t_1)\delta(t_2 - t_1),\\
\mathrm{E}\,\langle v(t_1)v^{\mathrm{T}}(t_2)\rangle &= R(t_1)\delta(t_2 - t_1).\\
\mathrm{E}\,\langle w(t_1)v^{\mathrm{T}}(t_2)\rangle &= M(t_1)\delta(t_2 - t_1)
\end{aligned}$$

$$\delta(t) = \begin{cases} 1, & t = 0,\\ 0, & t \neq 0. \end{cases} \qquad \text{(Dirac delta function)}$$

The symbols Q, R, and M represent $r \times r$, $l \times l$, and $r \times l$ matrices, respectively, and δ represents the Dirac delta "function" (a measure). The values over time of the variate $X(t)$ in the differential equation model define vector-valued Markov processes. This model is a fairly accurate and useful representation for many real-world processes, including stationary Gaussian and nonstationary Gaussian processes, depending on the statistical properties of the RVs and the temporal properties of the deterministic variables. (The function $u(t)$ usually represents a known control input. For the rest of the discussion in this chapter, we will assume that $u(t) = 0$.)

Example 4.3 (Exponentially Correlated Process) Continuing with Example 4.1, let the RP $X(t)$ be a zero-mean stationary normal RP having autocorrelation

$$\psi_x(\tau) = \sigma^2 e^{-\alpha|\tau|}. \tag{4.65}$$

The corresponding PSD is

$$\Psi_x(\omega) = \frac{2\sigma^2\alpha}{\omega^2 + \alpha^2}. \tag{4.66}$$

This type of RP can be modeled as the output of a linear system with input $w(t)$, a zero-mean white Gaussian noise with PSD equal to unity. Using Equation 4.61, one can derive the transfer function $H(j\omega)$ for the following model:

$$H(j\omega)H(-j\omega) = \frac{\sqrt{2\alpha}\sigma}{\alpha + j\omega} \cdot \frac{\sqrt{2\alpha}\sigma}{\alpha - j\omega}.$$

Take the stable portion of this system transfer function as

$$H(s) = \frac{\sqrt{2\alpha}\sigma}{s + \alpha}, \tag{4.67}$$

which can be represented as

$$\frac{x(s)}{w(s)} = \frac{\sqrt{2\alpha}\sigma}{s + \alpha}, \tag{4.68}$$

By taking the inverse Laplace transform of both sides of this last equation, one can obtain the following sequence of equations:

$$\begin{aligned} \dot{x}(t) + \alpha x(t) &= \sqrt{2\alpha}\sigma w(t), \\ \dot{x}(t) &= -\alpha x(t) + \sqrt{2\alpha}\sigma w(t), \\ z(t) &= x(t), \end{aligned}$$

with $\sigma_x^2(0) = \sigma^2$. The parameter $1/\alpha$ is called the *correlation time* of the process.

The block diagram representation of the process in Example 4.3 is shown in Table 4.1. This is called a *shaping filter*. Some other examples of differential equation models are also given in Table 4.1.

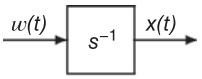
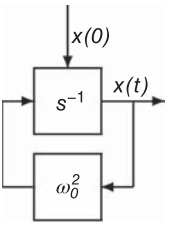
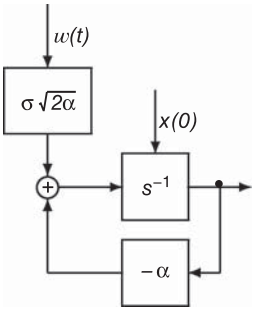
4.4.2 Discrete-Time Models for Random Sequences (RS)

RPs in discrete time are also called *random sequences*. A vector-valued discrete-time recursive equation for modeling an RS with initial conditions can be given in the form

$$\begin{aligned} x_k &= \Phi_{k-1}x_{k-1} + G_{k-1}w_{k-1} + \Gamma_{k-1}u_{k-1}, \\ z_k &= H_k x_k + \upsilon_k + D_k u_k. \end{aligned} \tag{4.69}$$

This is the complete model with deterministic inputs u_k as discussed in Chapter 2, Equations 2.38 and 2.39, and RS noise w_k and υ_k as described in Chapter 5,

TABLE 4.1 System Models of Random Processes

Random Process	Autocorrelation Function and Power Spectral Density	Shaping Filter	State-Space Model
White noise	$\psi_x(\tau) = \sigma^2\delta^2(\tau)$ $\Psi_x(\omega) = \sigma^2$	None	Plant noise $w(t)$
Random walk	$\psi_x(\tau) = \mathrm{E}\langle x(t_1)x(t_2)\rangle$ $= \sigma^2 t_1$ if $t_1 < t_2$ $\Psi_x(\omega) \propto \sigma^2/\omega^2$		$\dot{x} = w(t)$ $\sigma_x^2(0) = 0$
Random constant	$\psi_x(\tau) = \sigma^2$ $\Psi_x(\omega) = 2\pi\sigma^2\delta(\omega)$	None	$\dot{x} = 0$ $\sigma_x^2(0) = \sigma^2$
Sinusoid	$\psi_x(\tau) = \sigma^2\cos(\omega_0\tau)$ $\Psi_x(\omega) = \pi\sigma^2\delta(\omega - \omega_0)$		$\dot{x} = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{bmatrix} x$ $P(0) = \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$
Exponentially correlated	$\psi_x(\tau) = \sigma^2 e^{-\alpha\lvert\tau\rvert}$ $\Psi_x(\omega) = \frac{2\sigma^2\alpha}{\omega^2+\alpha^2}$ $1/\alpha$ = correlation time		$\dot{x} = -\alpha x + \sigma\sqrt{2\alpha}w(t)$ $\sigma_x^2(0) = \sigma^2$

where

x_k is an $n \times 1$ state vector,
z_k is an $\ell \times 1$ measurement vector,
u_k is an $r \times 1$ deterministic input vector,
Φ_{k-1} is an $n \times n$ time-varying matrix,
G_{k-1} is an $n \times r$ time-varying matrix,
H_k is an $\ell \times n$time-varying matrix,
D_k is an $\ell \times r$ time-varying matrix,
Γ_{k-1} is an $n \times r$ time-varying matrix,

and the expected values

$$\begin{aligned}
\mathrm{E}\,\langle w_k \rangle &= 0 \\
\mathrm{E}\,\langle v_k \rangle &= 0 \\
\mathrm{E}\,\langle w_{k_1} w_{k_2}^{\mathrm{T}} \rangle &= Q_{k_1} \Delta(k_2 - k_1) \\
\mathrm{E}\,\langle v_{k_1} v_{k_2}^{\mathrm{T}} \rangle &= R_{k_1} \Delta(k_2 - k_1) \\
\mathrm{E}\,\langle w_{k_1} v_{k_2}^{\mathrm{T}} \rangle &= M_{k_1} \Delta(k_2 - k_1).
\end{aligned}$$

Example 4.4 (Exponentially Correlated Sequence) Let the $\{x_k\}$ be a zero-mean stationary Gaussian RS with autocorrelation

$$\psi_x(k_2 - k_1) = \sigma^2 e^{-\alpha|k_2 - k_1|}.$$

This type of RS can be modeled as the output of a linear system with input w_k being zero-mean white Gaussian noise with PSD equal to unity.

A difference equation model for this type of process can be defined as

$$x_k = \Phi x_{k-1} + G w_{k-1}, \quad z_k = x_k. \tag{4.70}$$

In order to use this model, we need to solve for the unknown parameters Φ and G as functions of the parameter α. To do so, we first multiply Equation 4.19 by x_{k-1} on both sides and take the expected values to obtain the equations

$$\begin{aligned}
\mathrm{E}\,\langle x_k x_{k-1} \rangle &= \Phi\,\mathrm{E}\,\langle x_{k-1} x_{k-1} \rangle + G\,\mathrm{E}\,\langle w_{k-1} x_{k-1} \rangle, \\
\sigma^2 e^{-\alpha} &= \Phi \sigma^2,
\end{aligned}$$

assuming the w_k are uncorrelated and $\mathrm{E}\,\langle w_K \rangle = 0$, so that $\mathrm{E}\,\langle w_{k-1} x_{k-1} \rangle = 0$. One obtains the solution

$$\Phi = e^{-\alpha}. \tag{4.71}$$

Next, square the state variable defined by Equation 4.70 and take its expected value:

$$\mathrm{E}\,\langle x_k^2 \rangle = \Phi^2 \mathrm{E}\,\langle x_{k-1} x_{k-1} \rangle + G^2 \mathrm{E}\,\langle w_{k-1} w_{k-1} \rangle, \tag{4.72}$$

$$\sigma^2 = \sigma^2 \Phi^2 + G^2, \tag{4.73}$$

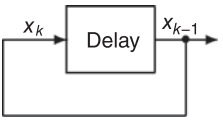
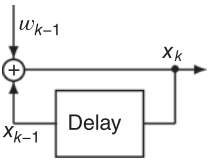
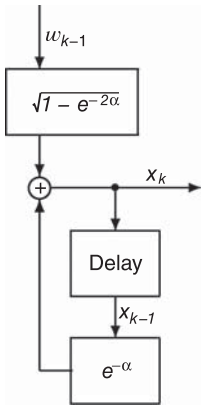
because the variance $\mathrm{E}\,\langle w_{k-1}^2 \rangle = 1$ and the parameter $G = \sigma\sqrt{1 - e^{-2\alpha}}$.

The complete model is then

$$x_k = e^{-\alpha} x_{k-1} + \sigma\sqrt{1 - e^{-2\alpha}}\, w_{k-1}$$

with $\mathrm{E}\,\langle w_k \rangle = 0$ and $\mathrm{E}\,\langle w_{k_1}\, w_{k_2} \rangle = \Delta(k_2 - k_1)$.

TABLE 4.2 Stochastic System Models for Discrete Random Sequences

Process Type	Autocorrelation Function	Block Diagram	State-Space Model
Random constant	$\psi_x(k_1 - k_2) = \sigma^2$		$x_k = x_{k-1}$ $\sigma_x^2(0) = \sigma^2$
Random walk	$\rightarrow +\infty$		$x_k = x_{k-1} + w_{k-1}$ $\sigma_x^2(0) = 0$
Exponentially correlated	$\psi_x(k_2 - k_1) = \sigma^2 e^{-\alpha\lvert k_2 - k_1\rvert}$		$x_k = e^{-\alpha} x_{k-1} + \sigma\sqrt{1 - e^{-2\alpha}}\, w_{k-1}$ $\sigma_x^2(0) = \sigma^2$

The dynamic process model derived in Example 4.4 is called a shaping filter. Block diagrams of this and other shaping filters are given in Table 4.2, along with their difference equation models.

4.4.3 Autoregressive Processes and Linear Predictive Models

A *linear predictive model* for a signal is a representation in the form

$$\acute{x}_{k+1} = \sum_{i=1}^{n} a_i \acute{x}_{k-i+1} + \acute{u}_k, \tag{4.74}$$

where $\acute{x}_k$ is the *prediction error*. Successive samples of the signal are predicted as linear combinations of the n previous values.

An *autoregressive process* has the same formulation, except that $\acute{u}_k$ is a white Gaussian noise sequence. Note that this formula for an autoregressive process can be rewritten in state-transition matrix (STM) form as

$$\begin{bmatrix} \acute{x}_{k+1} \\ \acute{x}_k \\ \acute{x}_{k-1} \\ \vdots \\ \acute{x}_{k-n+2} \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \acute{x}_k \\ \acute{x}_{k-1} \\ \acute{x}_{k-2} \\ \vdots \\ \acute{x}_{k-n+1} \end{bmatrix} + \begin{bmatrix} \acute{x} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{4.75}$$

$$x_{k+1} = \Phi x_k + u_k, \tag{4.76}$$

where the "state" is the n-vector of the last n samples of the signal and the covariance matrix Q_k of the associated process noise u_k will be filled with zeros, except for the term $Q_{11} = \mathrm{E}\,\langle \acute{u}_k^2 \rangle$.

4.5 SHAPING FILTERS (SF) AND STATE AUGMENTATION

Shaping filters have a long and interesting history. In the early years of "wireless telegraphy" (radio) transmission, a spark provided a white-noise source which could be input to an LC "tank circuit" (a shaping filter), the output of which was the radio frequency (RF) transmission frequency band of interest.

The focus of this section is on the use of shaping filters to model nonwhite-noise models for stationary RPs, using white-noise processes as inputs. For many physical systems encountered in practice, it may not be justified to assume that all noises are white Gaussian noise processes. It can be useful to generate an autocorrelation function or PSD from real data and then develop an appropriate noise model using differential or difference equations. These models are called *shaping filters,* a concept introduced in works by Hendrik Wade Bode (1905–1982) and Claude Elwood Shannon (1916–2001) [12] and by Lotfi Asker Zadeh and John Ralph Ragazzini (1912–1988) [13]. They are filters driven by noise with a flat spectrum (white-noise processes), which they shape to represent the spectrum of the actual system. It was shown in the previous section that a linear time-invariant system (shaping filter) driven by WSS white Gaussian noise provides such a model. The state vector can be "augmented" by appending to it the state vector components of the shaping filter, with the resulting model having the form of a linear dynamic system driven by white noise.

4.5.1 Correlated Process Noise Models

4.5.1.1 Shaping Filters for Dynamic Disturbance Noise Let a system model be given by

$$\dot{x}(t) = F(t)x(t) + G(t)w_1(t), \quad z(t) = H(t)x(t) + v(t), \tag{4.77}$$

where $w_1(t)$ is nonwhite, for example, correlated Gaussian noise. As given in the previous section, *v(t)* is a zero-mean white Gaussian noise. Suppose that $w_1(t)$ can be modeled by a linear shaping filter:[4]

$$\dot{x}_{\mathrm{SF}}(t) = F_{\mathrm{SF}}(t)x_{\mathrm{SF}}(t) + G_{\mathrm{SF}}(t)w_2(t) \tag{4.78}$$

$$w_1(t) = H_{\mathrm{SF}}(t)x_{\mathrm{SF}}(t), \tag{4.79}$$

where SF denotes the shaping filter and $w_2(t)$ is zero-mean white Gaussian noise. Now define a new augmented state vector

$$X(t) = [x(t)x_{\mathrm{SF}}(t)]^{\mathrm{T}}. \tag{4.80}$$

Equations 4.77 and 4.79 can be combined into the matrix form

$$\begin{bmatrix} \dot{x}(t) \\ \dot{x}_{\mathrm{SF}}(t) \end{bmatrix} = \begin{bmatrix} F(t) & G(t)H_{\mathrm{SF}}(t) \\ 0 & F_{\mathrm{SF}}(t) \end{bmatrix} \begin{bmatrix} x(t) \\ x_{\mathrm{SF}}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ G_{\mathrm{SF}}(t) \end{bmatrix} w_2(t), \tag{4.81}$$

$$\dot{X}(t) = F_T(t)X(t) + G_T(t)w_2(t), \tag{4.82}$$

and the output equation can be expressed in compatible format as

$$z(t) = [H(t)\ 0] \begin{bmatrix} x(t) \\ x_{\mathrm{SF}}(t) \end{bmatrix} + v(t) \tag{4.83}$$

$$= H_T(t)X(t) + v(t). \tag{4.84}$$

This total system given by Equations 4.82 and 4.84 is a linear differential equation model driven by white Gaussian noise. (See Figure 4.5 for a non-white noise model.)

4.5.2 Correlated Measurement Noise Models

4.5.2.1 Shaping Filters for Measurement Noise A similar development is feasible for the case of time-correlated measurement noise $v_1(t)$:

$$\dot{x}(t) = F(t)x(t) + G(t)w(t),$$

$$z(t) = H(t)x(t) + v_1(t). \tag{4.85}$$

In this case, let $v_2(t)$ be zero-mean white Gaussian noise and let the measurement noise $v_1(t)$ be modeled by

$$\dot{x}_{\mathrm{SF}}(t) = F_{\mathrm{SF}}(t)x_{\mathrm{SF}}(t) + G_{\mathrm{SF}}(t)v_2(t),$$

$$v_1(t) = H_{\mathrm{SF}}(t)x_{\mathrm{SF}}(t). \tag{4.86}$$

[4]See Example 4.2 for WSS processes.

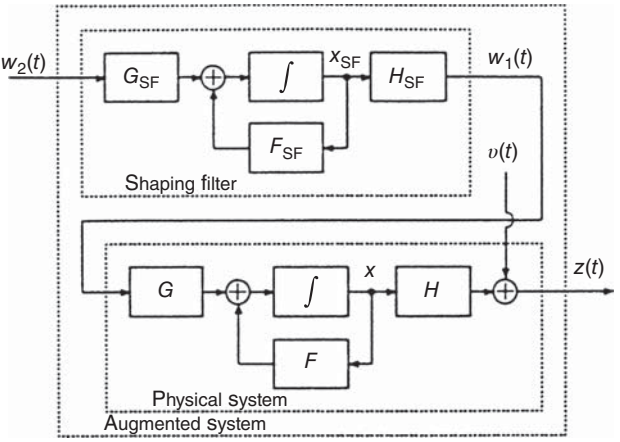


Figure 4.5 Shaping filter model for nonwhite noise.

The total augmented system is given by

$$\begin{bmatrix} \dot{x}(t) \\ \dot{x}_{SF}(t) \end{bmatrix} = \begin{bmatrix} F(t) & 0 \\ 0 & F_{SF}(t) \end{bmatrix} \begin{bmatrix} x(t) \\ x_{SF}(t) \end{bmatrix} + \begin{bmatrix} G(t) & 0 \\ 0 & G_{SF}(t) \end{bmatrix} \begin{bmatrix} w(t) \\ v_2(t) \end{bmatrix},$$

$$z(t) = [H(t)\ H_{SF}(t)] \begin{bmatrix} x(t) \\ x_{SF}(t) \end{bmatrix}. \tag{4.87}$$

This is in the form of a linear system model driven by white Gaussian noise and output equation with no input noise.

These systems can be specialized to the WSS process for continuous and discrete cases as by shaping filters shown in Tables 4.1 and 4.2.

Example 4.5 (Accelerometer Error Model) The "midpoint acceleration" error for an acceleration sensor (accelerometer) is defined as the effective acceleration error at the midpoint of the sampling period. The associated error model for an accelerometer in terms of unknown parameters of the sensor is as follows:

$$\Delta_{\beta_m} = \beta_m \otimes \zeta + b_A + h_A \beta_m + \beta_m^2 (FI1 - FX1) + \delta\beta,$$

$$h_A = \begin{bmatrix} S_1 & \delta_{12} & \delta_{13} \\ 0 & S_2 & \delta_{23} \\ 0 & 0 & S_3 \end{bmatrix}$$

$$\beta_m^2 = \begin{bmatrix} \beta_m^2 & 0 & 0 \\ 0 & \beta_2^2 & 0 \\ 0 & 0 & \beta_3^2 \end{bmatrix}$$

where

Δ_{β_m} is the midpoint acceleration error,
$\otimes$ is the cross product for 3-vectors,
ζ is a 3×1 vector representing attitude alignment errors between "platform" axes and computational axes,
b_A is a 3×1 vector of unknown accelerometer biases, normalized to the magnitude of gravity,
S_i are unknown accelerometer scale factor errors ($i = 1, 2, 3$),
δ_{ij} are unknown accelerometer axes nonorthogonalities,
$\delta\beta$ are other error terms, some of which are observable; for reason of practicality in our example, they are not estimated, only compensated with factory-calibrated values,
$FI1$ is a 3×1 unknown acceleration-squared nonlinearity for acceleration along the accelerometer input axis,
$FX1$ is a 3×1 unknown acceleration-squared nonlinearity for acceleration normal to the accelerometer input axis, and
β_m is a 3×1vector$(\beta_1, \beta_2, \beta_3)^{\mathrm{T}}$ of midpoint components of acceleration in platform coordinates.

The 12×1 accelerometer state vector x^A is composed of the subvectors and scalars

$$(x^A) = \left[\underbrace{b_A^{\mathrm{T}}}_{1\times3} \quad S_1 \quad \delta_{12} \quad S_2 \quad \delta_{13} \quad \delta_{23} \quad S_3 \quad \underbrace{(FX1 - FI1)^{\mathrm{T}}}_{1\times3} \right]^{\mathrm{T}}.$$

The 12 unknown parameters can be modeled as random walks (see Table 4.1) for the parameter identification problem to be discussed in Chapter 10.

An even larger sensor model is described in the next example.

Example 4.6 (Gyroscope Error Model) A 48-state gyroscope drift error model is given as follows:

$$\varepsilon = b_g + h_g\omega + U_g\beta + K_g\beta^1 + \mathrm{diag}|\omega|T_g + b_{gt}t + U_{gt}t\beta,$$

$$h_g = \begin{bmatrix} S_{g1} & \Delta_{12} & \Delta_{13} \\ \Delta_{21} & S_{g2} & \Delta_{23} \\ \Delta_{31} & \Delta_{32} & S_{g3} \end{bmatrix}$$

$$U_g = \begin{bmatrix} d_{I1} & d_{01} & d_{S1} \\ d_{S2} & d_{I2} & d_{02} \\ d_{03} & d_{S3} & d_{I3} \end{bmatrix}$$

$$K_g = \begin{bmatrix} k_{II1} & k_{001} & k_{SS1} & I_{I01} & k_{IS1} & k_{S01} \\ k_{SS2} & k_{II2} & k_{002} & k_{IS2} & k_{S02} & k_{I02} \\ k_{003} & k_{SS3} & k_{II3} & k_{S03} & k_{I03} & k_{IS3} \end{bmatrix}$$

$$\beta^1 = \begin{bmatrix} \beta_1^2 & \beta_2^2 & \beta_3^2 & \beta_1\beta_2 & \beta_1\beta_3 & \beta_2\beta_3 \end{bmatrix}^{\mathrm{T}}$$

$$x^g(t) = \begin{bmatrix} 1\times 3 & 1\times 9 & 1\times 9 & 1\times 15 & 1\times 3 & 1\times 3 & 1\times 3 \\ b_g^{\mathrm{T}} & h_g^{1T} & U_g^{1\mathrm{T}} & K_g^{1\mathrm{T}} & T_g^{\mathrm{T}} & b_{gt}^{\mathrm{T}} & U_{gt}^{1\mathrm{T}} \end{bmatrix}^{\mathrm{T}},$$

where

x^g is the 48-state gyroscope subsystem state vector of drifting error parameters,

b_g is a 3×1 vector of unknown gyroscope fixed drift parameters,

h_g is a 3×1 matrix containing unknown scale factor (S_{gi}) and linear axes alignment errors (Δ_{ij}) as components ($i, j = 1, 2, 3$),

T_g is a 3×1 vector of unknown nonlinear gyroscope torquer scale factor errors, with elements δS_{gi},

$diag\ |\omega|$ is a 3×3 diagonal matrix composed of absolute values of the components of ω (platform inertial angular rate) on the corresponding diagonal element,

U_g is a 3×3 matrix of unknown gyroscope mass unbalance parameters ($d_{k,j}$), indices I, 0, and S denoting input, output, and spin axes, respectively, for each gyroscope 1, 2 and 3,

K_g is a 3×6 matrix of unknown gyroscope compliance (g-squared) errors k_{kji},

b_{gt} is a 3×1 vector of unknown gyroscope fixed drift trend parameters,

U_{gt} is a 3×6 matrix of unknown gyroscope mass unbalance trend parameters, and

β is a 3×1 vector of vertical direction cosines (normalized gravity) $(\beta_1, \beta_2, \beta_3)^{\mathrm{T}}$

β^1 is a 6×1 vector.

The 48 unknown parameters are modeled as random walks and random ramps (see Table 4.1) for the parameter identification problem to be discussed in Chapter 10 (GNSS/INS integration).

4.6 MEAN AND COVARIANCE PROPAGATION

The time propagation of the means and covariances of linear stochastic systems is fundamental to Kalman filtering.

This section combines the fundamental equations for solutions of nonhomogeneous linear differential equations in Chapter 2 and equations for linear transformations of means and covariances derived in Chapter 3 to obtain solutions to generic linear stochastic differential equations of the sort

$$\frac{d}{dt}x(t) = F(t)\ x(t) + G(t)\ w(t),$$

where $w(t)$ is a zero-mean white-noise process.

The following subsections address these issues for linear stochastic systems in continuous time and in discrete time.

4.6.1 Propagating the Mean

4.6.1.1 General Solution If the dynamics of an n-dimensional state vector x can be modeled by a nonhomogeneous linear differential equation with disturbance $u(t)$ of the sort

$$\frac{d}{dt}x(t) = F(t)x(t) + G(t)w(t), \tag{4.88}$$

then the general solution is given by Equation 2.26 in terms of its initial value at $t = t_0$ as

$$x(t) = \Phi(t, t_0)\ x(t_0) + \int_{t+0}^{t} \Phi(t, s)\ G(s)\ w(s)\ ds \tag{4.89}$$

$$\Phi\,(t, s) \stackrel{\text{def}}{=} \exp\ \left[\int_{s}^{t} F\,(\tau)\ d\tau\right]. \tag{4.90}$$

Consequently, for $x \in X$, a random variate, the mean can be propagated as

$$\mu_x(t) \stackrel{\text{def}}{=} \underset{x\in X}{\mathrm{E}}\langle x(t)\rangle \tag{4.91}$$

$$= \underset{x}{\mathrm{E}}\left\langle \Phi\left(t, t_0\right)\ x(t) + \int_{t+0}^{t} \Phi(t, s)\ G(s)\ w(s)\ ds \right\rangle \tag{4.92}$$

$$= \Phi(t, t_0)\ \underset{x}{\mathrm{E}}\,\langle x(t_0)\rangle + \int_{t+0}^{t} \Phi(t, s)\ G(s)\ w(s)\ ds \tag{4.93}$$

$$= \Phi(t, t_0)\ \mu_x(t_0) + \int_{t+0}^{t} \Phi(t, s)\ G(s)\ w(s)\ ds. \tag{4.94}$$

Generic Linear Stochastic Differential Equation Solution If the disturbance function $u(t) = w(t)$, a *white-noise process* with mean $\mathrm{E}_w\langle w(t)\rangle = 0$, the generic linear differential equation becomes a generic linear *stochastic* differential equation. In that case, the expected solution for the mean of x would be

$$\mu_x(t) = \Phi(t, t_0)\ \underset{x}{\mathrm{E}}\,\langle x(t_0)\rangle + \int_{t+0}^{t} \Phi(t, s)\ G(s)\ \underset{w}{\mathrm{E}}\,\langle w(s)\rangle\ ds \tag{4.95}$$

$$= \Phi(t, t_0)\ \mu_x(t_0) + \int_{t+0}^{t} \Phi(t, s)\ G(s)\ (0)\ ds \tag{4.96}$$

$$= \Phi(t, t_0)\ \mu_x(t_0), \tag{4.97}$$

$$= \exp\ \left[\int_{t_0}^{t} F\,(\tau)\ d\tau\right] \mu_x(t_0), \tag{4.98}$$

the general solution for the mean for the generic linear stochastic differential equation model.

4.6.1.2 Differential Equation Model for the Mean Equation 4.98 can be differentiated with respect to t to obtain a differential equation,

$$\frac{d}{dt}\mu_x(t) = \frac{d}{dt}\exp\left[\int_{t_0}^{t} F(\tau)\ d\tau\right]\mu_x(t_0) \tag{4.99}$$

$$= F(t)\ \exp\left[\int_{t_0}^{t} F(\tau)\ d\tau\right]\mu_x(t_0) \tag{4.100}$$

$$= F(t)\ \mu_x(t) \tag{4.101}$$

for propagating the mean in continuous time.

4.6.1.3 Discrete-Time Model for the Mean Equation 4.97 is also the solution in discrete time if we set $t_0 = t_{k-1}$ and $t = t_k$, in which case

$$\mu_{x,k} \overset{\text{def}}{=} \mathop{\mathrm{E}}_{x}\langle x_k\rangle \tag{4.102}$$

$$= \Phi_{k-1}\mu_{x,k-1} \tag{4.103}$$

$$\Phi_{k-1} \overset{\text{def}}{=} \Phi(t_k, t_{k-1}) \tag{4.104}$$

$$= \exp\left[\int_{t_{k-1}}^{t_k} F(s)\ ds\right]. \tag{4.105}$$

4.6.2 Propagating the Covariance

4.6.2.1 General Solution for the Stochastic System The covariance matrix

$$P_{xx}(t) \overset{\text{def}}{=} \mathop{\mathrm{E}}_{x}\langle[x(t) - \mu_x(t)][x(t) - \mu_x(t)]^{\mathrm{T}}\rangle, \tag{4.106}$$

where $x(t)$ is given by Equation 4.89 with $u(t) = w(t)$, $\mu_x(t)$ is given by Equation 4.97, and their difference

$$x(t) - \mu_x(t) = \Phi(t, t_0)\ [x(t_0) - \mu_x(t_0)] + \int_{t_0}^{t} \Phi(t, s)\ G(s)\ w(s)\ ds \tag{4.107}$$

$$\Phi(t, s) \overset{\text{def}}{=} \exp\left[\int_{s}^{t} F(\tau)\ d\tau\right]. \tag{4.108}$$

As a consequence, the covariance matrix

$$P_{xx}(t) = \mathrm{E}_{x,w}\left\langle\left[\Phi\left(t, t_0\right)\ [x(t_0) - \mu_x(t_0)] + \int_{t_0}^{t} \Phi(t, s)\ G(s)\ w(s)\ ds\right]\right.$$

$$\left.\times\left[\Phi\left(t, t_0\right)\ [x(t_0) - \mu_x(t_0)] + \int_{t_0}^{t} \Phi(t, s)\ G(s)\ w(s)\ ds\right]^{\mathrm{T}}\right\rangle \tag{4.109}$$

$$= \underset{x}{\mathrm{E}} \langle [\Phi(t,t_0)\ [x(t_0) - \mu_x(t_0)]][\Phi(t,t_0)\ [x(t_0) - \mu_x(t_0)]]^{\mathrm{T}} \rangle$$
$$+ \underset{w}{\mathrm{E}} \left\langle \left[\int_{t_0}^{t} \Phi\left(t, s_1\right)\ G(s_1)\ w(s_1)\ ds_1 \right] \right.$$
$$\left. \times \left[\int_{t_0}^{t} \Phi\left(t, s_2\right)\ G(s_2)\ w(s_2)\ ds_2 \right]^{\mathrm{T}} \right\rangle \tag{4.110}$$
$$= \Phi(t,t_0) P_{xx}(t_0) \Phi^{\mathrm{T}}(t,t_0)$$
$$+ \int_{t_0}^{t} \Phi(t,s)\ G(s) \underset{w}{\mathrm{E}} \langle w(s)\ w^{\mathrm{T}}(s) \rangle G^{\mathrm{T}}(s)\ \Phi^{\mathrm{T}}(t,s)\ ds \tag{4.111}$$
$$= \Phi(t,t_0) P_{xx}(t_0) \Phi^{\mathrm{T}}(t,t_0) + \int_{t_0}^{t} \Phi(t,s)\ G(s)\ Q_t(s)\ G^{\mathrm{T}}(s)\ \Phi^{\mathrm{T}}(t,s)\ ds, \tag{4.112}$$

where we have used the notation Q_t to denote the dynamic disturbance noise covariance in continuous time, which has units of squared deviation per unit of time. This is to distinguish it from the covariance of dynamic disturbance noise in discrete time, which has units of squared deviation.

In the sequence of equations from Equation 4.109 to Equation 4.112,

(a) The transition between Equations 4.109 and 4.110 used the fact that

$$\underset{w}{\mathrm{E}} \langle w(s) \rangle = 0.$$

(b) The transition between Equations 4.110 and 4.111 used the fact that

$$\underset{w}{\mathrm{E}} \langle w(s_1) w^{\mathrm{T}}(s_2) \rangle = 0$$

unless $s_1 = s_2$, which collapses the double integral down to a single integral.

(c) The transition between Equations 4.111 and 4.112 used the fact that

$$\underset{w}{\mathrm{E}} \langle w(s) w^{\mathrm{T}}(s) \rangle = Q_t(s).$$

4.6.2.2 Differential Equation Model If one writes Equation 4.112 in the alternative form

$$P_{xx}(t) = \exp\ \left[\int_{t_0}^{t} F(\tau)\ d\tau \right]\ P_{xx}(t_0)\ \exp\ \left[\int_{t_0}^{t} F(\tau)\ d\tau \right]$$
$$+ \int_{t_0}^{t} \exp\ \left[\int_{s}^{t} F(\tau)\ d\tau \right]\ G(s)\ Q_t(s)\ G^{\mathrm{T}}(s)\ \exp\ \left[\int_{s}^{t} F^{\mathrm{T}}(\tau)\ d\tau \right]\ ds, \tag{4.113}$$

one can differentiate both sides with respect to t to obtain

$$\begin{aligned}\frac{d}{dt}P_{xx}(t) &= \frac{d}{dt}\left\{\exp\left[\int_{t_0}^{t} F(s)\,ds\right] P_{xx}(t_0)\,\exp\left[\int_{t_0}^{t} F(s)\,ds\right]\right\} \\ &\quad + \frac{d}{dt}\left\{\int_{t_0}^{t}\exp\left[\int_{s}^{t} F(\tau)\,d\tau\right] G(s)\,Q_t(s)\,G^{\mathrm{T}}(s)\right. \\ &\quad \left.\times\exp\left[\int_{s}^{t} F^{\mathrm{T}}(\tau)\,d\tau\right] ds\right\} \qquad (4.114) \\ &= F(t)\exp\left[\int_{t_0}^{t} F(s)\,ds\right] P(t_0)\,\exp\left[\int_{t_0}^{t} F^{\mathrm{T}}(s)\,ds\right] \\ &\quad + \exp\left[\int_{t_0}^{t} F(s)\,ds\right] P(t_0)\,\exp\left[\int_{t_0}^{t} F^{\mathrm{T}}(s)\,ds\right] F^{\mathrm{T}}(t) \\ &\quad + F(t)\left\{\int_{t_0}^{t}\exp\left[\int_{s}^{t} F(\tau)\,d\tau\right] G(s)Q(s)G^{\mathrm{T}}(s)\right. \\ &\quad \left.\times\exp\left[\int_{s}^{t} F^{\mathrm{T}}(\tau)\,d\tau\right] ds\right\} \\ &\quad + \left\{\exp\left[\int_{s}^{t} F(\tau)\,d\tau\right] G(s)Q(s)G^{\mathrm{T}}(s)\exp\left[\int_{s}^{t} F^{\mathrm{T}}(\tau)\,d\tau\right] ds\right\} F^{\mathrm{T}}(t) \\ &\quad + G(t)Q(t)G^{\mathrm{T}}(t) \qquad (4.115) \\ &= F(t)P_{xx}(t) + P_{xx}(t)F^{\mathrm{T}}(t) + G(t)Q_t(t)G^{\mathrm{T}}\,(t). \qquad (4.116)\end{aligned}$$

4.6.2.3 Discrete-Time Model For the generic discrete-time process model

$$x_k = \Phi_{k-1}x_{k-1} + G_{k-1}w_{k-1} \qquad (4.117)$$

$$\mathop{\mathrm{E}}_{w}\langle w_i w_j^{\mathrm{T}}\rangle = \begin{cases} 0, & i \neq j \\ Q_i, & i = j, \end{cases} \qquad (4.118)$$

propagation of the mean μ_k is modeled by Equation 4.103, so that the central difference and its outer product

$$x_k - \mu_k = \Phi_{k-1}[x_{k-1} - \mu_{k-1}] + G_{k-1}w_{k-1} \qquad (4.119)$$

$$\begin{aligned}[x_k - \mu_k][x_k - \mu_k]^{\mathrm{T}} &= \Phi_{k-1}[x_{k-1} - \mu_{k-1}][x_{k-1} - \mu_{k-1}]^{\mathrm{T}}\Phi_{k-1}^{\mathrm{T}} \\ &\quad + G_{k-1}w_{k-1}w_{k-1}^{\mathrm{T}}G_{k-1}^{\mathrm{T}} \\ &\quad + \Phi_{k-1}[x_{k-1} - \mu_{k-1}]w_{k-1}^{\mathrm{T}}G_{k-1}^{\mathrm{T}} \\ &\quad + G_{k-1}w_{k-1}[x_{k-1} - \mu_{k-1}]^{\mathrm{T}}\Phi_{k-1}^{\mathrm{T}}, \qquad (4.120)\end{aligned}$$

and the expected values

$$P_k \overset{\text{def}}{=} \underset{x,w}{\mathrm{E}} \langle [x_k - \mu_k][x_k - \mu_k]^{\mathrm{T}} \rangle \tag{4.121}$$

$$\begin{aligned} &= \underset{x}{\mathrm{E}} \langle \Phi_{k-1}[x_{k-1} - \mu_{k-1}][x_{k-1} - \mu_{k-1}]^{\mathrm{T}} \Phi_{k-1}^{\mathrm{T}} \rangle \\ &\quad + \underset{w}{\mathrm{E}} \langle G_{k-1} w_{k-1} w_{k-1}^{\mathrm{T}} G_{k-1}^{\mathrm{T}} \rangle \\ &\quad + \underset{x}{\mathrm{E}} \langle \Phi_{k-1}[x_{k-1} - \mu_{k-1}] \rangle \underbrace{\underset{w}{\mathrm{E}} \langle w_{k-1}^{\mathrm{T}} G_{k-1}^{\mathrm{T}} \rangle}_{=0} \\ &\quad + \underbrace{\underset{w}{\mathrm{E}} \langle G_{k-1} w_{k-1} \rangle}_{=0} \underset{x}{\mathrm{E}} \langle [x_{k-1} - \mu_{k-1}]^{\mathrm{T}} \Phi_{k-1}^{\mathrm{T}} \rangle \end{aligned} \tag{4.122}$$

$$= \Phi_{k-1} P_{k-1} \Phi_{k-1}^{\mathrm{T}} + G_{k-1} Q_{k-1} G_{k-1}^{\mathrm{T}}. \tag{4.123}$$

4.6.3 Steady-State Solutions

One of the remarkable features of the Kalman filter is that the covariance of estimation uncertainty *with measurements* can have a finite steady-state value even if the covariance equation without sampling is unstable.

The issue considered here is whether the covariance equations *without measurement* are stable, in that their solutions approach a finite constant value as $t \to +\infty$ (continuous-time model) or $k \to +\infty$ (discrete-time model).

4.6.3.1 Continuous Time A steady-state solution for the covariance is a constant value which the covariance approaches as $t \to +\infty$, independent of the initial value of the covariance. Some time-varying and time-invariant systems may have steady-state solutions, but not all time-varying or time-invariant systems have steady-state solutions.

The time-invariant continuous-time model with zero-mean white noise

$$\dot{x}(t) = Fx(t) + w(t) \tag{4.124}$$

$$\underset{w}{\mathrm{E}} \langle w(t) w^{\mathrm{T}}(t) \rangle = Q_t \tag{4.125}$$

will have a steady-state covariance solution only if the steady-state equation

$$0 = \lim_{t \to +\infty} \frac{d}{dt} P(t) \tag{4.126}$$

$$= FP(\infty) + P(\infty) F^{\mathrm{T}} + Q_t \tag{4.127}$$

has a nonnegative definite solution $P(\infty)$. As a rule, it will have such solutions only if the eigenvalues of F have negative real parts.

4.6.3.2 Discrete Time The equivalent steady-state covariance equation in discrete time would be

$$P_\infty = \Phi P_\infty \Phi^{\mathrm{T}} + Q, \tag{4.128}$$

which may or may not have a solution, depending on the eigenvalues of Φ. As a rule, it will have a solution only if all the eigenvalues of Φ are inside the unit circle in the complex plane.

4.6.4 Results

The main results of this section are summarized in Table 4.3.

Example 4.7 (Scalar System in Continuous Time) A dynamic model is given by the scalar differential equation

$$\begin{aligned} \dot{x}(t) &= -x(t) + w(t) \\ \mathrm{E}x(0) &= 0 \\ P(0) &= 0 \\ F &= -1 \\ G &= 1 \\ w(t) &\equiv 1 \\ Q(t) &= 1. \end{aligned}$$

TABLE 4.3 Propagating Essential Moments with White-Noise Inputs

Continuous time equation	$\dot{x}(t) = F(t)x(t) + G(t)w(t)$
Mean value equation	$\frac{d}{dt}\mathrm{E}\langle x(t)\rangle = F(t)\mathrm{E}\langle x(t)\rangle$
Initial value	$\mathrm{E}\langle x(t_o)\rangle$
covariance equation	$\dot{P}(t) = F(t)P(t) + P(t)F^{\mathrm{T}}(t) + G(t)Q_t(t)G^{\mathrm{T}}(t)$
Initial value	$P(t_0)$
steady-state equation	$0 = FP(\infty) + P(\infty)F^{\mathrm{T}} + GQ_tG^{\mathrm{T}}$
(algebraic equation)	(for time-invariant model only)
Discrete-time equation	$x_k = \Phi_{k-1}x_{k-1} + G_{k-1}w_{k-1}$
Mean value equation	$\mathrm{E}\langle x_k\rangle = \Phi_{k-1}\mathrm{E}\langle x_{k-1}\rangle$
Covariance equation	$P_k = \Phi_{k-1}P_{k-1}\Phi_{k-1}^{\mathrm{T}} + G_{k-1}Q_{k-1}G_{k-1}^{\mathrm{T}}$
Initial value	P_0
steady-state equation	$P_\infty = \Phi P_\infty \Phi^{\mathrm{T}} + GQG^{\mathrm{T}}$
(algebraic equation)	(for time-invariant model only)

From Table 4.3,

$$\begin{aligned} \mathrm{E}\dot{x}(t) &= -Ex(t) + 1 \\ \mathrm{E}x(0) &= 0 \\ \mathrm{E}x(t) &= 1 - e^{-t},\ t \geq 0. \end{aligned}$$

The covariance equation is then given by

$$\begin{aligned} \dot{P}(t) &= -2P(t) + 1 \\ P(0) &= 0 \\ P(t) &= e^{-2t}P(0) + \int_0^t e^{-2(t-\tau)}1 \quad d\tau \\ &= \frac{1}{2}(1 - e^{-2t}). \end{aligned}$$

The steady-state covariance is then

$$\begin{aligned} P(\infty) &= \lim_{t\to+\infty} \frac{1}{2}(1 - e^{-2t}) \\ &= \frac{1}{2}. \end{aligned}$$

Example 4.8 (Scalar System in Discrete Time) A discrete-time model is given as

$$\left.\begin{array}{rcll} x_k^1 & = & -x_{k-1}^1 + w_{k-1}^1 & \\ \mathrm{E}w_{k-1}^1 & = & 0 & \text{(non-white)} \\ \Psi_w^1 & = & e^{-(k_1-k_2)} & \text{(autocorrelation)} \end{array}\right\} (a)$$

The model of non-white noise evolves into scalar difference equation (given in Table 4.1), called a shaping filter:

$$x_k^2 = \mathrm{e}^{-1}x_{k-1}^2 + \sqrt{1 - e^{-1}} \qquad w_{k-1} \quad (b),$$

where w_{k-1} is white noise with zero mean and covariance equal to 1.

Combining equation (a) and (b), as given by Equation 4.98,

$$x_k = \begin{bmatrix} -1 & 1 \\ 0 & e^{-1} \end{bmatrix} x_{k-1} + \begin{bmatrix} 0 \\ \sqrt{1 - e^{-1}} \end{bmatrix} w_{k-1}$$

$$P_k = \begin{bmatrix} -1 & 1 \\ 0 & e^{-1} \end{bmatrix} \quad P_{k-1} \begin{bmatrix} -1 & 0 \\ 1 & e^{-1} \end{bmatrix} + \begin{bmatrix} 0 \\ \sqrt{1 - e^{-1}} \end{bmatrix} 1 \begin{bmatrix} 0 & \sqrt{1 - e^{-1}} \end{bmatrix}$$

$$P_k^{11} = +P_{k-1}^{11} - 2P_{k-1}^{12} + P_k^{22}$$

$$P_k^{12} = -e^{-1} P_{k-1}^{12} + e^{-1} P_{k-1}^{22}$$
$$P_k^{22} = e^{-2} P_{k-1}^{22} + (1 - e^{-1})$$
$$P_0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The steady-state solution will be

$$P_\infty^{22} = \frac{(1 - e^{-1})}{1 - e^{-2}}$$
$$P_\infty^{12} = \frac{\left[e^{-1} \frac{(1-e^{-1})}{1-e^{-2}}\right]}{1 + e^{-1}}$$
$$P_\infty^{11} = \text{underdetermined.}$$

Example 4.9 (Steady-State Covariance of a Harmonic Resonator) The stochastic system model

$$\dot{x}(t) = Fx(t) + w(t),$$
$$\mathrm{E}\,\langle w(t_1) w^{\mathrm{T}}(t_2)\rangle = \delta(t_1 - t_2) Q,$$
$$Q = \begin{bmatrix} 0 & 0 \\ 0 & q \end{bmatrix}$$

for an underdamped harmonic resonator driven by zero-mean white *acceleration* noise $w(t)$.

It is of interest to find whether the covariance of the process $x(t)$ reaches a finite *steady-state* value $P(\infty)$. Not every RP has a finite steady-state value, but it will turn out to be finite in this example.

Recall that the state-space model for the mass-spring harmonic resonator from Examples 2.2, 2.3, and 2.7 has as its dynamic coefficient matrix the 2×2 constant matrix

$$F = \begin{bmatrix} 0 & 1 \\ -\frac{k_s}{m} & -\frac{k_d}{m} \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 1 \\ -\omega_r^2 - \omega_d^2 & -2\omega_d \end{bmatrix},$$

where m is the supported mass, k_s is the spring elastic constant, and k_d is the dashpot damping coefficient. The alternate model parameters are

$$\omega_r = \sqrt{\frac{k_s}{m} - \frac{k_d^2}{4m^2}} \text{ (undamped resonant frequency),}$$

$$\tau_d = \frac{2m}{k_d} \text{ (damping time constant)},$$

$$\omega_d = \frac{1}{\tau} \text{ (damping frequency)}.$$

If the covariance matrix $P(t)$ reaches some steady-state value $P(\infty)$ as $t \to +\infty$, the asymptotic covariance dynamic equation becomes

$$\begin{aligned} 0 &= \lim_{t\to\infty} \frac{d}{dt} P(t) \\ &= FP(\infty) + P(\infty)F^{\mathrm{T}} + Q, \end{aligned}$$

which is an *algebraic equation.* That is, it includes only algebraic operations (multiplies and adds) and no derivatives or integrals from the calculus. Moreover, it is a linear equation of the unknown elements p_{11}, $p_{12} = p_{21}$, p_{22} of the 2×2 symmetric matrix $P(\infty)$. Equation 4.129 is equivalent to three scalar linear equations:

$$\begin{aligned} 0 &= \sum_{k=1}^{2} f_{1k} p_{k1} + \sum_{k=1}^{2} p_{1k} f_{k1} + q_{11}, \\ 0 &= \sum_{k=1}^{2} f_{1k} p_{k2} + \sum_{k=1}^{2} p_{1k} f_{k2} + q_{12}, \\ 0 &= \sum_{k=1}^{2} f_{2k} p_{k2} + \sum_{k=1}^{2} p_{2k} f_{k2} + q_{22}, \end{aligned}$$

with known parameters

$$\begin{array}{llllllll} q_{11} &= 0, & q_{12} &= 0, & q_{22} &= q, & & \\ f_{11} &= 0, & f_{12} &= 1, & f_{21} &= -\omega_r^2 - \omega_d^2, & f_{22} &= -2\omega_d. \end{array}$$

The linear system of equations above can be rearranged into the nonsingular 3×3 system of equations

$$\begin{bmatrix} 0 \\ 0 \\ q \end{bmatrix} = -\begin{bmatrix} 0 & 2 & 0 \\ -(\omega_r^2 + \tau_d^{-2}) & -\frac{2}{\tau_d} & 1 \\ 0 & -2(\omega_r^2 + \tau_d^{-2}) & -\frac{4}{\tau_d} \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ p_{22} \end{bmatrix}$$

with solution

$$\begin{bmatrix} p_{11} \\ p_{12} \\ p_{22} \end{bmatrix} = q \begin{bmatrix} \frac{\tau_d}{4(\omega_r^2 + \tau_d^{-2})} \\ 0 \\ \frac{\tau_d}{4} \end{bmatrix},$$

$$P(\infty) = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \frac{q\tau_d}{4(\omega_r^2 + \tau_d^{-2})} \begin{bmatrix} 1 & 0 \\ 0 & (\omega_r^2 + \tau_d^{-2}) \end{bmatrix}.$$

Note that the steady-state state covariance depends linearly on the process noise covariance q. The steady-state covariance of velocity also depends linearly on the damping time constant τ. The dimensionless quantity $2\pi\omega_r\tau_d$ is called the *quality factor, Q-factor*, or simply "the Q" of a resonator. It equals the number of cycles that the unforced resonator will go through before its amplitude falls to $1/e \approx 37\%$ of its initial amplitude.

4.7 RELATIONSHIPS BETWEEN MODEL PARAMETERS

4.7.1 Parameters of Continuous and Discrete Models

The mathematical forms of the application models used in Kalman filtering (in discrete time) and Kalman–Bucy filtering (in continuous time) are summarized in Table 4.4. All these models are written in general-purpose mathematical forms, with the dimensions and values of the model parameters F, Q, H, R, and Φ depending on the application. The relationship between Φ and F (already shown in the table) is derived in Chapter 2, and the solution for nonhomogeneous differential equations derived in Chapter 2 were used—with a little stochastic calculus—to derive the relationship between the parameters $Q(t)$ and Q_{k-1} in Equation 4.131.

4.7.1.1 Sensor Models The last column of Table 4.4 includes a new type of model, labeled "output model" at the top of the table. This type of model was introduced in Section 2.3.4. Output models, which are for the sensors used in estimating the state vector x of a Kalman filter, contain the parameters H and R. The relationship between $H(t)$ and H_k is already shown in the table.

The relationship between $R(t)$ and R_k depends on the type of filtering used before sampling the measurements z_k at discrete times t_k. These types of relationships depend on the same stochastic calculus used in deriving the relationships between $Q(t)$ and Q_{k-1}. In Section 4.7.3, we will derive the $R(t) \rightarrow R_k$ relationship for some common types of anti-aliasing filters.

4.7.1.2 The Utility of Continuous-Time Models These relationships are important for many applications of Kalman filtering where the source models are in continuous time, and the Kalman filter must be implemented on a computer in discrete time. Engineers and scientists often find it more natural and reliable to start with the model in continuous time and then convert the model to discrete time after they have full confidence in the model they have developed. The risk of blunders in the filter derivation can usually be reduced by following the maxim:

Think continuously. Act discretely.

TABLE 4.4 Parameters of Stochastic System Models

Filter Type	Dynamic Model	Output Model
Kalman–Bucy	$\dot{x}(t) = F(t)\, x(t) + w(t)$	$z(t) = H(t)\, x(t) + v(t)$
Parameters	$F(t) = \frac{\partial \dot{x}}{\partial x}$	$H(t)\frac{\partial z}{\partial x}$
	$Q(t) = \mathrm{E}\,\langle w(t) w^{\mathrm{T}}(t)\rangle$	$R(t) = \mathrm{E}\,\langle v(t) v^{\mathrm{T}}(t)\rangle$
Kalman	$x_k = \Phi_{k-1}\, x_{k-1} + w_{k-1}$	$z_k = H_k\, x_k + v_k$
Parameters	$\Phi_{k-1} = \exp\left(\int_{t_{k-1}}^{t_k} F(s)\, ds\right)$	$H_k = H(t_k)$
	$Q_{k-1} = \mathrm{E}\,\langle w_{k-1} w_{k-1}^{\mathrm{T}}\rangle$	$R_k = \mathrm{E}\,\langle v_k v_k^{\mathrm{T}}\rangle$

4.7.2 Relationship between $Q(t)$ and Q_{k-1}

If we let $t = t_k$ and $t_0 = t_{k-1}$, then Equation 4.112 becomes

$$P_k = \Phi_{k-1} P_{k-1} \Phi_{k-1} + G_{k-1} Q_{k-1} G_{k-1}^{\mathrm{T}} \tag{4.129}$$

$$G_{k-1} Q_{k-1} G_{k-1}^{\mathrm{T}} = \int_{t_{k-1}}^{t_k} \exp\left[\int_s^{t_k} F(\tau)\, d\tau\right] G(s)\, Q_t(s)\, G^{\mathrm{T}}(s) \times \exp\left[\int_s^{t_k} F(\tau)\, d\tau\right]^{\mathrm{T}} ds, \tag{4.130}$$

where the last equation defines the relationship between the equivalent white-noise processes covariances in continuous time (Q_t) and discrete time (Q_k). In the case that $G_{k-1} = I = G(s)$, this becomes

$$Q_{k-1} = \int_{t_{k-1}}^{t_k} \exp\left[\int_s^{t_k} F(\tau)\, d\tau\right] Q_t(s)\, \exp\left[\int_s^{t_k} F(\tau)\, d\tau\right]^{\mathrm{T}} ds. \tag{4.131}$$

Solutions for some time-invariant system models are shown in Table 4.5

4.7.2.1 Covariance Units Note that the process noise covariance matrices $Q(s)$ and Q_{k-1} have different physical units. The units of $w(t) \propto \dot{x}$ are the units of the state vector x divided by time, and the units of $w_{k-1} \propto x$ are the same as those of the state vector x. One might then expect that the units of $Q(s) = \mathrm{E}\,\langle w(s) w^{\mathrm{T}}(s)\rangle$ would be the units of xx^{T} divided by time squared. However, the expectancy operator combined with the stochastic integral (of Itô or Stratonovich) makes the units of $Q(s)$ be the units of xx^{T} divided by time—a potential source of confusion. Integration of $Q(s)$ over time then makes the units of Q_k equal to those of xx^{T}.

Example 4.10 (Q_{k-1} for the Harmonic Resonator Model) Consider the problem of determining the covariance matrix Q_{k-1} for the equivalent *discrete-time model*

$$x_k = \Phi_{k-1} x_{k-1} + w_{k-1},$$
$$\mathrm{E}\,\langle w_{k-1} w_{k-1}^{\mathrm{T}}\rangle = Q_{k-1}$$

for a harmonic resonator driven by white acceleration noise, given the variance q of the process noise in its *continuous-time model*

TABLE 4.5 Equivalent Constant Parameters, Continuous- and Discrete-Time models

Continuous Time		Discrete Time	
F	Q_t	Φ_k	Q_k
$[0]$	$[q]$	$[1]$	$[q\ \Delta t]$
$[-1/\tau]$	$[q]$	$[\exp\ (-\Delta t/\tau)]$	$\left[\frac{q\tau\ \left[1-\exp\ (-2\ \Delta t/\tau)\right]}{2}\right]$
$\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}$	$\begin{bmatrix}0 & 0\\ 0 & q\end{bmatrix}$	$\begin{bmatrix}1 & \Delta t\\ 0 & 1\end{bmatrix}$	$q\begin{bmatrix}\frac{(\Delta t)^3}{3} & \frac{(\Delta t)^2}{2}\\ \frac{(\Delta t)^2}{2} & \Delta t\end{bmatrix}$
$\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{bmatrix}$	$\begin{bmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & q\end{bmatrix}$	$\begin{bmatrix}1 & \Delta t & \frac{(\Delta t)^2}{2}\\ 0 & 1 & \Delta t\\ 0 & 0 & 1\end{bmatrix}$	$q\begin{bmatrix}\frac{(\Delta t)^5}{20} & \frac{(\Delta t)^4}{8} & \frac{(\Delta t)^3}{6}\\ \frac{(\Delta t)^4}{8} & \frac{(\Delta t)^3}{3} & \frac{(\Delta t)^2}{2}\\ \frac{(\Delta t)^3}{6} & \frac{(\Delta t)^2}{2} & \Delta t\end{bmatrix}$

$$\frac{d}{dt}x(t) = Fx(t) + w(t),$$
$$\mathrm{E}\,\langle w(t_1)\ w^{\mathrm{T}}(t_2)\rangle = \delta(t_1 - t_2)Q,$$
$$Q = \begin{bmatrix}0 & 0\\ 0 & q\end{bmatrix},$$

where ω is the resonant frequency, τ is the damping time constant, ξ is the corresponding damping "frequency" (i.e., has frequency units), and q is the process noise covariance in continuous time. (This stochastic system model for a harmonic resonator driven by white acceleration noise $w(t)$ is derived in Examples 4.2 and 4.9.)

Following the derivation of Example 2.6, the fundamental solution matrix for the *unforced* dynamic system model can be expressed in the form

$$\begin{aligned}\Phi(t) &= e^{Ft}\\ &= e^{-t/\tau}\begin{bmatrix}\frac{S(t)+C(t)\omega\tau}{\omega\tau} & \frac{S(t)}{\omega}\\ -\frac{S(t)(1+\omega^2\tau^2)}{\omega\tau^2} & \frac{-S(t)+C(t)\omega\tau}{\omega\tau}\end{bmatrix},\\ S(t) &= \sin(\omega t),\\ C(t) &= \cos(\omega t),\end{aligned}$$

in terms of its resonant frequency ω and damping time constant τ. Its matrix inverse

$$\Phi^{-1}(s) = \frac{e^{s/\tau}}{\omega\tau^2}\begin{bmatrix}\tau[\omega\tau C(s) - S(s)] & -\tau^2 S(s)\\ (1+\omega^2\tau^2)S(s) & \tau[\omega\tau C(s) + S(s)]\end{bmatrix}$$

at time $t = s$. Consequently, the indefinite integral matrix

$$\begin{aligned}
\Psi(t) &= \int_0^t \Phi^{-1}(s) \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \Phi^{\mathrm{T}-1}(s)\, ds \\
&= \frac{q}{\omega^2\tau^2} \int_0^t \begin{bmatrix} \tau^2 S(s)^2 & -\tau S(s)[\omega\tau C(s) + S(s)] \\ -\tau S(s)[\omega\tau C(s) + S(s)] & [\omega\tau C(s) + S(s)]^2 \end{bmatrix} e^{2s/\tau}\, ds \\
&= \begin{bmatrix} \dfrac{q\tau\{-\omega^2\tau^2 + [2S(t)^2 - 2C(t)\omega S(t)\tau + \omega^2\tau^2]\zeta^2\}}{4\omega^2(1+\omega^2\tau^2)} & \dfrac{-qS(t)^2\zeta^2}{2\omega^2} \\ \dfrac{-qS(t)^2\zeta^2}{2\omega^2} & \dfrac{q\{-\omega^2\tau^2 + [2S(t)^2 + 2C(t)\omega S(t)\tau + \omega^2\tau^2]\zeta^2\}}{4\omega^2\tau} \end{bmatrix}, \\
\zeta &= e^{t/\tau}.
\end{aligned}$$

The discrete-time covariance matrix Q_{k-1} can then be evaluated as (see Section 4.7)

$$\begin{aligned}
Q_{k-1} &= \Phi(\Delta t)\Psi(\Delta t)\Phi^{\mathrm{T}}(\Delta t) \\
&= \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix}, \\
q_{11} &= \frac{q\tau\{\omega^2\tau^2(1 - e^{-2\Delta t/\tau}) - 2S(\Delta t)e^{-2\Delta t/\tau}[S(\Delta t) + \omega\tau C(\Delta t)]\}}{4\omega^2(1+\omega^2\tau^2)}, \\
q_{12} &= \frac{qe^{-2\Delta t/\tau}S(\Delta t)^2}{2\omega^2}, \\
q_{21} &= q_{12}, \\
q_{22} &= \frac{q\{\omega^2\tau^2(1 - e^{-2\Delta t/\tau}) - 2S(\Delta t)e^{-2\Delta t/\tau}[S(\Delta t) - \omega\tau C(\Delta t)]\}}{4\omega^2\tau}.
\end{aligned}$$

Note that the *structure* of the discrete-time process noise covariance Q_{k-1} for this example is quite different from the continuous-time process noise Q. In particular, Q_{k-1} is a full matrix, although Q is a sparse matrix.

4.7.2.2 First-order approximation of Q_k for constant F and G The justification of a truncated power series expansion for Q_k when F and G are constant is as follows:

$$Q_k = \sum_{i=1}^{\infty} \frac{Q^i \Delta t^i}{i!}. \tag{4.132}$$

Consider the Taylor series expansion of Q_k about t_{k-1}, where

$$Q^i = \left.\frac{d^i Q}{dt^i}\right|_{t=t_{k-1}},$$

$$\begin{aligned}
\dot{Q} &= FQ_k + Q_k F^{\mathrm{T}} + GQ(t)G^{\mathrm{T}}, \\
Q^{(1)} &= \dot{Q}(t_{k-1}) = GQ(t)G^{\mathrm{T}} \text{ since } Q(t_{k-1}) = 0, \\
Q^{(2)} &= \ddot{Q}(t_{k-1}) = F\dot{Q}(t_{k-1}) + \dot{Q}(t_{k-1})F^{\mathrm{T}}, \\
&= FQ^{(1)} + Q^{(1)}F^{\mathrm{T}}, \\
&\vdots \\
Q^{(i)} &= FQ^{(i-1)} + Q^{(i-1)}F^{\mathrm{T}}, \quad i = 1, 2, 3, \ldots
\end{aligned}$$

Taking only first-order terms in the above series,

$$Q_k \approx GQ(t)G^{\mathrm{T}}\Delta t. \tag{4.133}$$

This is not always a good approximation, as is shown in the following example.

Example 4.11 (First-Order Approximation of Q_k for the Harmonic Resonator) Let us see what happens if this first-order approximation

$$Q_k \approx Q\Delta t$$

is applied to the previous example of the harmonic resonator with acceleration noise.

The solution to the steady-state "state covariance" equation (i.e., the equation of covariance of the state vector itself, not the estimation error)

$$P_\infty = \Phi P_\infty \Phi^{\mathrm{T}} + Q\Delta t$$

has the solution (for $\theta = 2\pi f_{\text{resonance}}/f_{\text{sampling}}$)

$$\begin{aligned}
\{P_\infty\}_{11} &= q\Delta t e^{-2\Delta t/\tau} \sin(\theta)^2 (e^{-2\Delta t/\tau} + 1)/D, \\
D &= \omega^2 (e^{-2\Delta t/\tau} - 1) \\
&\quad \times (e^{-2\Delta t/\tau} - 2e^{-2\Delta t/\tau}\cos(\theta) + 1)(e^{-2\Delta t/\tau} + 2e^{-2\Delta t/\tau}\cos(\theta) + 1)
\end{aligned}$$

for its upper-left element, which is the *steady-state MS resonator displacement*. Note, however, that

$$\{P_\infty\}_{11} = 0 \quad \text{if} \quad \sin(\theta) = 0,$$

which would imply that there is *no* displacement if the sampling frequency is twice the resonant frequency. This is absurd, of course. This proves by contradiction that

$$Q_k \neq Q\Delta t$$

in general—even though it may be a reasonable approximation in some instances.

Example 4.12 (Q_k Versus Q_t for Exponentially Correlated Noise Models) The exponentially correlated noise model is linear and time invariant, and the correspondence between the discrete time Q_k and the analogous Q_t in continuous time can be derived in closed form. Starting with the continuous-time model

$$\dot{x} = \frac{-1}{\tau}x(t) + w(t), \text{ for } w(t) \in \mathcal{N}(0, Q_t),$$

$$F = \frac{-1}{\tau}$$

$$\Phi_k = \Phi(\Delta t) = \exp\left(\int_0^{\Delta t} F\, ds\right) = \exp\,(-\Delta t/\tau)$$

$$Q_k = \Phi_k^2 \int_0^{\Delta t} \Phi^{-1}(s) Q_t \Phi^{-\mathrm{T}}(s)\, ds = \frac{\tau}{2}[1 - \exp\,(-2\,\Delta t/\tau)]Q_t$$

$$Q_t = \frac{2}{\tau[1 - \exp\,(-2\,\Delta t/\tau)]} Q_k.$$

4.7.2.3 Van Loan's Method for Computing Q_k from Continuous Q For the general linear time-invariant model in continuous time

$$\dot{x}(t) = Fx(t) + Gw(t) \text{ with } w(t) \in \mathcal{N}(0, Q_t),$$

and with n state variables, form the $2n \times 2n$ matrix

$$M = \Delta t \begin{bmatrix} -F & GQ_tG^{\mathrm{T}} \\ 0 & F^{\mathrm{T}} \end{bmatrix}, \tag{4.134}$$

partitioned into a 2×2 array of $n \times n$ blocks.

A 1978 article by Van Loan [14] shows how the $2n \times 2n$ matrix exponential $\exp(M)$, which can be computed numerically using the MATLAB function `expm`, can also be partitioned into a 2×2 array of $n \times n$ blocks, where

$$\exp(M) = \begin{bmatrix} \Psi & \Phi_k^{-1}Q_k \\ 0 & \Phi^{\mathrm{T}} \end{bmatrix}, \tag{4.135}$$

where Ψ is not used, but the equivalent discrete-time model parameters

$$\texttt{Phik} \stackrel{\text{def}}{=} \Phi_k$$

$$\texttt{Qk} \stackrel{\text{def}}{=} Q_k$$

can be computed in MATLAB software by the script

```
M    = DeltaT*[-F,G*Qc*G';zeros(n),F'];
N    = expm(M);
```

```
Phik = N(n+1:2*n,n+1:2*n)';
Qk   = Phik*N(1:n,n+1:2*n);
```

where the MATLAB variable `Qc` $\stackrel{\text{def}}{=} Q_t$.

Van Loan's method is coded in the MATLAB function `VanLoan` on the Wiley web site.

Example 4.13 (Using Van Loan's Method) Let the continuous-time model for a linear time-invariant stochastic system have parameter values

$$F = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, GQ_tG^{\mathrm{T}} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

and its equivalent implementation in discrete time require time-step $\Delta t = 1$. Then

$$M = \Delta t \begin{bmatrix} -F & GQ_tG^{\mathrm{T}} \\ 0 & F^{\mathrm{T}} \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\begin{aligned} \exp(M) &= I + M + \frac{1}{2!}M^2 + \frac{1}{3!}M^3 + \frac{1}{4!}M^4 + \cdots \\ &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &\quad + \frac{1}{6}\begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \frac{1}{24}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \text{more zero terms} \\ &= \begin{bmatrix} 1 & -1 & -\frac{1}{6} & -\frac{1}{2} \\ 0 & 1 & \frac{1}{2} & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \Psi & \Phi_k^{-1}Q_k \\ 0 & \Phi^{\mathrm{T}} \end{bmatrix}, \end{aligned}$$

from which

$$\Phi_k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

$$Q_k = \Phi_k\{\Phi_k^{-1}Q_k\} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}\left\{\begin{bmatrix} -\frac{1}{6} & -\frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}\right\} = \begin{bmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}.$$

Because $F^2 = 0$, the value of Φ_k obtained above can easily be verified by calculating it as

$$\Phi_k = \exp\ (\Delta t\ F)$$

$$= I + \Delta t\, F + \frac{1}{2!}\underbrace{(\Delta t\, F)^2}_{=0} + \cdots$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

which can be verified in MATLAB, as well:

```
>> F = [0,1;0,0],
F =
     0     1
     0     0
>> Qc = [0,0;0,1],
Qc =
     0     0
     0     1
>> DeltaT = 1;
>> M = DeltaT*[-F,Qc;zeros(2),F'],
M =
     0    -1     0     0
     0     0     0     1
     0     0     0     0
     0     0     1     0
```

```
>> N = expm(M),
N =
    1.0000   -1.0000   -0.1667   -0.5000
         0    1.0000    0.5000    1.0000
         0         0    1.0000         0
         0         0    1.0000    1.0000
>> Phi = N(3:4,3:4)',
Phi =
     1     1
     0     1
>> Qk = Phi*N(1:2,3:4),
Qk =
    0.3333    0.5000
    0.5000    1.0000
```

4.7.3 Relationship between R(t) and R_k

Sensor Noise Calibration. In practice, $R(t)$ and/or R_k should be determined by measuring the actual sensor outputs with constant inputs—as part of sensor characterization and/or calibration. In either case (continuous time or discrete time), the noise measurements would also be used to determine whether the appropriate noise model is zero-mean white noise. If not, then the same noise data (or its PSD) can be used to devise a suitable shaping filter (i.e., an auxiliary noise model as a stochastic system with white-noise inputs). The values of $R(t)$ or R_k would then be those which approximate—at the output of the shaping filter—the statistical properties of the measured sensor noise.

Sensor Output Bias. This sort of noise calibration is also needed to detect and correct any nonzero mean of the noise, which is usually called "sensor output bias." Sensor output bias can also be estimated and compensated by making it a system state variable, to be estimated by a Kalman filter. This approach also works for biases that are not constant.

4.7.3.1 Integrating Sensors Even if $R(t)$ and R_k are already zero-mean white noise, the relationship between them depends upon the way that the discrete-time sensor processes internal continuous-time noise. If it is an integrating sensor or uses an integrate-and-hold circuit before sampling, then

$$v_k = \int_{t_{k-1}}^{t_k} v(t)\, dt, \tag{4.136}$$

$$R_k = \overline{R} \tag{4.137}$$

$$= \frac{1}{t_k - t_{k-1}} \int_{t_{k-1}}^{t_k} R(t)\, dt, \tag{4.138}$$

where $\overline{R}$ is the time-averaged value of $R(t)$ over the interval $t_{k-1} < t \leq t_k$. If the analog noise $v(t)$ is zero-mean and white, then the integrate-and-hold circuit is the ideal low pass anti-aliasing filter. It integrates only over the intersampling interval, which keeps the output noise white (i.e., introduces no time correlation).

4.7.3.2 Other Anti-Alias Filters *Tapped Delay Lines.* Filters implemented with tapped delay lines can also keep total delays with an intersampling interval by keeping the total delay less than or equal to the output intersampling interval. The internal sampling rate can still be much higher than the rate at which the output is sampled.

IIR Filters. Analog RC filters, on the other hand, have infinite impulse response (IIR). This introduces some amount of lag beyond the intersampling interval, which creates time correlation in the output noise—even if the input noise is white.

Calculating R_k. In general, the filtered output noise covariance R_k can be calculated by using the input noise model and the transfer function of the anti-aliasing filter.

4.8 ORTHOGONALITY PRINCIPLE

4.8.1 Estimators Minimizing Expected Quadratic Loss Functions

A block diagram representation of an estimator of the state of a system represented by Equation 4.64 is shown in Figure 4.6. The estimate $\hat{x}(t)$ of $X(t)$ will be the output of a Kalman filter.

The *estimation error* is defined as the difference between the "true" value of a RV $X(t)$ and an "estimate" $\hat{x}(t)$.

A quadratic "loss" function of the estimation error has the form

$$[x(t) - \hat{x}(t)]^{\mathrm{T}} M [x(t) - \hat{x}(t)], \tag{4.139}$$

where M is an $n \times n$ symmetric, positive-definite "weighting matrix."

An "optimal" estimator for a particular quadratic loss function is defined as that estimate $\hat{x}(t)$ minimizing the *expected value* of the loss, with the probabilities conditioned on the observation $z(t)$. It will be shown that the optimal estimate of $X(t)$

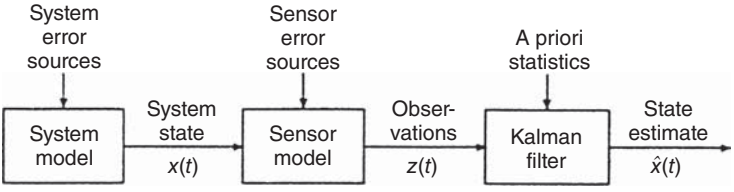


Figure 4.6 Block diagram of estimator model.

(minimizing the average of a quadratic cost function) is the conditional expectation of $X(t)$ given the observation $z(t)$:

$$\hat{x} = \mathrm{E}\,\langle x(t)|z(t)\rangle \quad \text{minimizes}$$
$$\mathrm{E}\,\langle [x(t)-\hat{x}(t)]^{\mathrm{T}} M[x(t)-\hat{x}(t)]|z(t)\rangle. \tag{4.140}$$

Let $z(t), 0 \le t \le t_1$, be the observed quantity and it is desired to estimate $X(t)$ at $t = t_2$. Then Equation 4.140 assumes the form

$$\hat{x}(t_2) = \mathrm{E}\,\langle x(t_2)|z(t), 0 \le t \le t_1\rangle \tag{4.141}$$

and the corresponding equation for a similar discrete model is

$$\hat{x}_{k_2} = \mathrm{E}\,\langle \hat{x}_{k_2}|z_1, z_2, \ldots, z_{k_1}\rangle, \quad 1 \le k_2 \le k_1. \tag{4.142}$$

Let

$$J = \mathrm{E}\,\langle [x(t)-\hat{x}(t)]^{\mathrm{T}} M[x(t)-\hat{x}(t)]|z(t)\rangle. \tag{4.143}$$

Recall that $\hat{x}(t)$ is a nonrandom function of the observations

$$0 = \frac{dJ}{d\hat{x}} \tag{4.144}$$
$$= -2M\,\mathrm{E}\,\langle [x(t)-\hat{x}(t)]|z(t)\rangle, \tag{4.145}$$
$$\mathrm{E}\,\langle \hat{x}(t)|z(t)\rangle = \hat{x}(t) = \mathrm{E}\,\langle x(t)|z(t)\rangle. \tag{4.146}$$

This proves the result of Equation 4.140. If $X(t)$ and $z(t)$ are jointly normal (Gaussian), the nonlinear minimum variance and linear minimum variance estimators coincide

$$\mathrm{E}\,\langle x_{k_2}|z_1, z_2, \ldots, z_{k_1}\rangle = \sum_{i=1}^{k_1} \alpha_i z_i \tag{4.147}$$

and

$$\mathrm{E}\,\langle x(t_2)|z(t), 0 \le t \le t_1\rangle = \int_0^{t_1} \alpha(t,\tau) z(\tau)\, d\tau. \tag{4.148}$$

4.8.1.1 Proof for the Discrete Case Let the probability density

$$p[x_{k_2}|z_{k_1}] \tag{4.149}$$

be Gaussian and let $\alpha_1, \alpha_2, \ldots, \alpha_{k_1}$ satisfy

$$\mathrm{E}\left\langle \left[x_{k_2} - \sum_{i=1}^{k_1} \alpha_i z_i \right] z_j^{\mathrm{T}} \right\rangle = 0, \quad j = 1, \ldots, k_1, \tag{4.150}$$

and

$$k_1 < k_2, \quad k_1 = k_2, \quad \text{or} \quad k_1 > k_2. \tag{4.151}$$

The existence of vectors α_i satisfying this equation is guaranteed because the covariance $[z_i, z_j]$ is nonsingular.

The vectors

$$\left[x_{k_2} - \sum \alpha_i z_i\right] \tag{4.152}$$

and z_i are independent. Then it follows from the zero-mean property of the sequence x_k that

$$\begin{aligned} \mathrm{E}\left\langle \left[x_{k_2} - \sum_{i=1}^{k_1} \alpha_i z_i\right] \middle| z_1, \ldots, z_{k_1} \right\rangle &= \mathrm{E}\left\langle x_{k_2} - \sum_{i+1}^{k_1} \alpha_i z_i \right\rangle \\ &= 0, \end{aligned}$$

$$\mathrm{E}\left\langle x_{k_2} | z_1, z_2, \ldots, z_{k_1} \right\rangle = \sum_{i=1}^{k_1} \alpha_i z_i.$$

The proof of the continuous case is similar.

The linear minimum variance estimator is unbiased, that is,

$$\mathrm{E}\left\langle x(t) - \hat{x}(t) \right\rangle = 0, \tag{4.153}$$

where

$$\hat{x}(t) = \mathrm{E}\left\langle x(t) | z(t) \right\rangle. \tag{4.154}$$

In other words, an unbiased estimate is one whose expected value is the same as that of the quantity being estimated.

4.8.2 Orthogonality Principle

The nonlinear solution $\mathrm{E}\langle x|z\rangle$ of the estimation problem is not simple to evaluate. If x and z are jointly normal, then $\mathrm{E}\langle x|z\rangle = \alpha_1 z + \alpha_0$.

Let x and z be scalars and M be a 1×1 weighting matrix. The constants α_0 and α_1 that minimize the *mean-squared* (MS) *error*

$$e = \mathrm{E}\left\langle [x - (\alpha_0 + \alpha_1 z)]^2 \right\rangle = \int_{\infty}^{\infty} \int_{\infty}^{\infty} [x - (\alpha_0 + \alpha_1 z)]^2 p(x, z)\, dx\, dz \tag{4.155}$$

are given by

$$\begin{aligned} \alpha_1 &= \frac{r\sigma_x}{\sigma_z}, \\ \alpha_0 &= \mathrm{E}\left\langle x \right\rangle - \alpha_1 \mathrm{E}\left\langle z \right\rangle, \end{aligned}$$

and the resulting minimum MS error $e_{\min}$ is

$$e_{\min} = \sigma_x^2 (1 - r^2), \tag{4.156}$$

where the ratio

$$r = \frac{\mathrm{E}\,\langle xz\rangle}{\sigma_x \sigma_z} \tag{4.157}$$

is called the *correlation coefficient* of x and z, and σ_x, σ_z are standard deviations of x and z, respectively.

Suppose α_1 is specified. Then

$$\frac{d}{d\alpha_0}\mathrm{E}\,\langle [x - \alpha_0 - \alpha_1 z]^2\rangle = 0 \tag{4.158}$$

and

$$\alpha_0 = \mathrm{E}\,\langle x\rangle - \alpha_1 \mathrm{E}\,\langle z\rangle. \tag{4.159}$$

Substituting the value of α_0 in $\mathrm{E}\,\langle [x - \alpha_0 - \alpha_1 z]^2\rangle$ yields

$$\begin{aligned}
\mathrm{E}\,\langle [x - \alpha_0 - \alpha_1 z]^2\rangle &= \mathrm{E}\,\langle [x - \mathrm{E}\,\langle x\rangle - \alpha_1 (z - \mathrm{E}\,\langle z\rangle)]^2\rangle \\
&= \mathrm{E}\,\langle [(x - \mathrm{E}\,\langle x\rangle) - \alpha_1 (z - \mathrm{E}\,\langle z\rangle)]^2\rangle \\
&= \mathrm{E}\,\langle [x - \mathrm{E}\,\langle x\rangle]^2\rangle + \alpha_1^2 \mathrm{E}\,\langle [z - \mathrm{E}\,\langle z\rangle]^2\rangle \\
&\quad - 2\alpha_1 \mathrm{E}\,\langle (x - \mathrm{E}\,\langle x\rangle)(z - \mathrm{E}\,\langle z\rangle)\rangle,
\end{aligned}$$

and differentiating with respect to α_1 as

$$\begin{aligned}
0 &= \frac{d}{d\alpha_1}\mathrm{E}\,\langle [x - \alpha_0 - \alpha_1 z]^2\rangle \\
&= 2\alpha_1 \mathrm{E}\,\langle (z - \mathrm{E}\,\langle z\rangle)^2\rangle - 2\mathrm{E}\,\langle (x - \mathrm{E}\,\langle x\rangle)(z - \mathrm{E}\,\langle z\rangle)\rangle, \qquad (4.160) \\
\alpha_1 &= \frac{\mathrm{E}\,\langle (x - \mathrm{E}\,\langle x\rangle)(z - \mathrm{E}\,\langle z\rangle)\rangle}{\mathrm{E}\,\langle (z - \mathrm{E}\,\langle z\rangle)^2\rangle} \\
&= \frac{r\sigma_x \sigma_z}{\sigma_z^2} \\
&= \frac{r\sigma_x}{\sigma_z}, \qquad (4.161) \\
e_{\min} &= \sigma_x^2 - 2r^2\sigma_x^2 + r^2\sigma_x^2 \\
&= \sigma_x^2 (1 - r^2).
\end{aligned}$$

Note that if one assumes that x and z have zero means,

$$\mathrm{E}\,\langle x\rangle = \mathrm{E}\,\langle z\rangle = 0, \tag{4.162}$$

then we have the solution

$$\alpha_0 = 0. \tag{4.163}$$

Orthogonality Principle The constant α_1 that minimizes the MS error

$$e = \mathrm{E}\,\langle [x - \alpha_1 z]^2 \rangle \tag{4.164}$$

is such that $x - \alpha_1 z$ is *orthogonal* to z. That is,

$$\mathrm{E}\,\langle [x - \alpha_1 z] z \rangle = 0, \tag{4.165}$$

and the value of the minimum MS error is given by the formula

$$e_m = \mathrm{E}\,\langle (x - \alpha_1 z) x \rangle. \tag{4.166}$$

4.8.3 A Geometric Interpretation of Orthogonality

Consider all RVs as vectors in abstract vector spaces. The inner product of x and z is taken as the second moment $\mathrm{E}\,\langle xz \rangle$. Thus

$$\mathrm{E}\,\langle x^2 \rangle = \mathrm{E}\,\langle x^{\mathrm{T}} x \rangle \tag{4.167}$$

is the square of the length of x. The vectors $x, z, \alpha_1 z$, and $x - \alpha_1 z$ are as shown in Figure 4.7.

The MS error $\mathrm{E}\,\langle (x - \alpha_1 z)^2 \rangle$ is the square of the length of $x - \alpha_1 z$. This length is minimum if $x - \alpha_1 z$ is orthogonal (perpendicular) to z,

$$\mathrm{E}\,\langle (x - \alpha_1 z) z \rangle = 0. \tag{4.168}$$

We will apply the orthogonality principle to derive Kalman estimators in Chapter 5.

4.9 SUMMARY

4.9.1 Important Points to Remember

Events Form a Sigma Algebra of Outcomes of an Experiment. A *statistical experiment* is an undertaking with an uncertain outcome. The set of *all* possible outcomes of an

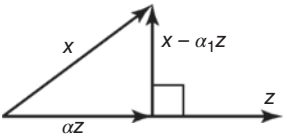


Figure 4.7 Orthogonality diagram.

experiment is called a *sample space*. An event is said to *occur* if the outcome of an experiment is an element of the event.

Independent Events. A collection of events is called *mutually independent* if the occurrence or nonoccurrence of any finite number of them has no influence on the possibilities for occurrence or nonoccurrence of the others.

Random Variables Are Functions. A *scalar RV* is a real-valued function defined on the sample space of a probability space such that, for every open interval (a, b), $-\infty < a \leq b < +\infty$, the set

$$f^{-1}((a, b)) = \{s \in S | a < f(s) < b\}$$

is an event (i.e., is in the sigma algebra of events). A vector-valued RV has scalar RVs as its components. An RV is also called a *variate*.

Random processes (RPs) are functions of time with RVs as their values. A *process* is the evolution over time of a system. If the future state of the system can be predicted from its initial state and its inputs, then the process is considered *deterministic*. Otherwise, it is called *nondeterministic*. If the possible states of a nondeterministic system at any time can be represented by an RV, then the evolution of the state of the system is an RP, or a *stochastic process*. Formally, an RP or a stochastic process is a function f defined on a time interval with *RVs* as its values $f(t)$.

An RP is called

A *Bernoulli process*, or *independent, identically distributed (i.i.d.)* process if the probability distribution of its values at any time is independent of its values at any other time.

A *Markov process* if, for any time t, the probability distribution of its state at any time $\tau > t$, given its state at time t, is the same as its probability distribution given its state at all times $s \leq t$.

A *Gaussian process* if the probability distribution of its possible values at any time is a Gaussian distribution.

Stationary if certain statistics of its probability distributions are invariant under shifts of the time origin. If only its first and second moments are invariant, it is called *wide-sense stationary* (WSS) or *weak-sense stationary*. If all its statistics are invariant, it is called *strict-sense stationary*.

Ergodic if the probability distribution of its values at any one time, over the ensemble of sample functions, equals the probability distribution over all time of the values of randomly chosen member functions.

Orthogonal to another RP if the expected value of their pointwise product is zero.

4.9.2 Important Equations to Remember

The density function of an n-vector-valued (or *multivariate*) *Gaussian probability distribution* $\mathcal{N}(\overline{x}, P)$ has the functional form

$$p(x) = \frac{1}{\sqrt{(2\pi)^n \det P}} e^{-(1/2)(x-\overline{x})^{\mathrm{T}} P^{-1} (x-\overline{x})},$$

where $\bar{x}$ is the *mean* of the distribution and P is the covariance matrix of deviations from the mean.

A *linear stochastic process in continuous time* with state x and state covariance P has the model equations

$$\begin{aligned}\dot{x}(t) &= F(t)x(t) + G(t)w(t),\\ z(t) &= H(t)x(t) + v(t),\\ \dot{P}(t) &= F(t)P(t) + P(t)F^{\mathrm{T}}(t) + G(t)Q(t)G^{\mathrm{T}}(t),\end{aligned}$$

where $Q(t)$ is the covariance of zero-mean *plant noise* $w(t)$. A *discrete-time linear stochastic process* has the model equations

$$\begin{aligned}x_k &= \Phi_{k-1}x_{k-1} + G_{k-1}\, w_{k-1},\\ z_k &= H_k x_k + v_k,\\ P_k &= \Phi_{k-1}P_{k-1}\Phi_{k-1}^{\mathrm{T}} + G_{k-1}Q_{k-1}G_{k-1}^{\mathrm{T}},\end{aligned}$$

where x is the system state, z is the system output, w is the zero-mean uncorrelated *plant noise*, Q_{k-1} is its covariance of w_{k-1}, and v is the zero-mean uncorrelated *measurement noise*. Plant noise is also called *process noise*. These models may also have known inputs. *Shaping filters* are models of these types that are used to represent RPs with certain types of spectral properties or temporal correlations.

PROBLEMS

4.1 For the discrete-time model

$$x_k = m^2 x_{k-1} + w_{k-1},\ Ew_{k-1} = 0,\ Q_{k-1} = \Delta(k-k),\ P_0 = 1, \text{ and } \mathop{\mathrm{E}}_{x}\langle x_0\rangle = 1,$$

find $E\langle x_3\rangle$, P_k, and P_∞.

4.2 An amplitude-modulated signal is specified by

$$y(t) = [1 + mx(t)]\cos(\Omega t + \lambda).$$

Here $X(t)$ is a WSS RP independent of λ, which is an RV uniformly distributed over $[0, 2\pi]$. We are given that

$$\psi_x(\tau) = \frac{1}{\tau^2 + 1}.$$

(a) Verify that $\psi_x(\tau)$ is an autocorrelation.

(b) Let *x(t)* have the autocorrelation given above. Using the direct method for computing the spectral density, calculate Ψ_y.

4.3 Let $R(T)$ be an arbitrary autocorrelation function for an mean-square continuous stochastic process $X(t)$ and let $\Psi(\omega)$ be the PSD for the process $X(t)$. Is it true that

$$\lim_{|\omega|\to\infty} \Psi(\omega) = 0?$$

Justify your answer.

4.4 Find the state-space models for longitudinal, vertical, and lateral turbulence for the following PSD of the "Dryden" turbulence model:

$$\Psi(\omega) = \sigma^2 \left(\frac{2L}{\pi V}\right)\left(\frac{1}{1+(L\omega/V)^2}\right),$$

where

ω = frequency in radians per second,

σ = root-mean-square (RMS) turbulence intensity,

L = scale length in feet, and

V = airplane velocity in feet per second (290 ft/s)

(a) For longitudinal turbulence,

$$L = 600 \text{ ft}$$

$$\sigma_u = 0.15 \text{ mean head wind or tail wind (knots).}$$

(b) For vertical turbulence,

$$L = 300 \text{ ft}$$

$$\sigma_w = 1.5 \text{ knots.}$$

(c) For lateral turbulence,

$$L = 600 \text{ ft}$$

$$\sigma_v = 0.15 \text{ mean cross-wind (knots).}$$

4.5 Consider the RP

$$x(t) = \cos(\omega_0 t + \theta_1)\cos(\omega_0 t + \theta_2),$$

where θ_1 and θ_2 are independent RVs uniformly distributed between 0 and 2π.

(a) Show that $X(t)$ is WSS.

(b) Calculate $\psi_x(\tau)$ and $\Psi_x(\omega)$.

(c) Discuss the ergodicity of $X(t)$.

4.6 Let $\psi_x(\tau)$ be the autocorrelation of a WSS RP. Is the real part of $\psi_x(\tau)$ necessarily also an autocorrelation? If your answer is affirmative, prove it; if negative, give a counterexample.

4.7 Assume $X(t)$ is WSS

$$y(t) = x(t)\cos(\omega t + \theta),$$

where ω is a constant and θ is a uniformly distributed $[0, 2\pi]$ random phase. Find $\psi_{xy}(\tau)$.

4.8 The RP $X(t)$ has mean zero and autocorrelation function

$$\psi_x(\tau) = e^{-|\tau|}.$$

Find the autocorrelation function for

$$y(t) = \int_0^t x(u)\, du, \quad t > 0.$$

4.9 Assume $X(t)$ is WSS with PSD

$$\Psi_x(\omega) = \begin{cases} 1, & -a \le \omega \le a, \\ 0, & \text{otherwise.} \end{cases}$$

Sketch the spectral density of the process

$$y(t) = x(t)\cos(\Omega t + \theta),$$

where θ is a uniformly distributed random phase and $\Omega > a$.

3.19

(a) Define a WSS RP.

(b) Define a strict-sense stationary RP.

(c) Define a realizable linear system.

(d) Is the following an autocorrelation function?

$$\psi(\tau) = \begin{cases} 1 - |\tau|, & |\tau| < 1, \\ 0 & \text{otherwise.} \end{cases}$$

Explain.

4.10 Assume $X(t)$ is a stationary RP with autocorrelation function

$$\psi_x(\tau) = \begin{cases} 1 - |\tau|, & -1 \le \tau \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Find the spectral density $\Psi_y(\omega)$ for

$$y(t) = x(t)\cos(\omega_0 t + \lambda)$$

when ω_0 is a constant and λ is an RV uniformly distributed on the interval $[0, 2\pi]$.

4.11 An RP $X(t)$ is defined by

$$x(t) = \cos(t + \theta),$$

where θ is an RV uniformly distributed on the interval $[0, 2\pi]$. Calculate the autocorrelation function $\psi_y(t, s)$ for

$$y(t) = \int_0^t x(u)\, du.$$

4.12 Let ψ_1 and ψ_2 be two arbitrary continuous, absolutely integrable autocorrelation functions. Are the following necessarily autocorrelation functions? Briefly explain your answer.

(a) $\psi_1 \cdot \psi_2$.

(b) $\psi_1 + \psi_2$.

(c) $\psi_1 - \psi_2$.

(d) $\psi_1 * \psi_2$(the convolution of $\psi_1$with ψ_2).

4.13 Give a short reason for each answer:

(a) If $f(t)$ and $g(t)$ are autocorrelation functions, $f^2(t) + g(t)$ is (necessarily, perhaps, never) an autocorrelation function.

(b) As in (a), $f^2(t) - g(t)$ is (necessarily, perhaps, never) an autocorrelation function.

(c) If $X(t)$ is a strictly stationary process, $x^2(t) + 2x(t-1)$ is (necessarily, perhaps, never) strictly stationary.

(d) The function

$$\omega(\tau) = \begin{cases} \cos\tau, & -\frac{9}{2}\pi \le \tau \le \frac{9}{2}\pi, \\ 0 & \text{otherwise}, \end{cases}$$

(is, is not) an autocorrelation function.

(e) Let $X(t)$ be strictly stationary and ergodic and α be a Gaussian RV with mean 0 and variance 1 and α is independent of $X(t)$. Then $y(t) = \alpha x(t)$ is (necessarily, perhaps, never) ergodic.

4.14 Which of the following functions is an autocorrelation function of a WSS process? Give a brief reason for each answer.

(a) $e^{-|\tau|}$.

(b) $e^{-|\tau|}\cos\tau$.

(c) $\Gamma(t) = \begin{cases} 1, & |t| < a, \\ 0 & |t| \geq a. \end{cases}$

(d) $e^{-|\tau|} \sin \tau$.

(e) $\frac{3}{2}e^{-|\tau|} - e^{-2|\tau|}$.

(f) $2e^{-2|\tau|} - e^{-|\tau|}$.

4.15 Is the following function an autocorrelation function of a WSS process?

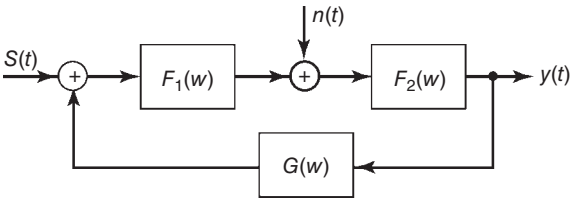
$$\psi_x(\tau) = 1.5e^{-|\tau|} + (11/3)e^{-3|\tau|}.$$

4.16 Discuss each of the following:

(a) The distinction between stationarity and wide-sense stationarity.

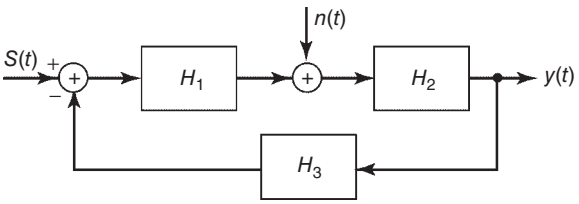
(b) The periodic character of the cross-correlation function of two processes that are themselves periodic with periods mT and nT, respectively.

4.17 A system transfer function can sometimes be experimentally determined by injecting white noise $w(t)$ and measuring the cross-correlation between the system output and the white noise. Here we consider the following system:



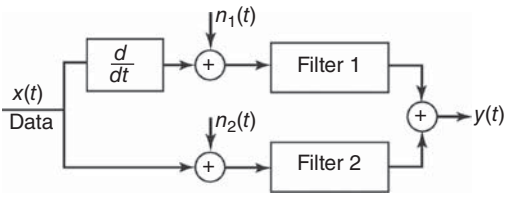
We assume $\Psi_s(\omega)$ known, $S(t)$ and $w(t)$ independent, and $\Psi_n(\omega) = 1$. Find $\Psi_{yn}(\omega)$. *Hint:* Write $y(t) = y_S(t) + y_n(t)$, where y_S and y_n are the parts of the output due to S and n, respectively.

4.18 Let $S(t)$ and $w(t)$ be real stationary uncorrelated RPs, each with mean zero.



Here, $H_1(j2\pi\omega), H_2(j2\pi\omega)$, and $H_3(j2\pi\omega)$ are transfer functions of time-invariant linear systems and $S_0(t)$ is the output when $w(t)$ is zero and $n_0(t)$ is the output when $S(t)$ is zero. Find the output signal-to-noise ratio, defined as $\mathrm{E}\,\langle S_0^2(t)\rangle / \mathrm{E}\,\langle n_0^2(t)\rangle$.

4.19 A single random data source is measured by two different transducers, and their outputs are suitably combined into a final measurement $y(t)$. The system is as pictured below:



Assume that $n_1(t)$ and $n_2(t)$ are uncorrelated RPs, data and noises are uncorrelated, filter 1 has transfer function $Y(s)/s$, and filter 2 has transfer function $1 - Y(s)$. Suppose that it is desired to determine the MS error of measurement, where the error is defined by $e(t) = x(t) - y(t)$. Calculate the mean-square value of the error in terms of $Y(s)$ and the spectral densities Ψ_x, Ψ_{n_1}, and Ψ_{n_2}.

4.20 Let $X(t)$ be the solution of

$$\dot{x} + x = w(t),$$

where $w(t)$ is white noise with spectral density 2π.

(a) Assuming that the above system has been operating since $t = -\infty$, find $\psi_x(t_1 t_2)$. Investigate whether $X(t)$ is WSS, and if so, express ψ_x accordingly.

(b) Instead of the system in (a), consider

$$\dot{x} + x = \begin{cases} w(t), & t \geq 0, \\ 0, & t < 0, \end{cases}$$

where $x(0) = 0$. Again, compute $\psi_x(t_1, t_2)$.

(c) Let $y(t) = \int_0^t x(\tau) d\tau$. Find $\psi_{xy}(t_1, t_2)$ for both of the systems described in (a) and (b).

(d) It is desired to predict $x(t + \alpha)$ from $x(t)$, that is, a future value of the process from its present value. A possible predictor $\hat{x}(t + \alpha)$ is of the form

$$\hat{x}(t + \alpha) = ax(t).$$

Find that a which will give the smallest mean-square prediction error, that is, that minimizes

$$\mathrm{E}\,\langle |\hat{x}(t + \alpha) - x(t + \alpha)|^2 \rangle,$$

where $x(t)$ is as in part *(a)*.

4.21 Let *x(t)* be the solution of

$$\dot{x} + x = w(t)$$

with initial condition $x(0) = x_0$. It is assumed that $w(t)$ is white noise with spectral density 2π and is turned on at $t = 0$. The initial condition x_0 is an RV independent of $w(t)$ and with zero mean.

(a) If x_0 has variance σ^2, what is $\psi_x(t_1, t_2)$? Derive the result.

(b) Find that value of σ (call it σ_0) for which $\psi_x(t_1, t_2)$ is the same for all $t \geq 0$. Determine whether, with $\sigma = \sigma_0$, $\psi_x(t_1, t_2)$ is a function only of $t_1 - t_2$.

(c) If the white noise had been turned on at $t = -\infty$ and the initial condition has zero mean and variance σ_0^2 as above, is *x(t)* WSS? Justify your answer by appropriate reasoning and/or computation.

4.22 Let

$$\dot{x}(t) = F(t)x(t) + w(t),$$
$$x(a) = x_a, t \geq a,$$

where x_a is a zero-mean RV with covariance matrix P_a and

$$\mathrm{E}\,\langle w(t)\rangle = 0 \quad \forall t,$$
$$\mathrm{E}\,\langle w(t)w^{\mathrm{T}}(s)\rangle = Q(t)\delta(t-s) \quad \forall t, s$$
$$\mathrm{E}\,\langle x(a)w^{\mathrm{T}}(t)\rangle = 0 \quad \forall t.$$

(a) Determine the mean m(t) and covariance P(t, t) for the process x(t).

(b) Derive a differential equation for P(t, t).

4.23 Find the covariance matrix P(t) and its steady-state value $P(\infty)$ for the following continuous systems:

(a) $\dot{x} = \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix} x + \begin{bmatrix} 1 \\ 1, \end{bmatrix} w(t), \quad P(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1, \end{bmatrix},$

(b) $\dot{x} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} x + \begin{bmatrix} 5 \\ 1 \end{bmatrix} w(t), \quad P(0) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,where $w \in \mathcal{N}(0, 1)$ and white.

4.24 For the continuous-time system

$$\dot{x}(t) = -x(t) + w(t)$$
$$x(0) = 1$$
$$P(0) = 1$$
$$w(t) = 2$$
$$Q(t) = e^{-2}\delta(t - \tau).$$

Find $\mathrm{E}\langle x(t)\rangle$, $P_x(t)$, P_∞.

4.25 Find the covariance matrix P_k and its steady-state value P_∞ for the following discrete system:

$$x_{k+1} = \begin{bmatrix} 0 & \frac{1}{2} \\ -\frac{1}{2} & 2 \end{bmatrix} x_k + \begin{bmatrix} 1 \\ 1 \end{bmatrix} w_k, \quad P_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

where $w_k \in N(0, 1)$ and white.

4.26 Find the steady-state covariance for the state-space model given in Example 3.4.

4.27 Show that the continuous-time steady-state algebraic equation

$$0 = FP(\infty) + P(\infty)F^{\mathrm{T}} + GQG^{\mathrm{T}}$$

has no nonnegative solution for the scalar case with $F = Q = G = 1$.

4.28 Show that the discrete-time steady-state algebraic equation

$$P_\infty = \Phi P_\infty \Phi^{\mathrm{T}} + Q$$

has no solution for the scalar case with $\Phi = Q = 1$.

4.29 Find the covariance of x_k as a function of k and its steady-state value for the system

$$x_k = -2x_{k-1} + w_{k-1},$$

where $\mathrm{E}w_{k-1} = 0$ and $\mathrm{E}(w_k w_j) = e^{-|k-j|}$. Assume the initial value of the covariance (P_0) is 1.

4.30 Find the covariance of *x(t)* as a function of t and its steady-state value for the system

$$\dot{x}(t) = -2x(t) + w(t),$$

where $\mathrm{E}w(t) = 0$ and $\mathrm{E}(w(t_1) \quad w(t_2)) = e^{-|t_1 - t_2|}$. Assume the initial value of the covariance (P_0) is 1.

4.31 Suppose that *x(t)* has autocorrelation function $\psi_x(\tau) = e^{-c|\tau|}$. It is desired to predict $x(t + \alpha)$ on the basis of the past and present of *x(t)*, that is, the predictor may use *x(s)* for all $s \leq t$.

(a) Show that the minimum mean-square error linear prediction is

$$\hat{x}(t + \alpha) = e^{-c\alpha} x(t).$$

(b) Find the MS error corresponding to the above. *Hint:* Use the orthogonality principle.

REFERENCES

[1] E. Allen, *Modeling with Itô Stochastic Differential Equations*, Springer, Berlin, 2007.

[2] L. Arnold, *Stochastic Differential Equations: Theory and Applications*, Kreiger, Malabar, FL, 1992.

[3] J. Baras and V. Mirelli, *Recent Advances in Stochastic Calculus*, Springer-Verlag, New York, 1990.

[4] K. Itô and H. P. McKean Jr., *Diffusion Processes and Their Sample Paths*, Academic Press, New York, 1965.

[5] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications*, Springer, Berlin, 2007.

[6] K. Sobczyk, *Stochastic Differential Equations with Applications to Physics and Engineering*, Kluwer Academic Press, Dordrecht, The Netherlands, 1991.

[7] R. L. Stratonovich, *Topics in the Theory of Random Noise* (R. A. Silverman, Ed.), Gordon & Breach, New York, 1963.

[8] A. Einstein, "Über die von molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen," *Annelen der Physik*, Vol. 17, pp. 549–560, 1905.

[9] P. Langevin, "Sur la théory du muovement brownien," *Comptes Rendus de l'Academie des Sciences*, Vol. 146, pp. 530–533, Paris, 1908.

[10] R. S. Bucy and P. D. Joseph, *Filtering for Stochastic Processes with Applications to Guidance*, American Mathematical Society, Chelsea Publishing, Providence, RI, 2005.

[11] K. Itô, *Lectures on Stochastic Processes*, Tata Institute of Fundamental Research, Bombay (Mumbai), 1961.

[12] H. W. Bode and C. E. Shannon, "A simplified derivation of linear least-squares smoothing and prediction," *IRE Proceedings*, Vol. 48, pp. 417–425, 1950.

[13] L. A. Zadeh and J. R. Ragazzini, "An extension of Wiener's theory of prediction," *Journal of Applied Physics*, Vol. 21, pp. 645–655, 1950.

[14] C. F. Van Loan, "Computing integrals involving the matrix exponential," *IEEE Transactions on Automatic Control*, Vol. AC-23, No. 3, pp. 395–404, 1978.