

## CHAPTER 6

# Stationary Stochastic Processes

### 6.1. Weak Definition of a Stochastic Process

This chapter is devoted to further topics in the theory of stochastic processes and their applications. We start with a weaker definition of a stochastic process that is sufficient in the study of stationary processes. We said before that a stochastic process is a function  $u$  of both a variable  $\omega$  in a probability space and a continuous parameter  $t$ , making  $u$  a random variable for each  $t$  and a function of  $t$  for each  $\omega$ . We made statements about the kind of function of  $t$  that was obtained for each  $\omega$ . The definition here is less specific about what happens for each  $\omega$ .

Consider a collection of random variables  $u(t, \omega) \in \mathbb{C}$  parametrized by  $t$ .

**DEFINITION.** We say that  $u(t, \omega)$  is a real-valued stochastic process if for every finite set of points  $t_1, \dots, t_n$ , the joint distribution of  $u(t_1, \omega), \dots, u(t_n, \omega)$  is known:

$$F_{t_1, \dots, t_n}(y_1, \dots, y_n) = P(u(t_1) \leq y_1, \dots, u(t_n) \leq y_n).$$

The family of functions  $F_{t_1, \dots, t_n}(y_1, \dots, y_n)$  must satisfy some natural requirements:

1.  $F \geq 0$ .
2.  $F(\infty, \dots, \infty) = 1$  and  $F(-\infty, \dots, -\infty) = 0$ .
3.  $F_{t_1, \dots, t_n}(y_1, \dots, y_m, \infty, \dots, \infty) = F_{t_1, \dots, t_m}(y_1, \dots, y_m)$ .
4. If  $(i_1, \dots, i_n)$  is a permutation of  $(1, \dots, n)$ , then

$$F_{t_{i_1}, \dots, t_{i_n}}(y_{i_1}, \dots, y_{i_n}) = F_{t_1, \dots, t_n}(y_1, \dots, y_n).$$

This definition has a natural extension to complex-valued processes, in which one assumes that one knows the joint distribution of the real and complex parts of  $u$ .

A moment of  $u(t, \omega)$  of order  $q$  is an object of the form

$$M_{i_1, \dots, i_n} = E[u^{i_1}(t_1) \cdots u^{i_n}(t_n)], \quad \sum_{j=1}^n i_j = q.$$

If a stochastic process has finite moments of order  $q$ , it is a process of order  $q$ . The moment

$$E[u(t, \omega)] = m(t)$$

is the mean of  $u$  at  $t$ . The function

$$E \left[ (u(t_1, \omega) - m(t_1)) \overline{(u(t_2, \omega) - m(t_2))} \right] = R(t_1, t_2)$$

is the covariance of  $u$ . Let us list the properties of the covariance of  $u$ :

1.  $R(t_1, t_2) = \overline{R(t_2, t_1)}$ .
2.  $R(t_1, t_1) \geq 0$ .
3.  $|R(t_1, t_2)| \leq \sqrt{R(t_1, t_1)R(t_2, t_2)}$ .
4. For all  $t_1, \dots, t_n$  and all  $z_1, \dots, z_n \in \mathbb{C}$ ,

$$\sum_{i=1}^n \sum_{j=1}^n R(t_i, t_j) z_i \overline{z_j} \geq 0.$$

The first three properties are easy to establish; the fourth is proved as follows: For any choice of complex numbers  $z_j$ , the sum

$$\sum_{i=1}^n \sum_{j=1}^n R(t_i, t_j) z_i \overline{z_j}$$

is by definition equal to

$$E \left[ \left| \sum_{j=1}^n (u(t_j) - m(t_j)) z_j \right|^2 \right] \geq 0$$

(i.e., to the expected value of a nonnegative quantity).

**DEFINITION.** A process is stationary in the strict sense if for every  $t_1, \dots, t_n$  and  $T \in \mathbb{R}$ ,

$$F_{t_1, \dots, t_n}(y_1, \dots, y_n) = F_{t_1+T, \dots, t_n+T}(y_1, \dots, y_n).$$

For a stochastic process that is stationary in this sense, all moments are constant in time, and in particular,  $m(t) = m$  and  $R(t_1, t_2) =$

$R(t_1+T, t_2+T)$  for all  $T$ . Choose  $T = -t_2$ ; then  $R(t_1, t_2) = R(t_1-t_2, 0)$ , and it becomes reasonable to define

$$R(t_1 - t_2) = R(t_1, t_2),$$

where the function  $R$  on the left side, which has only one argument, is also called  $R$  in the hope that there is no ambiguity. Note that  $R(T) = R(t+T, t)$ .

The above properties become, for the new function  $R$ ,

1.  $R(t) = \overline{R(-t)}$ .
2.  $R(0) \geq 0$ .
3.  $|R(t)| \leq R(0)$ .
4. For all  $t_1, \dots, t_n$  and all  $z_1, \dots, z_n \in \mathbb{C}$ ,

$$\sum_i^n \sum_j^n R(t_i - t_j) z_i \overline{z_j} \geq 0. \quad (6.1)$$

DEFINITION. A stochastic process is stationary in the wide sense if it has a constant mean and its covariance depends only on the difference between the arguments, i.e.,

1.  $m(t) = m$ .
2.  $R(t_1, t_2) = R(t_1 - t_2)$ .

If a stochastic process is stationary in the wide sense and Gaussian, then it is stationary in the strict sense (because a Gaussian process is fully determined by its mean and covariances). Brownian motion is not stationary. White noise is stationary (but ill defined without appeal to distributions).

We now consider some instances of processes that are stationary in the wide sense. Pick  $\xi \in \mathbb{C}$  to be a random variable and  $h(t)$  a nonrandom function of time, and consider the process  $u(t, \omega) = \xi h(t)$ . Assume for simplicity that  $h(t)$  is differentiable, and determine when a process of this type is stationary in the wide sense. Its mean is

$$m(t) = E[\xi h(t)] = h(t)E[\xi],$$

which is constant if and only if  $h(t)$  is constant or  $E[\xi] = 0$ . Suppose  $E[\xi] = 0$ . The covariance

$$R(t_1, t_2) = E[\xi h(t_1) \overline{\xi h(t_2)}] = E[\xi \overline{\xi} h(t_1) \overline{h(t_2)}]$$

must depend only on the difference  $t_1 - t_2$ . Consider the special case  $t_1 = t_2 = t$ . In this case, the covariance  $E[\xi \overline{\xi} h(t) \overline{h(t)}]$  must be  $R(0)$ ;

hence  $h(t)\overline{h(t)}$  must be constant. Therefore,  $h(t)$  is of the form

$$h(t) = Ae^{i\phi(t)}, \quad (6.2)$$

where  $A$  is a constant and  $\phi(t)$  a function of  $t$  that remains to be determined. Now we narrow the possibilities some more. Suppose  $A \neq 0$ . Then

$$R(t_1 - t_2) = |A|^2 E[\xi\bar{\xi}]e^{i\phi(t_1) - i\phi(t_2)}.$$

Set  $t_1 - t_2 = T$  and  $t_2 = t$ . Then

$$R(T) = |A|^2 E[\xi\bar{\xi}]e^{i[\phi(t+T) - \phi(t)]}$$

for all  $t, T$ . Since  $R(T) = \overline{R(-T)}$ , we see that

$$\frac{\phi(t+T) - 2\phi(t) + \phi(t-T)}{T^2} = 0.$$

Letting  $T \rightarrow 0$  gives  $\phi''(t) = 0$  for all  $t$ , so  $\phi(t) = \lambda t + \beta$ , where  $\lambda, \beta$  are constants. Also  $e^{i\beta}$  is a constant. We have therefore shown that the process  $u(t, \omega) = \xi h(t)$  is stationary in the wide sense if  $h(t) = Ce^{i\lambda t}$  (where  $C, \lambda$  are constants) and  $E[\xi] = 0$ .

## 6.2. Covariance and Spectrum

In the last section, we presented an example of a stationary stochastic process in the wide sense, given by  $u(t, \omega) = \xi e^{i\lambda t}$ , where  $\xi$  is a random variable with mean 0. This stochastic process has a covariance of the form

$$R(T) = R(t_1, t_2) = R(t_1 - t_2) = E[|\xi|^2]e^{i\lambda T},$$

where  $T = t_1 - t_2$ . Now we want to generalize this example. First, we try to construct a process of the form

$$u(t, \omega) = \xi_1 e^{i\lambda_1 t} + \xi_2 e^{i\lambda_2 t},$$

with  $\lambda_1 \neq \lambda_2$ . Then  $E[u] = E[\xi_1]e^{i\lambda_1 t} + E[\xi_2]e^{i\lambda_2 t}$ , which is independent of  $t$  if  $E[\xi_1] = E[\xi_2] = 0$ . The covariance is

$$\begin{aligned} E [ & (\xi_1 e^{i\lambda_1 t_1} + \xi_2 e^{i\lambda_2 t_1})(\bar{\xi}_1 e^{-i\lambda_1 t_2} + \bar{\xi}_2 e^{-i\lambda_2 t_2}) ] \\ & = E [ |\xi_1|^2 e^{i\lambda_1 T} + |\xi_2|^2 e^{i\lambda_2 T} + \xi_1 \bar{\xi}_2 e^{i\lambda_1 t_2 - i\lambda_2 t_2} + \bar{\xi}_1 \xi_2 e^{i\lambda_1 t_1 - i\lambda_2 t_2} ], \end{aligned}$$

which can be stationary only if  $E[\xi_1 \bar{\xi}_2] = 0$ . Then  $u(t, \omega)$  is stationary and

$$R(T) = E[|\xi_1|^2]e^{i\lambda_1 T} + E[|\xi_2|^2]e^{i\lambda_2 T}.$$

More generally, a process  $u = \sum_j \xi_j e^{i\lambda_j t}$  is stationary in the wide sense if  $E[\xi_j \bar{\xi}_k] = 0$  when  $j \neq k$  and  $E[\xi_i] = 0$ . In this case,

$$R(T) = \sum E[|\xi_j|^2] e^{i\lambda_j T}.$$

This expression can be rewritten in a more useful form as a Stieltjes integral. Recall that when  $q$  is a nondecreasing function of  $x$ , the Stieltjes integral of a function  $h$  with respect to  $q$  is defined to be

$$\int h dq = \lim_{\max\{x_{i+1}-x_i\} \rightarrow 0} \sum h(x_i^*) [q(x_{i+1}) - q(x_i)],$$

where  $x_i \leq x_i^* \leq x_{i+1}$ . If  $q$  is differentiable, then

$$\int_a^b h dq = \int_a^b h q' dx.$$

Suppose  $q(x)$  is the step function

$$q(x) = \begin{cases} 0, & x < c, \\ q_0 & x \geq c, \end{cases}$$

with  $a \leq c \leq b$ . Then  $\int_a^b h dq = h(c)q_0$  if  $h$  is continuous at  $c$ . We define the function  $G = G(k)$  by

$$G(k) = \sum_{\{j|\lambda_j \leq k\}} E[|\xi_j|^2];$$

i.e.,  $G(k)$  is the sum of the expected values of the squares of the amplitudes of the complex exponentials with frequencies less than or equal to  $k$ . Then  $R(T)$  becomes

$$R(T) = \int_{-\infty}^{+\infty} e^{ikT} dG(k).$$

We shall now see that under some technical assumptions, this relation holds for all stochastic processes that are stationary in the wide sense. Indeed, we have the following theorem.

THEOREM 6.1 (Khinchin).

1. If  $R(T)$  is the covariance of a stochastic process  $u(t, \omega)$ , stationary in the wide sense such that

$$\lim_{h \rightarrow 0} E [|u(t+h) - u(t)|^2] = 0,$$

then  $R(T) = \int e^{ikT} dG(k)$  for some nondecreasing function  $G(k)$ .

2. If a function  $R(T)$  can be written as  $\int e^{ikT} dG(k)$  for some nondecreasing function  $G$ , then there exists a stochastic process, stationary in the wide sense, satisfying the condition in part (1) of the theorem, that has  $R(T)$  as its covariance.

Khinchin's theorem follows from the inequalities we have proved for  $R$ ; indeed, one can show (but we will not do so here) that a function that satisfies these inequalities is the Fourier transform of a nonnegative function. If it so happens that  $dG(k) = g(k) dk$ , then  $R(T) = \int e^{ikT} g(k) dk$ , and  $g(k)$  is called the spectral density of the process. Thus, Khinchin's theorem states that the covariance function is a Fourier transform of the spectral density. Hence, if we know  $R(T)$ , we can compute the spectral density by

$$g(k) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-ikT} R(T) dT.$$

For a nonrandom periodic function, one can define an energy per wave number  $k$  as the squared amplitude of the  $k$ th Fourier coefficient; for a nonrandom aperiodic function, one can define the energy per wave number as the squared magnitude of the Fourier transform. The samples of a stationary stochastic process do not have Fourier transforms in the usual sense, because they do not tend to zero at  $\pm\infty$ , but one can still define an average energy per wave number for a stationary stochastic process by the Fourier transform of the covariance.

EXAMPLE. Consider white noise, the derivative (in a sense we have not discussed) of Brownian motion. One can show that  $R(T) = \delta(T)$  (see the exercises). Its spectral density (interpreted carefully) is  $\phi(k) = 1/2\pi$ ; thus, all frequencies have the same amplitude. The adjective "white" comes from the fact that in white light, all frequencies are present with the same amplitude. A stationary random function that is not white noise is called colored noise.

### 6.3. The Inertial Spectrum of Turbulence

To illustrate these constructions, we now derive the spectrum of fully developed turbulence. We do not write down the equations of motion; the only properties of these equations that will be used here are that (a) they are nonlinear, and (b) energy dissipation by viscosity is proportional to an integral over the domain of the sum of the squares of the derivatives of the velocity field (a quantitative description of this property will be given below).

Consider turbulence in a fluid, far from any solid boundaries, with the Reynolds number  $Re = U\ell_0/\nu$  very large, where  $U$  is a typical velocity difference in the flow,  $\ell_0$  is a length scale for the flow, and  $\nu$  is the viscosity; the dimensionless number  $Re$  is large when the velocity differences are large and the viscosity is small, which are the circumstances when turbulence appears;  $U$  is chosen to be a typical velocity difference rather than a typical velocity because a velocity component common to the whole flow field is not relevant when one is studying turbulence. The large scales of turbulent flow are typically driven by large-scale forcing (e.g., in the case of meteorology, by the rotation of the earth around its axis and around the sun); turbulence is characterized by the transfer of energy from large scales to smaller scales at which the energy is dissipated. One usually assumes that as the energy moves to large wave numbers  $k$  (i.e., small scales), the specifics of the forcing are forgotten and the flow can be viewed as approximately homogeneous (translation-invariant) and isotropic (rotation-invariant) at small scales, and that the properties of the flow at small scales are universal (i.e., independent of specific geometry and forcing). One further assumes that the solutions of the equations of fluid mechanics can be viewed as random; how nonrandom equations produce solutions that can be viewed as random is an interesting question that we will not discuss here.

Assume that the velocity field is homogeneous, i.e., statistically translation-invariant in space (not in time, as was implicitly assumed in the previous section through the choice of the letter  $t$  for the parameter). The velocity field in three space dimensions is a vector quantity:  $\mathbf{u} = (u_1, u_2, u_3)$ . Each of these components is a function of the three spatial variables  $x_1, x_2, x_3$ . A Fourier transform in three-dimensional space can be defined and is a function of three Fourier

variables  $k_1, k_2, k_3$  that correspond to each of the spatial variables, and we write  $\mathbf{k} = (k_1, k_2, k_3)$ . One can define a covariance matrix

$$R_{ij}(\mathbf{r}) = E[u_i(\mathbf{x})u_j(\mathbf{x} + \mathbf{r})],$$

where  $\mathbf{r}$  is a three-component vector; then Khinchin's theorem becomes

$$R_{ii}(\mathbf{r}) = \int_{-\infty}^{\infty} e^{i\mathbf{k}\cdot\mathbf{r}} dG_{ii}(\mathbf{k}), \quad (6.3)$$

where  $\mathbf{k} = (k_1, k_2, k_3)$ ,  $\mathbf{k}\cdot\mathbf{r}$  is the ordinary Euclidean inner product, and the functions  $G_{ii}$  are nondecreasing. Without loss of generality in what follows, one can write  $dG_{ii}(\mathbf{k}) = g_{ii}(\mathbf{k}) dk_1 dk_2 dk_3$  (this is so because all we will care about is the dimensions of the various quantities, which are not affected by a possible lack of smoothness). Setting  $\mathbf{r} = 0$  in Eq. (6.3) and summing over  $i$ , we find that

$$E[u_1^2 + u_2^2 + u_3^2] = \int_{-\infty}^{\infty} (g_{11} + g_{22} + g_{33}) dk_1 dk_2 dk_3.$$

We define the left-hand side of this equation to be the specific energy (i.e., energy per unit volume) of the flow and denote it by  $E[u^2]$ . Splitting the integration into an integration in a polar variable  $k$  and integrations over angular variables, one can write

$$E[u^2] = \int_0^{\infty} E(k) dk,$$

with

$$E(k) = \int_{k_1^2 + k_2^2 + k_3^2 = k^2} (g_{11} + g_{22} + g_{33}) dS(\mathbf{k}),$$

where  $dS(\mathbf{k})$  is an element of area on a sphere of radius  $k$ . We define  $E(k)$  to be the *energy spectrum*; it is a function only of  $k = \sqrt{k_1^2 + k_2^2 + k_3^2}$ . The energy spectrum can be thought of as the portion of the energy that can be imputed to motion with wave numbers of magnitude  $k$ .

The kinetic energy of the flow is proportional to the square of the velocity, whereas energy dissipation is proportional to the square of the derivatives of the velocity; in spectral variables (i.e., after Fourier transformation), differentiation becomes multiplication by  $k$ , the Fourier



variable. A calculation, which we skip, shows that  $D$ , the energy dissipation per unit volume  $D$ , can be written as

$$D = \int_0^{\infty} k^2 E(k) dk,$$

where  $E(k)$  is the energy spectrum. This calculation requires some use of the equations of motion, and this is the only place where those equations are made use of in the argument of this section.

It is plausible that when  $Re$  is large, the kinetic energy resides in a range of  $k$ 's disjoint from the range of  $k$ 's where the dissipation is taking place, and indeed, experimental data show it to be so; specifically, there exist wave numbers  $k_1$  and  $k_2$  such that

$$\int_0^{k_1} E(k) dk \sim \int_0^{\infty} E(k) dk, \quad \int_{k_2}^{\infty} k^2 E(k) dk \sim \int_0^{\infty} k^2 E(k) dk,$$

with  $k_1 \ll k_2$ . This observation roughly divides the spectrum into three pieces: (a) the range between 0 and  $k_1$ , the *energy range*, where most of the energy resides; what happens in this range depends on the boundary and initial conditions and must be determined separately for each turbulent flow; (b) the *dissipation range*  $k > k_2$ , where the energy is dissipated; and (c) the intermediate range between  $k_1$  and  $k_2$ ; this range is the conduit through which turbulence moves energy from the energy range to the dissipation range, and it is responsible for the enhanced dissipation produced by turbulence (see Fig. 6.1). One can hope that the properties of turbulence in the intermediate range are *universal*, i.e., independent of the particular flow one is studying. The nonlinearity of the equations couples the energy range to the intermediate range, and if one can find the universal properties of the intermediate range, one can use them to compute in the energy range. We now determine these universal properties.

We will be relying on dimensional analysis (see Chap. 1). The spectrum in the intermediate range  $E(k)$  is a function of  $k$ , the viscosity  $\nu$ , the length scale of the turbulence  $\ell_0$ , the amplitude  $U$  of the typical velocity difference in the flow, and the rate of energy dissipation  $\epsilon$ . This last variable belongs here because energy is transferred from the low- $k$  domain through the intermediate range into the large- $k$  domain, where it is dissipated; the fact that  $\epsilon$  belongs in the list was the brilliant insight of Kolmogorov.

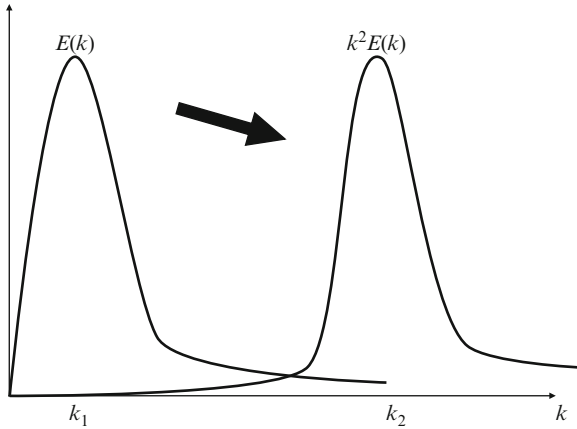


FIGURE 6.1. Sketch of the energy, inertial, and dissipation ranges in turbulence.

Our basic units are the units of length and of time. Suppose the former is reduced by a factor  $L$  and the latter by a factor  $T$ . The dimension of the viscosity is  $L^2/T$ , that of  $\epsilon$  is  $L^2/T^3$ , that of  $k$  is  $1/L$ , and the equation  $E[u^2] = \int E(k) dk$  shows that the dimension of  $E$  is  $L^3/T^2$ . Dimensional analysis yields  $E(k)(\epsilon^{-2/3}k^{5/3}) = \Phi(Re, \ell_0 k)$  for some unknown function  $\Phi$  of the two large arguments  $Re$  and  $\ell_0 k$ ;  $Re$  is large because this is the condition for fully developed turbulence to appear, and  $\ell_0 k$  is large in the intermediate range of scales. If the function  $\Phi$  has a finite nonzero limit  $C$  as its arguments grow (an assumption of complete similarity), one can deduce  $E(k) = C\epsilon^{2/3}k^{-5/3}$ , which is the famous Kolmogorov–Obukhov scaling law for the intermediate range of fully developed turbulence, the cornerstone of turbulence theory. Note that the viscosity has dropped out from this result, leading to the conclusion that the dynamics of the intermediate range are purely *inertial*, i.e., independent of viscosity; this is why the intermediate range is usually called the *inertial range*.

This law is not fully satisfactory for various reasons, and a number of correction schemes have been proposed over the years. In recent years, it has been argued that the unknown function  $\Phi$  behaves, as its arguments tend to infinity, like  $C(Re)(\ell_0 k)^{-d/\log(Re)}\Phi_0(Re, \ell_0 k)$ , where it is  $\Phi_0$  that tends to a nonzero constant as its arguments grow,  $C(Re)$  is a function of  $Re$ , and  $d$  is a positive constant; the exponent  $-d/\log(Re)$

is an anomalous exponent. This is an assumption of incomplete similarity, which leads, for large  $Re$  and  $\ell_0 k$ , to the relation

$$E(k) = C(Re)\epsilon^{2/3}k^{-5/3}(\ell_0 k)^{-d/\log(Re)}.$$

The exponent  $-5/3$  is corrected by the small quantity  $-d/\log(Re)$ ; this quantity is a function of the Reynolds number  $Re$ , but its variation with  $Re$  is slow. However, this correction violates the assumption that the intermediate range is purely inertial. Other proposals for the anomalous exponent, without a dependence on  $Re$ , have also been made.

### 6.4. Time Series

Suppose we are observing a stochastic process  $u(t, \omega)$ , have been observing it long enough to know that it is stationary and to determine its temporal covariances, and suppose we are given observed values  $U(s)$  of  $u(t, \omega)$  for  $s \leq t$  (we denote observed values by capital letters). The question we address in this section is how to predict a value for  $u(t + T, \omega)$  based on the information we have. For simplicity, we shall do so only for a stationary random sequence.

**DEFINITION.** A stationary random sequence is a collection  $u(t, \omega)$  of random variables for  $t = 0, 1, 2, 3, \dots$  as well as for  $t = -1, -2, -3, \dots$  such that the joint distribution of every subset is known, subject to the obvious compatibility conditions, and such that all the distributions are invariant under the transformation  $t \rightarrow t + T$  for  $T$  an integer. Such sequences are also known as *time series*.

Assume  $E[u(t)] = 0$ . The covariance is

$$R(T) = E[u(t + T)\overline{u(t)}],$$

where  $T \in \mathbb{Z}$  satisfies, as before, the following conditions:

1.  $R(0) \geq 0$ .
2.  $|R(T)| \leq R(0)$ .
3.  $R(T) = \overline{R(-T)}$ .
4.  $\sum_{i,j} R(i - j)z_i \overline{z_j} \geq 0$ .

If  $u(t, \omega) = \xi(\omega)h(t)$  is stationary, we can repeat the arguments in Sect. 4.1. Since  $R(0) = E[|u|^2] = E[|\xi|^2]|h(t)|^2$ , we see that  $h(t) = Ae^{i\phi(t)}$  for  $t = 0, \pm 1, \dots$ . Since  $R(1) = \overline{R(-1)}$ , we obtain

$$\phi(t + 1) - \phi(t) = -(\phi(t - 1) - \phi(t)) \pmod{2\pi}$$

for  $t = 0, \pm 1, \dots$ . Setting  $\phi(0) = \alpha$  and  $\phi(0) - \phi(-1) = \lambda$ , we find by induction that  $\phi(t) = \alpha + \lambda t \pmod{2\pi}$ . Consequently,  $h(t) = Ae^{i(\alpha + \lambda t)} = Ce^{i\lambda t}$  for all integers  $t$ , where  $C = Ae^{i\alpha}$  is a possibly complex constant and  $\lambda$  is an integer.

Define a periodic function  $g$  of the argument  $k$  by

$$g(k) = \frac{1}{2\pi} \sum_{T=-\infty}^{+\infty} R(T)e^{-iT k},$$

where  $T$  takes on integer values. Note that if  $R(T)$  does not converge rapidly enough to 0 as  $|T|$  increases,  $g$  may not be smooth. Then  $R(T) = \int_{-\pi}^{\pi} e^{iT k} g(k) dk$ . (The factor  $2\pi$  of Fourier theory is broken up here differently from how we did it before.)

One can show that if  $R(T)$  is a covariance for a time series, then  $g \geq 0$ . Conversely, if  $R(T)$  is given for all integers  $T$ , and if  $\frac{1}{2\pi} \sum_T R(T)e^{-iT k} \geq 0$ , then there exists a time series for which  $R(T)$  is the covariance. This is Khinchin's theorem for a time series.

Consider the problem of finding an estimate for  $u(t + m, \omega)$  when one has values  $u(t - n), u(t - (n - 1)), \dots, u(t - 1)$ . Nothing is assumed here about the mechanism that produces these values; all we are going to use is the assumed fact that the time series is stationary, and that we know the covariance. If the covariance vanishes whenever  $T \neq 0$ , then the  $u(t)$  are uncorrelated, and no useful prediction can be made. We would like to find a random variable  $\hat{u}(t + m, \omega)$  with  $m = 0, 1, 2, \dots$  such that

$$E [ |u(t + m, \omega) - \hat{u}(t + m, \omega)|^2 ]$$

is as small as possible. We know from Sect. 2.3 that

$$\hat{u}(t + m, \omega) = E[u(t + m, \omega) | u(t - 1), u(t - 2), \dots, u(t - n)].$$

The way to evaluate  $\hat{u}$  is to find a basis  $\{\phi_i\}$  in the space of functions of  $\{u(t - n), \dots, u(t - 1)\}$ , expand  $\hat{u}$  in this basis, i.e.,

$$\hat{u} = \sum_{j=1}^n a_j \phi_j(u(t - 1), \dots, u(t - n)),$$

and calculate the coefficients  $a_j$  of the expansion. This is hard in general. We simplify the problem by looking only for the best

approximation in the span of  $\{u(t-1), \dots, u(t-n)\}$ , i.e., we look for a random variable

$$\hat{u}(t+m, \omega) = \sum_{j=1}^n a_j u(t-j, \omega).$$

This is called *linear prediction*. The span  $L$  of the  $u(t-j, \omega)$  is a closed linear space; therefore, the best linear prediction minimizes

$$E[|u(t+m, \omega) - \hat{u}(t+m, \omega)|^2]$$

for  $\hat{u}$  in  $L$ . What we have to do is to find  $\{a_j\}_{j=1}^n$  such that

$$E \left[ \left| u(t+m, \omega) - \sum_{j=1}^n a_j u(t-j, \omega) \right|^2 \right]$$

is as small as possible. We have

$$\begin{aligned} & E[|u - \hat{u}|^2] \\ &= E \left[ \left( u(t+m) - \sum_j a_j u(t-j) \right) \overline{\left( u(t+m) - \sum_l a_l u(t-l) \right)} \right] \\ &= E \left[ u(t+m) \overline{u(t+m)} - \sum_l \bar{a}_l u(t+m) \overline{u(t-l)} \right. \\ &\quad \left. - \sum_j a_j \overline{u(t+m)} u(t-j) + \sum_j \sum_l a_j \bar{a}_l u(t-j) \overline{u(t-l)} \right] \\ &= R(0) - 2\operatorname{Re} \left( \sum_j \bar{a}_j R(m+j) \right) + \sum_j \sum_l a_j \bar{a}_l R(l-j), \end{aligned}$$

which is minimized when

$$\frac{1}{2} \frac{\partial E[|u - \hat{u}|^2]}{\partial \bar{a}_j} = -R(m+j) + \sum_{l=1}^n a_l R(j-l) = 0 \quad (6.4)$$

for  $j = 1, \dots, n$ . Here we use the fact that if  $q(x, y) = Q(x+iy, x-iy) = Q(z, \bar{z})$  is real, then  $q_x = q_y = 0$  if and only if  $Q_{\bar{z}} = 0$  or  $Q_z = 0$  (see also Exercise 6, Chap. 1). The uniqueness of the solution of the system (6.4) and the fact that this procedure gives a minimum are guaranteed by the orthogonal projection theorem for closed linear spaces (see Sect. 1.1). The problem of prediction for time

series has been reduced (in the linear approximation) to the solution of  $n$  linear equations in  $n$  unknowns. This concludes our general discussion of prediction for time series.

We now turn to a special case in which this linear system of equations can be solved analytically with the help of complex variables. The reader not familiar with contour integration should fast forward at this point to the next section. Rewrite (6.4) in terms of the Fourier transform. The spectral representation of  $R(T)$  is

$$R(T) = \int_{-\pi}^{\pi} e^{ikT} g(k) dk.$$

Then (6.4) becomes

$$\int_{-\pi}^{\pi} \left( -e^{i(j+m)k} + \sum_{l=1}^n a_l e^{i(j-l)k} \right) g(k) dk = 0.$$

Moving  $e^{ijk}$  outside the parentheses, we get

$$\int_{-\pi}^{\pi} e^{ijk} \left( e^{imk} - \sum_{l=1}^n a_l e^{-ilk} \right) g(k) dk = 0. \quad (6.5)$$

So far, (6.5) is just a reformulation of (6.4). To continue, we need an explicit representation of  $g(k)$ . Consider the special case  $R(T) = Ca^{|T|}$  for  $T = 0, \pm 1, \pm 2, \dots$ , where  $C > 0$  and  $0 < a < 1$ . Is  $R$  the covariance of a stationary process? It certainly satisfies conditions 1, 2, 3. To check condition 4, we compute

$$\begin{aligned} g(k) &= \frac{1}{2\pi} \sum_{n=-\infty}^{\infty} R(n) e^{-ink} \\ &= \frac{C}{2\pi} \left[ \sum_{n=1}^{\infty} (ae^{-ik})^n + 1 + \sum_{n=1}^{\infty} (ae^{ik})^n \right] \\ &= \frac{C}{2\pi} \left[ \frac{ae^{-ik}}{1 - ae^{-ik}} + 1 + \frac{ae^{ik}}{1 - ae^{ik}} \right] \\ &= \frac{C}{2\pi} \frac{1 - a^2}{(1 - ae^{-ik})(1 - ae^{ik})} > 0. \end{aligned}$$

This shows that  $R(T)$  is the Fourier transform of a nonnegative function, and consequently the covariance of a stationary process.

Assume for simplicity that  $C(1 - a^2)/(2\pi) = 1$ . We solve (6.5) using complex variables. Let  $e^{ik} = z$ . Then  $\bar{z} = z^{-1}$ ,  $dk = dz/(iz)$ , and (6.5) becomes

$$\frac{1}{2\pi} \int_{|z|=1} z^j \left( z^m - \sum_{\ell=1}^n a_\ell z^{-\ell} \right) \frac{1}{(z - a) \left( \frac{1}{z} - a \right)} \frac{dz}{iz} = 0$$

for  $j = 1, 2, \dots, n$ . We must therefore determine  $a_1, \dots, a_n$  such that

$$\sum_{\ell=1}^n a_\ell \frac{1}{2\pi i} \int_{|z|=1} \frac{z^{j-\ell}(1 - az)^{-1}}{z - a} dz = \frac{1}{2\pi i} \int_{|z|=1} \frac{z^{j+m}(1 - az)^{-1}}{z - a} dz.$$

We find the coefficients recursively by comparing two consecutive values of  $j$ , starting from the back. Let  $j = n$  and  $j = n - 1$ . Using residue theory, we get

$$\sum_{\ell=1}^n \frac{a_\ell a^{n-\ell}}{1 - a^2} = \frac{a^{n+m}}{1 - a^2},$$

$$\sum_{\ell=1}^{n-1} \frac{a_\ell a^{n-1-\ell}}{1 - a^2} + a_n \left[ \frac{a^{-1}}{1 - a^2} + \frac{(1 - a \cdot 0)^{-1}}{0 - a} \right] = \frac{a^{n-1+m}}{1 - a^2}.$$

Multiplying the last equation by  $a$  and subtracting, we get  $a_n = 0$ . This simplifies the next step with  $j = n - 1$  and  $j = n - 2$  substantially, and using similar arguments, we obtain  $a_{n-1} = 0$ . In the last step,

$$\frac{a_1}{2\pi i} \int_{|z|=1} \frac{z(1 - az)^{-1}}{z} \frac{dz}{z - a} = \frac{1}{2\pi i} \int_{|z|=1} \frac{z^{1+m}(1 - az)^{-1}}{z - a} dz,$$

which yields  $a_1(1 - a^2)^{-1} = a^{1+m}(1 - a^2)^{-1}$ , or  $a_1 = a^{1+m}$ . We have therefore shown that if  $R(T) = Ca^{|T|}$  with  $0 < a < 1$ , then the best approximation of  $u(t + m, \omega)$  for  $m = 0, 1, \dots$  is  $a^{1+m}u(t - 1, \omega)$ . This is intuitively obvious: the correlations between variables decays like  $a$  to the power of the distance between them, so the predictive power of the last-measured quantity decays in the same way.

### 6.5. Random Measures and Random Fourier Transforms

We showed previously that the covariance of a wide-sense stationary stochastic process can be written as the Fourier transform of a spectral density. We now use this fact to find useful representations for the process itself, including a stochastic generalization of the Fourier

transform that does not require that the process have samples to which the Fourier transform can be applied individually. These representations will be convolutions of nonrandom functions with certain simple processes.

The reader may wish to know that the material in the present section will not be used in the remainder of the book, and therefore can be skipped on a first reading.

Given a probability space  $(\Omega, \mathcal{B}, P)$ , consider the set of random variables  $f(\omega)$ , where  $\omega$  is in  $\Omega$ , such that  $E[ff] < \infty$ . We refer to this set as  $L_2(\Omega, \mathcal{B}, P)$ . We now construct a one-to-one mapping  $L_2(\Omega, \mathcal{B}, P) \rightarrow L_2(A, \mu)$ , where  $A$  is a subset of the  $t$ -axis and  $\mu$  is a measure on  $A$ . Consider  $\mathcal{A}$ , an algebra of subsets of  $A$ , i.e., a collection of sets with the property that if the sets  $A_i$  are in  $\mathcal{A}$ , then so are their complements, as well as their finite unions and intersections; an algebra is much like a  $\sigma$ -algebra, with the exception that we do not require that the union of a countably infinite family of subsets belong to the algebra, a detail that is important in a rigorous analysis, but which we will disregard here.

Consider the triple  $(A, \mathcal{A}, \mu)$ , where  $\mu$  is a rule that to each subset  $A_i \in \mathcal{A}$  assigns a number such that

1.  $\mu(A_i) \geq 0$ .
2.  $\mu(A_i)$  is finite.
3.  $\mu(\emptyset) = 0$ .
4.  $A_i \cap A_j = \emptyset \Rightarrow \mu(A_i \cup A_j) = \mu(A_i) + \mu(A_j)$ .

(Again, note that we are concerned only with finitely many  $A_i$ .) Next, construct a random variable  $\rho = \rho(A_i, \omega)$ , where  $A_i \in \mathcal{A}$  and  $\omega \in \Omega$  (recall that a random variable is a function defined on  $\Omega$ ), that has the following properties:

1.  $A_i \cap A_j = \emptyset \Rightarrow \rho(A_i \cup A_j, \omega) = \rho(A_i, \omega) + \rho(A_j, \omega)$ .
2.  $\rho(A_i, \omega)$  is square integrable, i.e.,  $E[\rho(A_i, \omega)\bar{\rho}(A_i, \omega)] < \infty$ .
3.  $\rho(\emptyset, \omega) = 0$ .
4.  $A_i, A_j \subset A \Rightarrow E[\rho(A_i, \omega)\bar{\rho}(A_j, \omega)] = \mu(A_i \cap A_j)$ .

The properties listed above imply that  $\mu(A_i) \geq 0$  for all  $A_i \in \mathcal{A}$ , since

$$\mu(A_i) = \mu(A_i \cap A_i) = E[\rho(A_i, \omega)\bar{\rho}(A_i, \omega)] \geq 0.$$



We call  $\mu$  the *structure function* of  $\rho$ . Just as a stochastic process is a function of both  $\omega$  and  $t$ , a random measure is a function of both  $\omega$  and the subsets  $A_i$  of  $A$ .

Now define  $\chi_{A_i} = \chi_{A_i}(t)$ , the characteristic function of the subset  $A_i$  of the  $t$ -axis, to be

$$\chi_{A_i}(t) = \begin{cases} 1, & t \in A_i, \\ 0, & \text{otherwise,} \end{cases}$$

and consider a function  $q(t)$  of the form

$$q(t) = \sum c_i \chi_{A_i}(t).$$

We consider the case in which  $\{A_i\}$  is a finite partition of  $A$ , i.e., there are only finitely many  $A_i$ ,  $A_i \cap A_j = \emptyset$ , for  $i \neq j$ , and  $\bigcup A_i = A$ . Thus,  $q(t)$  takes on only a finite number of values. To this function  $q(t)$  assign the random variable

$$f(\omega) = \sum c_i \rho(A_i, \omega).$$

Hence, each characteristic function of a subset is replaced by the random variable that the random measure assigns to the same subset; thus, this substitution transforms a function of  $t$  into a function of  $\omega$  (i.e., into a random variable).

Now consider the product  $q_1(t)\overline{q_2}(t)$  of two functions of the form

$$q_1 = \sum_{j=1}^n c_j \chi_{A_j}(t), \quad q_2 = \sum_{k=1}^m d_k \chi_{B_k}(t),$$

where  $\{B_i\}$  is another finite partition of  $A$ . It is not necessary for  $n$  and  $m$  to be equal. There is a finite number of intersections of the  $A_j$  and  $B_k$ , and on each of these subsets, the product

$$\begin{aligned} q_1 \overline{q_2} &= \left( \sum_{j=1}^n c_j \chi_{A_j} \right) \left( \sum_{k=1}^m \overline{d_k} \chi_{B_k} \right) \\ &= \sum_{j=1}^n \sum_{k=1}^m c_j \overline{d_k} \chi_{A_j \cap B_k}, \end{aligned}$$

takes on a constant value  $c_j \overline{d_k}$ . Thus, the same construction allows us to assign a random variable  $f_1 \overline{f_2}$  to the product  $q_1 \overline{q_2}$ . Since

$$f_1(\omega) = \sum c_j \rho(A_j, \omega), \quad f_2(\omega) = \sum d_k \rho(B_k, \omega),$$

we conclude that

$$\begin{aligned}
 E[f_1 \bar{f}_2] &= E \left[ \sum_{j=1}^n \sum_{k=1}^m c_j \bar{d}_k \rho(A_j, \omega) \bar{\rho}(B_k, \omega) \right] \\
 &= \sum_{j=1}^n \sum_{k=1}^m c_j \bar{d}_k E[\rho(A_j, \omega) \bar{\rho}(B_k, \omega)] \\
 &= \sum_{j=1}^n \sum_{k=1}^m c_j \bar{d}_k \mu(A_j \cap B_k) \\
 &= \int q_1 \bar{q}_2 \mu(dt). \tag{6.6}
 \end{aligned}$$

Thus we have established a mapping between random variables with finite mean squares and functions of time with finite square integrals (i.e., between the random variables  $f(\omega)$  and functions  $q(t)$  such that  $\int q_1(t) \bar{q}_2(t) \mu(dt)$  is finite). Although we have defined the mapping only for functions  $q(t) = \sum c_i \chi_{A_i}(t)$ , an argument that we omit enables us to extend the mapping to all random variables and functions of  $t$  with the square integrability properties listed above.

**EXAMPLE.** We now show in detail how this construction works for a very special case. Say we are given a probability space  $(\Omega, B, P)$  and three subsets of the  $t$ -axis:  $A_1 = [0, 1)$ ,  $A_2 = [1, 3)$ , and  $A_3 = [3, 3\frac{1}{2}]$ . Each  $A_i$  is assigned a real-valued random variable  $\rho_i(\omega) = \rho(A_i, \omega)$  that has mean 0 and variance equal to the length of  $A_i$ . For example,  $\rho_1(\omega)$  has mean 0 and variance 1, and so forth. The variables  $\rho_1$ ,  $\rho_2$ , and  $\rho_3$  are independent, and  $E[\rho_i \rho_j] = 0$  for  $i \neq j$ , where  $E[\rho_i^2]$  is the length of the  $i$ th interval. Moreover,

$$\begin{aligned}
 \chi_1(t) &= \begin{cases} 1, & 0 \leq t < 1, \\ 0, & \text{elsewhere,} \end{cases} \\
 \chi_2(t) &= \begin{cases} 1, & 1 \leq t < 3, \\ 0, & \text{elsewhere,} \end{cases} \\
 \chi_3(t) &= \begin{cases} 1, & 3 \leq t \leq 3\frac{1}{2}, \\ 0, & \text{elsewhere,} \end{cases}
 \end{aligned}$$

where  $\int \chi_i \chi_j dt = 0$  for  $i \neq j$  and  $\int \chi_i^2 dt$  is the length of the  $i$ th interval.

Now take a function of the form  $q_1(t) = \sum_i c_i \chi_i(t)$ , where the  $c_i$  are constants. Clearly,

$$f_1(\omega) = \sum_{i=1}^3 c_i \rho_i(\omega).$$

Suppose we have another function  $q_2(t)$  on the same partition:

$$q_2(t) = \sum_{j=1}^3 d_j \chi_j(t) \rightarrow f_2(\omega) = \sum_{j=1}^3 d_j \rho_j(\omega).$$

Then

$$\begin{aligned} E[f_1 \bar{f}_2] &= E \left[ \sum_{i=1}^3 \sum_{j=1}^3 c_i \bar{d}_j \rho_i \rho_j \right] \\ &= \sum_{j=1}^3 c_j \bar{d}_j E[\rho_j^2] \\ &= \sum_{j=1}^3 c_j \bar{d}_j \mu(A_j), \end{aligned}$$

where  $\mu(A_j)$  is the length of  $A_j$ . Notice also that

$$\begin{aligned} \int_0^{3\frac{1}{2}} q_1(t) \bar{q}_2(t) dt &= \int_0^{3\frac{1}{2}} \sum_{i=1}^3 \sum_{j=1}^3 c_i \bar{d}_j \chi_i(t) \chi_j(t) dt \\ &= \sum_j c_j \bar{d}_j \mu(A_j), \end{aligned}$$

which verifies that  $q(t) \rightarrow f(\omega)$ , so  $E[f_1 \bar{f}_2] = \int q_1(t) \bar{q}_2(t) \mu(dt)$  as in (6.6).

Now approximate every square integrable function  $q$  on  $A$  by a step function, construct the corresponding random variable, and take the limit, as the approximation improves, of the sequence of random variables obtained in this way. This makes for a mapping of square integrable functions on  $A$  onto random variables with finite mean squares. This mapping can be written as

$$f(\omega) = \int q(s) \rho(ds, \omega)$$

(the right-hand side is an integral with respect to the measure  $\rho$ ), where the variable  $t$  has been replaced by  $s$  for convenience. Now view a stochastic process  $u$  as a family of random variables labeled by the parameter  $t$  (i.e., there is a random variable  $u$  for every value of  $t$ ) and apply the representation just derived at each value of  $t$ . Therefore,

$$u(t, \omega) = \int q(t, s) \rho(ds, \omega).$$

Assume that  $u(t, \omega)$  is stationary in the wide sense. Then the covariance of  $u$  is

$$\begin{aligned} R(t_1 - t_2) &= E[u(t_1, \omega) \overline{u(t_2, \omega)}] \\ &= E \left[ \int q(t_1, s_1) \rho(ds_1) \int \overline{q(t_2, s_2)} \overline{\rho(ds_2)} \right] \\ &= E \left[ \int q(t_1, s_1) \overline{q(t_2, s_2)} \rho(ds_1) \overline{\rho(ds_2)} \right] \\ &= \int q(t_1, s_1) \overline{q(t_2, s_2)} E[\rho(ds_1) \overline{\rho(ds_2)}] \\ &= \int q(t_1, s) \overline{q(t_2, s)} \mu(ds). \end{aligned}$$

One can show that the converse is also true: if the last equation holds, then  $u(t, \omega) = \int q(t, s) \rho(ds, \omega)$  with  $E[\rho(ds) \overline{\rho(ds)}] = \mu(ds)$ . Note that in all of the above, equality holds in a mean square ( $L_2$ ) sense, and little can be said about the higher moments.

**EXAMPLE.** If  $u = u(t, \omega)$  is a wide-sense stationary stochastic process, then it follows from Khinchin's theorem that

$$R(T) = E[u(t+T, \omega) \overline{u(t, \omega)}] \tag{6.7}$$

$$= \int e^{ikT} dG(k). \tag{6.8}$$

Conversely, if  $E[\rho(dk) \overline{\rho(dk)}] = dG(k)$ , we see that if

$$u(t, \omega) = \int e^{ikt} \rho(dk, \omega),$$

then

$$\begin{aligned} E[u(t+T, \omega) \overline{u(t, \omega)}] &= \int e^{ik(t+T-t)} E[\rho(dk) \overline{\rho(dk)}] \\ &= \int e^{ikT} dG(k). \end{aligned}$$

We have just shown that  $dG(k)$  is the energy density in the interval  $dk$ . This  $\rho(k)$  is the stochastic Fourier transform of  $u$ . The inverse Fourier transform does not exist in the usual sense (i.e.,  $\int u(t, \omega) e^{-ikt} dt$  for each  $\omega$  does not exist), but for (6.5) to hold, it is sufficient for  $E[|u(t)|^2]$  to exist for each  $t$ .

One can summarize the construction of the stochastic Fourier transform as follows: For the ordinary Fourier transform, the Parseval identity is a consequence of the definitions. To generalize the Fourier transform, we started from a general form of Parseval's identity and found a generalized version of the Fourier transform that satisfies it.

EXAMPLE. Suppose  $dG(k) = g(k) dk$ . Then

$$\int e^{ik(t_2-t_1)} dG(k) = \int e^{ikt_2} \sqrt{g(k)} e^{-ikt_1} \sqrt{g(k)} dk.$$

Recall that  $g(k) \geq 0$ . Write  $\sqrt{g(k)} = \hat{h}(k) = \widehat{h(t)}(k)$ , where  $h(t)$  is the inverse Fourier transform of  $\hat{h}(k)$ ,  $\hat{h}(k) = \frac{1}{\sqrt{2\pi}} \int h(t) e^{-ikt} dt$ . Then

$$\begin{aligned} e^{-ikt_2} \sqrt{g(k)} &= e^{-ikt_2} \frac{1}{\sqrt{2\pi}} \int h(t) e^{-ikt} dt \\ &= \frac{1}{\sqrt{2\pi}} \int h(t) e^{-ik(t+t_2)} dt \\ &= \frac{1}{\sqrt{2\pi}} \int h(t-t_2) e^{-ikt} dt \\ &= \widehat{h(t-t_2)}(k), \end{aligned}$$

where the  $(k)$  at the very end is there to remind you that  $\widehat{h(t-t_2)}$  is a function of  $k$ . Since the Fourier transform preserves inner products, we find that

$$R(t_1, t_2) = \int \bar{h}(t-t_1) h(t-t_2) dt,$$

and by changing  $t$  to  $s$ , we obtain

$$R(t_1, t_2) = \int \bar{h}(s - t_1)h(s - t_2)\mu(ds),$$

where  $\mu(ds) = ds$ . Applying our representation, we get  $u(t, \omega) = \int \bar{h}(s - t)\rho(ds)$ , where  $E[|\rho(ds)|^2] = ds$ . The random measure constructed as increments of Brownian motion at instants  $ds$  apart has this property. Thus, any wide-sense stationary stochastic process with  $dG(k) = g(k) dk$  can be approximated as a sum of translates (in time) of a fixed function, each translate multiplied by independent Gaussian random variables. This is the *moving average* representation.

### 6.6. Exercises

1. Find some way to show nonrigorously that the covariance function of white noise is a delta function. Suggestion: Approximate Brownian motion by a random walk with Gaussian increments of nonzero length, find the time series of the difference quotients of this walk, calculate its covariance, and take a formal limit.
2. Consider the stochastic process  $u = \xi \cos(t)$ , where  $\xi$  is a random variable with mean 0 and variance 1. Find the mean and the covariance functions. Obviously, this is not a stationary process. However,  $\cos(t) = (e^{it} + e^{-it})/2$ . How do you reconcile this with the construction we have of stationary processes as sums of exponentials?
3. Consider the differential equation  $(u^2)_x = \epsilon u_{xx}$  on the real line, with the boundary conditions  $u(-\infty) = u_0$ ,  $u(+\infty) = -u_0$ , where  $\epsilon$  and  $u_0$  are constants. Assume that  $u$  is a velocity, with dimension  $L/T$ , where  $L$  is the dimension of length and  $T$  the dimension of time. Find the dimension of  $\epsilon$ . Because of the boundary conditions,  $u$  does not have a usual Fourier transform, but one can define one by taking the Fourier transform of  $u'$  and dividing it by  $ik$ . Let  $\hat{u}(k)$  be this Fourier transform of  $u$ . Define the energy spectrum by  $E(k) = |\hat{u}(k)|^2$ . Find the dimension of  $E(k)$ ; show that the dimensionless quantity  $E(k)k^2/u_0^2$  must be a function of the variable  $k\epsilon/u_0$ . Assume complete similarity, and deduce that as you pass to the limit  $\epsilon \rightarrow 0$ , the spectrum converges to  $E(k) = C/k^2$  for some constant  $C$ .

4. Consider the wide-sense stationary stochastic process  $u = \xi e^{it}$ , where  $\xi$  is a Gaussian variable with mean 0 and variance 1. What is its stochastic Fourier transform? What is the measure  $\rho(dk)$ ?
5. Consider a stochastic process of the form  $u(\omega, t) = \sum_j \xi_j e^{i\lambda_j t}$ , where the sum is finite and the  $\xi_j$  are independent random variables with mean 0 and variance  $v_j$ . Calculate the limit as  $T \rightarrow \infty$  of the random variable  $(1/T) \int_{-T}^T |u(\omega, s)|^2 ds$ . How is it related to the spectrum as we have defined it? What is the limit of  $(1/T) \int_{-T}^T u ds$ ?
6. Suppose you have to construct on the computer (for example, for the purpose of modeling the random transport of pollutants) a Gaussian stationary stochastic process with mean 0 and a given covariance function  $R(t_1 - t_2)$ . Propose a construction.
7. Show that there is no stationary (in the wide sense) stochastic process  $u = u(\omega, t)$  that satisfies (for each  $\omega$ ) the differential equation  $y'' + 4y = 0$  as well as the initial condition  $y(t = 0) = 1$ .
8. Let  $\eta$  be a random variable. Its characteristic function is defined as  $\phi(\lambda) = E[e^{i\lambda\eta}]$ . Show that  $\phi(0) = 1$  and that  $|\phi(\lambda)| \leq 1$  for all  $\lambda$ . Show that if  $\phi_1, \phi_2, \dots, \phi_n$  are the characteristic functions of independent random variables  $\eta_1, \dots, \eta_n$ , then the characteristic function of the sum of these variables is the product of the  $\phi_i$ .
9. Show that if  $\phi(\lambda)$  is the characteristic function of  $\eta$ , then

$$E[\eta^n] = (-i)^n \frac{d^n}{d\lambda^n} \phi(0),$$

provided both sides of the equation make sense. Use this fact to show that if  $\xi_i$ ,  $i = 1, \dots, n$ , are Gaussian variables with mean 0, *not necessarily independent*, then

$$E[\xi_1 \xi_2 \cdots \xi_n] = \begin{cases} \sum \Pi E[\xi_{i_k} \xi_{j_k}], & n \text{ even,} \\ 0, & n \text{ odd.} \end{cases}$$

On the right-hand side,  $i_k$  and  $j_k$  are two of the indices, the product is over a partition of the  $n$  indices into disjoint groups of two, and the sum is over all such partitions (this is Wick's theorem). Hints: Consider the variable  $\sum \lambda_j \xi_j$ ; its moments can be calculated from the derivatives of its characteristic function. By assumption, this

variable is Gaussian and its characteristic function, i.e., the Fourier transform of its density, is given by a formula we have derived.

10. Consider the following functions  $R(T)$ ; which ones are the covariances of some wide-sense stationary stochastic process, and why? (here  $T = t_1 - t_2$ , as usual):

1.  $R(T) = e^{-T^2}$ .
2.  $R = Te^{-T^2}$ .
3.  $R = e^{-T^2/2}(T^2 - 1)$ .
4.  $R = e^{-T^2/2}(1 - T^2)$ .

### 6.7. Bibliography

- [1] G.I. BARENBLATT, *Scaling*, Cambridge University Press, Cambridge, 2004.
- [2] G.I. BARENBLATT AND A.J. CHORIN, A mathematical model for the scaling of turbulence, Proc. Natl. Acad. Sci. USA, 101 (2004), pp. 15,023–15,026.
- [3] R. GHANEM AND P. SPANOS, *Stochastic Finite Elements*, Dover, NY, 2012
- [4] I. GIKHMAN AND A. SKOROKHOD, *Introduction to the Theory of Random Processes*, Saunders, Philadelphia, 1965.
- [5] A. YAGLOM, *An Introduction to the Theory of Stationary Random Functions*, Dover, New York, 1962.