1

Introduction: Some basic concepts of the theory of random processes

Fluctuations in solids are typical random processes. By definition, a random process is a random function $x(t)$ of an independent variable t, which, in the theory of fluctuations, is the time. The theory of fluctuations is based on the highly developed mathematical theory of random processes. This chapter is an introduction to the basic concepts and methods of this theory which are used throughout this book. The list of recommended books and reviews is presented at the end of the book.

1.1 Probability density functions; moments; stationary processes

Let $P\{\ldots\}$ be the probability of an event indicated in the curly brackets. The distribution functions of the first, second, and higher order for a given random process $x(t)$ are defined as

$$
\begin{aligned}
W_1(x_1, t_1) &= P\{x(t_1) \leq x_1\} \\
W_2(x_1, t_1; x_2, t_2) &= P\{x(t_1) \leq x_1; x(t_2) \leq x_2)\} \\
W_n(x_1, t_1; \ldots; x_n, t_n) &= P\{x(t_1) \leq x_1; \ldots; x(t_n) \leq x_n\},
\end{aligned}
\tag{1.1.1}
$$

where x_i is the value of the random quantity at instant t_i.

For instance, $W_2(x_1, t_1; x_2, t_2)$ is the probability that the random quantity x at instant t_1 is less than a given value x_1, and at instant t_2 is less than x_2.

If $W_n(x_1, t_1; \ldots; x_n, t_n)$ are differentiable functions of the variables $x_1, \ldots, x_n$, one may introduce the corresponding probability density functions:

$$
\begin{aligned}
w_1(x_1, t_1) &= \frac{\partial W_1(x_1, t_1)}{\partial x_1}, \\
w_2(x_1, t_1; x_2, t_2) &= \frac{\partial^2 W_2(x_1, t_1; x_2, t_2)}{\partial x_1 \partial x_2}.
\end{aligned}
\tag{1.1.2}
$$

The functions w_n are called n-dimensional probability density functions

1

of the random quantity $x(t)$. They may be defined in a different way. For instance, the two-dimensional probability density function,

$$w_2(x_1, t_1; x_2, t_2)dx_1 dx_2 = P\{x_1 \leq x(t_1) < x_1 + dx_1; x_2 \leq x(t_2) < x_2 + dx_2\}.$$
(1.1.3)

The probability density functions must be consistent with each other, that is, each k-dimensional density function can be obtained from any n-dimensional function with $n > k$ by integration over 'extra' variables:

$$w_k(x_1, t_1; \ldots; x_k, t_k) = \int dx_{k+1} \ldots dx_n w_n(x_1, t_1; \ldots; x_n, t_n).$$
(1.1.4)

Random processes are called stationary if all distributions $w_n(x_1, t_1; \ldots; x_n, t_n)$ remain invariable under any identical shift of all time points $t_1, \ldots, t_n$. It means that for a stationary process the probability density function $w_1(x_1, t_1)$ does not depend at all on the time t_1, the probability density function $w_2(x_1, t_1; x_2, t_2)$ depends only on the difference $t_1 - t_2$, that is, may be written as $w_2(x_1, x_2, t_1 - t_2)$, and so on.

The mean value of the random variable $x(t)$ is equal to (for the methods of averaging random quantities see below)

$$\langle x(t) \rangle = \int_{-\infty}^{+\infty} x(t)dW(x) = \int_{-\infty}^{+\infty} dx\, x w_1(x, t).$$
(1.1.5)

The r-th order central moment is the average value of the random quantity $(\delta x(t))^r$, where $\delta x(t) = x(t) - \langle x \rangle$ is the deviation of the random quantity $x(t)$ from its mean value $\langle x \rangle$, i.e., the fluctuation:

$$\langle [\delta x(t)]^r \rangle = \int_{-\infty}^{+\infty} dx (\delta x)^r w_1(x, t).$$
(1.1.6)

The second-order central moment is the variance, i.e., the mean value of the fluctuation squared.

If the probability density $w_n(x_1, t_1; \ldots; x_n, t_n)$ is known for different instants $t_1, \ldots, t_n$, one can calculate the correlation function:

$$\langle \delta x(t_1) \ldots \delta x(t_n) \rangle = \int dx_1\, dx_2 \ldots dx_n\, \delta x_1 \cdots \delta x_n\, w_n(x_1, t_1; \ldots; x_n, t_n). \quad (1.1.7)$$

The characteristic function of a random quantity is defined as the mean value of the function $\exp(iux)$, where x is the random variable and u is a real parameter:

$$\phi_x(u) \equiv \langle e^{iux} \rangle = \int_{-\infty}^{+\infty} e^{iux}dW(x).$$
(1.1.8)

If the distribution function $W(x)$ is differentiable (see above), the characteristic function is simply the Fourier transform of the probability density

function, Eq. (1.1.2):

$$\phi_x(u) = \int_{-\infty}^{+\infty} dx e^{iux} w(x). \tag{1.1.9}$$

The characteristic function has a very useful property: if the random variable X is a sum of N *independent* random variables, i.e.,

$$X = \sum_{i=1}^{N} x_i, \tag{1.1.10}$$

then the characteristic function of X is simply the product of the characteristic functions of the variables x_i :

$$\phi_X(u) = \phi_{x_1}(u) \cdot \phi_{x_2}(u) \cdots \phi_{x_N}(u). \tag{1.1.11}$$

Therefore, in many problems in which the random variable is given by Eq. (1.1.10) it is more convenient to calculate $\phi_X(u)$ first and then the density function $w(X)$ using the inverse Fourier transformation (Eq. (1.1.9)):

$$w(x) = \int_{-\infty}^{+\infty} \frac{du}{2\pi} \phi_x(u) e^{-iux}. \tag{1.1.12}$$

In the theory of fluctuations an important distribution is the normal (Gaussian) distribution. It arises when the random quantity $x(t)$ is a sum of many ($N \gg 1$) independent and identically distributed random quantities. An example is the noise generated in a macroscopic specimen by the random motion of a great number of independent but identical defects, or spins. Let $\xi_1, \ldots, \xi_N$ be independent and identically distributed random quantities and let $x = \xi_1 + \cdots + \xi_N$. If the summands $\xi_1, \ldots, \xi_N$ are small enough and their number N is great enough, that is, if as $N \to \infty$ the mean value of $x(t)$ is equal to $\langle x \rangle$, and the variance of $x(t)$ is equal to σ^2, the one-dimensional probability density function is

$$w_1(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\delta x)^2}{2\sigma^2}\right], \tag{1.1.13}$$

where $\delta x = x - \langle x \rangle$ is the fluctuation. Such a distribution is called normal or Gaussian.

A random process is called Gaussian if *all* its probability density functions are normal for all $n = 1, \ldots$. The n-dimensional normal distribution reads:

$$w_n(x_1, \ldots, x_n) = \frac{1}{\sqrt{(2\pi)^n \det \hat{\lambda}}} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ij}^{-1} \delta x_i \delta x_j\right]. \tag{1.1.14}$$

Here $\delta x_i = x_i - \langle x_i \rangle$. The matrix $\hat{\lambda}$ is called the covariance matrix. Its

elements equal:

$$\lambda_{ij} = \langle \delta x_i \delta x_j \rangle = \lambda_{ji}. \tag{1.1.15}$$

In Eq. (1.1.14) $\det \hat{\lambda}$ is the determinant of the matrix $\hat{\lambda}$, λ_{ij}^{-1} are the elements of the matrix $\hat{\lambda}^{-1}$ which is the reciprocal of $\hat{\lambda}$.

In particular, the two-dimensional normal distribution equals:

$$
\begin{aligned}
w_2(x_1, x_2) = {} & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_{12}^2}} \exp\Big\{-\frac{1}{2(1-\rho_{12}^2)}\Big[\frac{(\delta x_1)^2}{\sigma_1^2} + \frac{(\delta x_2)^2}{\sigma_2^2} \\
& - 2\rho_{12}\frac{\delta x_1}{\sigma_1}\frac{\delta x_2}{\sigma_2}\Big]\Big\}.
\end{aligned}
\tag{1.1.16}
$$

Here $\sigma_i^2 = \langle(\delta x_i)^2\rangle$ are the variances, $\rho_{12} = \langle \delta x_1 \delta x_2 \rangle / \sigma_1 \sigma_2$ is the correlation coefficient.

Equations (1.1.7) and (1.1.14) imply that for Gaussian random processes all nonzero n-th order ($n > 2$) moments can be expressed in terms of the second-order moments, i.e., the covariances λ_{ij} (pair correlations). In other words, the measurement of higher-order correlations can not give any new information apart from that which is contained in the pair correlation (for the consequences of this property of Gaussian random processes see Sec. 1.2).

Many random processes are similar to electron emission from the cathode of a thermionic tube. The instants of emission of individual electrons are random because the emission events are uncorrelated. Such random processes are called Poissonian. The hops of identical noninteracting defects in a solid are another example of such processes. Let n_1 be the mean rate of individual events, that is, the mean number of events per unit time. In the case of electron emission, n_1 is the mean number of electrons emitted per unit time. The mean number of such events during a definite time t_m is obviously $\bar{n} = n_1 t_m$. Of course, the actual number n of events that happen to occur during the time t_m is a random quantity which fluctuates around its mean value $\bar{n}$. The problem is to find the probability $P(n)$ of exactly n individual events occurring in the time $t_m = \bar{n}/n_1$.

Let us divide the time t_m into N equal intervals. N is taken so large or, equivalently, the intervals t_m/N are taken so small that occurence of more than one event in one interval is improbable. For each small interval two alternatives are possible: either one individual event or no event may occur. The corresponding probabilities are $\bar{n}/N$ and $1 - \bar{n}/N$, respectively. The probability of each definite sequence of n 'full' and $N - n$ 'empty' intervals (one event or no event occured, respectively) equals $(\bar{n}/N)^n(1 - \bar{n}/N)^{N-n}$.

The total number of different sequences (arrangements) of n 'full' and $N - n$ 'empty' intervals is the number of combinations $N!/n!(N-n)!$. Thus the probability of n events occurring equals

$$\frac{N!}{n!(N-n)!}\Big(\frac{\bar{n}}{N}\Big)^n\Big(1-\frac{\bar{n}}{N}\Big)^{N-n}$$
$$= \frac{\bar{n}^n}{n!}\Big(1-\frac{\bar{n}}{N}\Big)^N \cdot 1 \cdot \Big(1-\frac{1}{N}\Big)\cdots\Big(1-\frac{n-1}{N}\Big)\Big(1-\frac{\bar{n}}{N}\Big)^{-n}.$$

The sought probability, which is called the Poisson distribution, follows from the last equation in the limit $N \to \infty$ at finite n. Since $\lim_{N\to\infty}(1-\bar{n}/N)^N = \exp(-\bar{n})$,

$$P(n) = \frac{\bar{n}^n}{n!}e^{-\bar{n}}. \tag{1.1.17}$$

The statistical characteristics of a random process are nonrandom quantities, which can be, in principle, measured by a corresponding averaging of the random quantity (or quantities). There are two possible procedures of averaging.

1. Averaging over time, i.e., over a sufficiently long record of the random process. In this case, for example, the distribution function $W_1(x_1)$ of a stationary process is equal to that part of the total time t_m of measurement of this process during which the random quantity $x < x_1$. The mean value of the random quantity is then

$$\bar{x} = \lim_{t_m\to\infty}\frac{1}{t_m}\int_{-t_m/2}^{t_m/2} dt\, x(t). \tag{1.1.18}$$

2. Averaging over an ensemble of a large number N of identical systems, i.e., systems in which identical conditions for the random process and identical means of its measurement are provided. In this case the distribution function, $W_1(x_1,t_1)$, is equal to the fraction of the total number of ensemble systems in which at instant t_1 the random quantity $x < x_1$. The mean value in this case (denoted by angle brackets) equals:

$$\langle x \rangle = \lim_{N\to\infty}\frac{1}{N}\sum_{i=1}^{N} x_i(t). \tag{1.1.19}$$

Here $x_i(t)$ is the value of the random quantity at the instant t in the i-th system of the ensemble.

One may ask: does a definite mean value exist if one increases the time t_m of averaging or the number N of identical systems of the ensemble? This question is fundamental. It is often reduced to the following one: do both averaging procedures (see above) yield identical mean values? If the

answers are positive, the random process is called ergodic. Ergodicity and nonergodicity are discussed in Sec. 1.4.

1.2 The correlation function

The correlation function is one of the most important characteristics of any random process. Let $x(t)$ be a quantity that randomly varies in time t and let $\langle x \rangle$ be its mean value. The variation in time of the fluctuation, i.e., the deviation of the random quantity from its mean value, $\delta x(t) = x(t) - \langle x \rangle$, is, of course, also random. The correlation function is a nonrandom characteristic of the kinetics of these random fluctuations: it shows how the fluctuations evolve in time *on average*.

Consider an ensemble of a large number N of systems in which identical conditions for the random process under consideration have been created. Let $\delta x_i(t_1)$ and $\delta x_i(t_2)$ be the fluctuations in the i-th system of the ensemble ($i = 1, \ldots, N$) at instants t_1 and t_2, respectively. The correlation function is defined then by the following equation:

$$\psi_x(t_1, t_2) \equiv \langle \delta x(t_1) \delta x(t_2) \rangle = \langle x(t_1) x(t_2) \rangle - \langle x(t_1) \rangle \langle x(t_2) \rangle$$
$$= \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \delta x_i(t_1) \delta x_i(t_2). \tag{1.2.1}$$

Here and below the angle brackets denote averaging over an ensemble of identical systems.

In a stationary system the correlation function remains invariable if both instants, t_1 and t_2, are shifted identically. In such systems ψ_x depends only on the difference $t_1 - t_2$ (Sec. 1.1).

The correlation function may be determined also by averaging over a sufficiently long record of the random process $x(t)$ in one system, i.e., by averaging over a long enough time of measurement t_m:

$$\psi_x(t_1 - t_2) \equiv \overline{\delta x(t_1) \delta x(t_2)} = \lim_{t_m \to \infty} \frac{1}{t_m} \int_{-t_m/2}^{t_m/2} dt \, \delta x(t_1 + t) \delta x(t_2 + t). \tag{1.2.2}$$

Both correlation functions (1.2.1) and (1.2.2) coincide if the system is ergodic (see the discussion of ergodicity and nonergodicity of random systems in Sec. 1.4).

According to Eqs. (1.2.1) or (1.2.2), at $t_1 = t_2$ the correlation function is simply the variance, i.e., the average fluctuation squared, $\psi_x(t, t) = \langle (\delta x)^2 \rangle$. As $|t_1 - t_2| \to \infty$ the correlation function in stable systems falls off to zero. Let the fluctuation in the i-th system of the ensemble at instant t_1 be $\delta x_i(t_1)$. If

the instant $t_2 > t_1$ but is close enough to t_1, the fluctuation in each system of the ensemble has no time to change its value appreciably, and the fluctuation $\delta x_i(t_2)$ most probably has the same sign as $\delta x_i(t_1)$. Therefore, the product $\delta x_i(t_1)\delta x_i(t_2) > 0$ in almost all systems of the ensemble. As the instants t_1 and t_2 become farther apart, the number of those systems in which the sign of $\delta x_i(t_2)$ is opposite to the sign of $\delta x_i(t_1)$ approaches the number of systems in which these signs coincide. Thus, in the sum of Eq. (1.2.1) besides positive terms negative ones also arise. At large enough $|t_1 - t_2|$ positive and negative summands are equally probable and $\psi_x(t_1 - t_2) \to 0$.

The decay of $\psi_x(t_1 - t_2)$ to zero as $|t_1 - t_2| \to \infty$ may be viewed as if the system 'forgets' its 'initial' fluctuation. In each physical system there is a characteristic time of 'forgetting' the initial fluctuation. This time is usually called the relaxation time. In real systems a number of kinetic processes (they may even be coupled) and, consequently, a number of relaxation times exist. Moreover, in many complex systems a continuous spectrum of such times even exists.

The correlation function can be written in terms of the two-dimensional probability density (Eq. 1.1.7):

$$\psi_x(t_1, t_2) = \int dx_1 \, dx_2 \, \delta x_1 \, \delta x_2 \, w_2(x_1, t_1; x_2, t_2). \qquad (1.2.3)$$

Here $\delta x_1 = x_1 - \langle x(t_1)\rangle$, $\delta x_2 = x_2 - \langle x(t_2)\rangle$.

The two-dimensional probability density may be represented in the form of the product:

$$w_2(x_1, t_1; x_2, t_2) = w_1(x_1, t_1) P(x_2, t_2|x_1, t_1). \qquad (1.2.4)$$

Here $w_1(x, t)$ is the one-dimensional probability density of the random quantity $x(t)$ at the instant t, $P(x_2, t_2|x_1, t_1)dx_2$ is the conditional probability of the quantity x at time t_2 being in the interval between x_2 and $x_2 + dx_2$, if in the previous instant t_1 the value of x was x_1. Let us denote by

$$\langle \delta x(t_2)|\delta x_1, t_1\rangle = \int dx_2 \delta x_2 P(\delta x_2, t_2|\delta x_1, t_1) \qquad (1.2.5)$$

the mean value of fluctuation δx at the instant t_2 under the condition that at a previous instant t_1 the value of the random quantity was x_1. In these notations

$$\psi_x(t_1, t_2) = \int dx_1 \delta x_1 w_1(\delta x_1, t_1)\langle \delta x(t_2)|\delta x_1, t_1\rangle. \qquad (1.2.6)$$

The quantity $\langle \delta x(t)|\delta x_0, t_0\rangle$ describes the average evolution (usually, a decay) of the fluctuation, the initial value of which at the instant t_0 is δx_0. Note that the absolute value of a fluctuation, for a given initial value, may

either decrease or increase in time. However, *on average* it usually (but not always) decreases for any given initial value. The correlation function is the value of $\delta x_1 \langle \delta x(t_2) | \delta x_1, t_1 \rangle$ averaged over the 'initial' values δx_1.

Let us select those systems of the ensemble in which the fluctuation at an instant t_1 is close to a definite value δx_1. After t_1, eventually the fluctuations in different systems (of the initially selected ones) will be quite different (they are randomized) and their distribution will be the same as that of the fluctuations in the entire ensemble. However, the mean value of the fluctuation for the entire ensemble is zero. Thus, the conditional average value of the fluctuation tends to zero:

$$\lim_{(t_2 - t_1) \to \infty} \langle \delta x(t_2) | \delta x_1, t_1 \rangle = 0.$$

In the case of a Gaussian random process the conditional mean fluctuation $\langle \delta x(t) | \delta x_1, 0 \rangle$ and the correlation function are connected by a simple equation. Substitution of $P(\delta x_2, t_2 | \delta x_1, t_1) = w_2(\delta x_1, t_1; \delta x_2, t_2)/w_1(x_1, t_1)$ for a Gaussian process (Eq. (1.1.16)) into Eq. (1.2.5) yields after integration over δx_2:

$$\frac{\langle \delta x(t) | \delta x_0, 0 \rangle}{\delta x_0} = \frac{\psi_x(t)}{\langle (\delta x)^2 \rangle}. \tag{1.2.7}$$

This relationship means that, for a Gaussian process, the time-dependence of the conditional mean fluctuation is the same as that of the correlation function at *any* initial fluctuation δx_0. If $\psi_x(t)$ is a monotonically decreasing function of $|t|$, the fluctuation $|\delta x(t)|$ at $t > 0$ *on average* decreases at any, arbitrarily small, $|\delta x(0)|$. Of course, this behavior takes place only on average, the absolute value of a random fluctuation may grow and exceed its value at a preceding instant of time.

Several random quantities $x_\alpha(t)$ ($\alpha = 1, \dots, M$) are often coupled (correlated) with one another. The correlation function is then a matrix:

$$\psi_{\alpha\beta}(t_1, t_2) = \langle \delta x_\alpha(t_1) \delta x_\beta(t_2) \rangle. \tag{1.2.8}$$

Any function with $\alpha = \beta$ is called an auto-correlation function, those with $\alpha \neq \beta$ are the cross-correlation functions of different physical quantities.

The above definition of the correlation function is meaningful only if the quantum-mechanical effects are insignificant and the random quantity $x(t)$ may be considered as a classical quantity (c-number). In quantum mechanics, a quantum-mechanical operator $\hat{x}$ is associated with any physical quantity x. Several representations (pictures) of quantum mechanics are known. They are, of course, equivalent. In the Schrödinger representation the wave functions of the system's states depend on time but the operators do not.

In the Heisenberg representation, on the contrary, the wave functions (state vectors) are independent of time, and the entire dependence on time t is transferred to the operators $\hat{x}(t)$, which obey the Heisenberg equation of motion:

$$\frac{\partial \hat{x}(t)}{\partial t} = \frac{i}{\hbar}(\hat{H}\hat{x} - \hat{x}\hat{H}). \tag{1.2.9}$$

Here, $\hat{H}$ is the operator of the system's energy (Hamiltonian) which is also expressed in terms of Heisenberg operators depending on the same time t, $\hbar = h/2\pi$, h is Planck's constant.

The correlation function $\psi_x(t_1, t_2)$ is more conveniently written in terms of Heisenberg operators of the fluctuating quantities. Obviously, it is expressed in terms of the product of two Heisenberg operators taken at instants t_1 and t_2, i.e., $\hat{x}(t_1)\hat{x}(t_2)$. However, the correlation function is an observable (measurable) physical quantity, and its values must be real. According to general rules of quantum mechanics, the corresponding operator of the correlation function must be Hermitian. The symmetrical combination,

$$\frac{1}{2}\{\hat{x}(t_1), \hat{x}(t_2)\} \equiv \frac{1}{2}(\hat{x}(t_1)\hat{x}(t_2) + \hat{x}(t_2)\hat{x}(t_1)), \tag{1.2.10}$$

meets this condition (Ekstein & Rostoker, 1955).

As in classical systems, where the correlation function is an average of the product $x(t_1)x(t_2)$, for quantum-mechanical systems it is equal to the expectation (mean value) of the operator given by Eq. (1.2.10). In general, the state of a quantum-mechanical system is determined by a density matrix $\hat{\rho}$, not by a wave function as a system in a 'pure' quantum-mechanical state. The expectation of a physical quantity, the operator of which is $\hat{x}$, equals $\langle \hat{x}(t) \rangle = \text{Tr}(\rho \hat{x}(t))$. Here, Tr denotes the trace (sum of the diagonal matrix elements) of the operator in brackets. In these notations, the correlation function may be written in the form:

$$\psi_x(t_1, t_2) = \frac{1}{2}\langle \{\hat{x}(t_1), \hat{x}(t_2)\} \rangle. \tag{1.2.11}$$

It follows from the very definition of $\psi_x(t_1, t_2)$ that it is a symmetrical function of the instants t_1 and t_2, i.e.,

$$\psi_x(t_1, t_2) = \psi_x(t_2, t_1). \tag{1.2.12}$$

If the system is a stationary one, the correlation function is a function of the difference $t_1 - t_2$ only (see above). According to Eq. (1.2.12), it is an even function of this difference.

In a more general case when there are several coupled (correlated) quantities x_α ($\alpha = 1, \ldots, M$), the correlation functions form a matrix (compare with

Eqs. (1.2.8) and (1.2.11)):

$$\psi_{\alpha\beta}(t_1, t_2) = \frac{1}{2}\langle(\hat{x}_\alpha(t_1)\hat{x}_\beta(t_2) + \hat{x}_\beta(t_2)\hat{x}_\alpha(t_1))\rangle. \tag{1.2.13}$$

Its elements obviously satisfy the relation

$$\psi_{\alpha\beta}(t_1, t_2) = \psi_{\beta\alpha}(t_2, t_1), \tag{1.2.14}$$

i.e., in a stationary system

$$\psi_{\alpha\beta}(t_1 - t_2) = \psi_{\beta\alpha}(t_2 - t_1). \tag{1.2.15}$$

There is another, more profound and less obvious, relationship between the correlation functions $\psi_{\alpha\beta}(t_1, t_2)$. It follows, as does the Onsager's principle of symmetry of kinetic coefficients (Onsager, 1931), from the symmetry of the equations of motion of any system of particles under the time reversal $t \to -t$. If the system under consideration is placed in a magnetic field **B**, this symmetry takes place if simultaneously the magnetic field is also reversed, i.e., **B** replaced by $-\mathbf{B}$. This relationship, when applied to stationary systems, takes the form:

$$\psi_{\alpha\beta}(t_1 - t_2; \mathbf{B}) = \pm\psi_{\alpha\beta}(t_2 - t_1; -\mathbf{B}). \tag{1.2.16}$$

The plus sign corresponds to the case when the physical quantities $x_\alpha(t)$ and $x_\beta(t)$ are either both invariant under reversal of time t, or both change their signs simultaneously. An example of a quantity that changes its sign under time inversion is the velocity of particles. The minus sign corresponds to the case when only one of the two quantities, x_α or x_β, changes its sign.

Obviously, in the absence of magnetic field ($\mathbf{B} = 0$) Eq. (1.2.16) becomes simpler. The correlation functions then are either strictly even or strictly odd functions of the difference $t_1 - t_2$.

Applying Eq. (1.2.16) to small times $t \equiv t_1 - t_2$, one obtains for $\mathbf{B} = 0$ two relationships for the positive and negative signs in the r.h.s. of Eq. (1.2.16), respectively. The first one is

$$\dot{\psi}_{\alpha\beta}(0) = 0, \tag{1.2.17}$$

where $\dot{\psi}$ is the time derivative of ψ, while the second one is

$$\psi_{\alpha\beta}(0) = 0. \tag{1.2.18}$$

Let $\psi_{\alpha\beta}(\omega)$ be the Fourier transform of the correlation function of a stationary system:

$$\psi_{\alpha\beta}(\omega) = \int_{-\infty}^{+\infty} d(t_1 - t_2)e^{i\omega(t_1 - t_2)}\psi_{\alpha\beta}(t_1 - t_2). \tag{1.2.19}$$

Because the correlation functions are real quantities, their Fourier transforms satisfy a relationship:

$$\psi^{*}_{\alpha\beta}(\omega) = \psi_{\alpha\beta}(-\omega). \qquad (1.2.20)$$

Two more relationships for $\psi_{\alpha\beta}(\omega)$ follow immediately from the relationships in Eqs.(1.2.15) and (1.2.16):

$$\psi_{\alpha\beta}(\omega) = \psi_{\beta\alpha}(-\omega) = \psi^{*}_{\beta\alpha}(\omega), \qquad (1.2.21)$$

$$\psi_{\alpha\beta}(\omega;\mathbf{B}) = \pm\psi_{\alpha\beta}(-\omega;-\mathbf{B}) = \pm\psi^{*}_{\alpha\beta}(\omega;-\mathbf{B}). \qquad (1.2.22)$$

The second Eq. (1.2.21) follows from Eq. (1.2.20). This implies that the matrix $\psi_{\alpha\beta}(\omega)$ is a Hermitian one. Equation (1.2.22) means that in the absence of a magnetic field the functions $\psi_{\alpha\beta}(\omega)$ are either purely real or purely imaginary, depending on the symmetry of the quantities $\hat{x}_{\alpha}(t)$ and $\hat{x}_{\beta}(t)$ under time reversal, see above.

The correlation of fluctuations at two different points $\mathbf{r}_1$ and $\mathbf{r}_2$ in a medium is negligible if the distance $R = |\mathbf{r}_1 - \mathbf{r}_2|$ between these points significantly exceeds some characteristic length L_c, which is called the correlation length. At $R \gg L_c$ the correlation function is proportional to $\exp(-R/L_c)$. A system, the dimensions L of which are many times greater than L_c, can be called macroscopic. This definition of a macroscopic system differs from the common one: a system with $L \simeq L_c$ or even with $L \ll L_c$ may consist of a huge number of atoms and be macroscopic from the point of view of atomic physics.

Macroscopic systems with dimensions $L \gg L_c$ possess an important property: the correlation function of relative (fractional) fluctuations, $\psi_x(t_1, t_2)/\langle x \rangle^2$, is inversely proportional to the system's volume V. If only two dimensions or one dimension are much greater than L_c, this correlation function is inversely proportional to the area A or the length L of the system, respectively. To prove this statement, one has to thoughtfully divide the system into blocks each of $\simeq L_c$ dimension. Then the fluctuations in different blocks are uncorrelated, and their contributions to the total fluctuation partially cancel one another. If $x(t) = \sum_i \xi_i(t)$, where $\xi_i(t)$ is the fluctuating quantity in the i-th block, then the mean value $\langle x \rangle = \sum_i \langle \xi_i \rangle = N \langle \xi \rangle$ (all N blocks are assumed to be identical). The fluctuation is also a sum over N blocks, $\delta x = \sum_i \delta \xi_i$. The correlation function then equals:

$$\psi_x(t_1, t_2) = \sum_{ij} \langle \delta\xi_i(t_1) \delta\xi_j(t_2) \rangle = \sum_i \psi_{\xi_i}(t_1, t_2) = N\psi_\xi(t_1, t_2). \qquad (1.2.23)$$

The cross-terms with $i \neq j$ dropped out owing to the absence of correlation

between different blocks. The correlation function of relative fluctuations,

$$\frac{\psi_x(t_1,t_2)}{\langle x \rangle^2} = \frac{1}{N} \frac{\psi_\xi(t_1,t_2)}{\langle \xi \rangle^2}, \tag{1.2.24}$$

is inversely proportional to the number of blocks, i.e., to V.

In the case of resistance fluctuations, the calculations are different because the total resistance is, in general, not a sum of the resistances of all blocks, but the result for a macroscopic resistor is the same: $\psi_R(t_1,t_2)/R^2 \propto 1/V$.

Usually only the lowest-order correlation functions are studied, both experimentally and theoretically. The higher-order correlations may also be of interest. As was mentioned in Sec. 1.1, in the case of Gaussian random processes these correlations are expressed in terms of pair correlations defined by Eq. (1.2.1). Specifically, the fourth-order correlation function equals:

$$\begin{aligned} \langle \delta x(t_1) \delta x(t_2) \delta x(t_3) \delta x(t_4) \rangle_{\text{Gauss}} &= \langle \delta x(t_1) \delta x(t_2) \rangle \langle \delta x(t_3) \delta x(t_4) \rangle \\ &+ \langle \delta x(t_1) \delta x(t_3) \rangle \langle \delta x(t_2) \delta x(t_4) \rangle + \langle \delta x(t_1) \delta x(t_4) \rangle \langle \delta x(t_2) \delta x(t_3) \rangle. \end{aligned} \tag{1.2.25}$$

The deviation from this relation indicates that the random process under investigation is not a Gaussian one (Sec. 1.3).

1.3 Spectral density of noise

In order to understand the meaning of the spectral density of noise and to derive the relationship between this quantity and the correlation function, we examine a basic circuit arrangement (Fig. 1.1). Consider a resistor connected to a spectrum analyzer which measures the voltage fluctuations across the resistor. A spectrum analyzer contains a band-pass adjustable-frequency filter with a narrow bandwidth and an output detector that responds to the mean square of the signal. In a real experiment the fluctuation signal is amplified, but we ignore these details. Let the central frequency and the bandwidth of the filter be $\bar{f}$ and Δf, respectively. Let the input fluctuation signal and the signal at the output of the filter be $\delta x(t)$ and $\delta x(t|\bar{f},\Delta f)$, respectively. The latter contains only the frequencies from $f_1 = \bar{f} - \Delta f/2$ to $f_2 = \bar{f} + \Delta f/2$.

The random noise signal measured in a long time $t_m \to \infty$, i.e., the input signal of the analyzer, can be represented as a Fourier integral:

$$\delta x(t) = \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} \delta x(\omega) e^{-i\omega t}. \tag{1.3.1}$$

Because $\delta x(t)$ is a real quantity, the Fourier amplitude $x(-\omega) = x^*(\omega)$. Taking into account this relationship, one can write the Fourier integral as
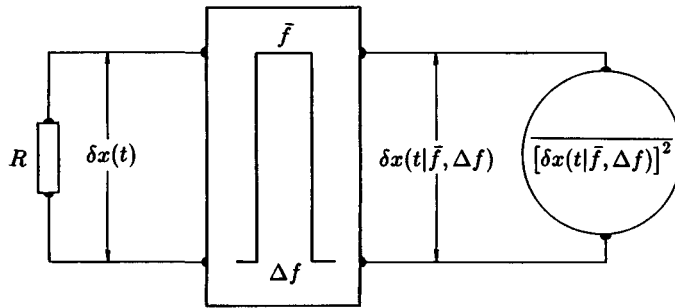
Fig. 1.1 Schematic diagram of a spectral analyzer for spectral density of noise measurement. R is the source of time-dependent fluctuations $\delta x(t)$. The rectangular box in the middle denotes a bandwidth filter with pass-band Δf and central frequency $\bar{f}$. $\delta x(t|\bar{f},\Delta f)$ is the output (filtered) random signal. $\overline{[\delta x(t|\bar{f},\Delta f)]^2}$ is the power of the filtered signal.

an integral over positive frequencies only:

$$\delta x(t) = \int_0^\infty \frac{d\omega}{2\pi}[\delta x(\omega)e^{-i\omega t} + \delta x^*(\omega)e^{i\omega t}]. \tag{1.3.2}$$

Similarly, the filter output signal

$$\delta x(t|\bar{f},\Delta f) = \int_{\bar{\omega}-\Delta\omega/2}^{\bar{\omega}+\Delta\omega/2} \frac{d\omega'}{2\pi}[\delta x(\omega')e^{-i\omega' t} + \delta x^*(\omega')e^{i\omega' t}]. \tag{1.3.3}$$

This signal squared, which is called noise power, depends on time and randomly fluctuates around its mean value $\overline{[\delta x(t|\bar{f},\Delta f)]^2}$ which is time-independent. Assuming that the averaging over the time of measurement, t_m, and over an ensemble of identical noise sources and spectral analyzers give identical results (Sec. 1.4), one can write:

$$\langle[\delta x(t|\bar{f},\Delta f)]^2\rangle = \int_{\bar{\omega}-\Delta\omega/2}^{\bar{\omega}+\Delta\omega/2} \frac{d\omega'}{2\pi}\frac{d\omega''}{2\pi}\langle[\delta x(\omega')e^{-i\omega' t} + \delta x^*(\omega')e^{i\omega' t}] \\ \times [\delta x(\omega'')e^{-i\omega'' t} + \delta x^*(\omega'')e^{i\omega'' t}]\rangle. \tag{1.3.4}$$

To find the result of averaging let us examine in the same way Eq. (1.2.1) for the correlation function of the same random quantity $\delta x(t)$:

$$\psi_x(t_1 - t_2) = \langle x(t_1)x(t_2)\rangle = \int_{-\infty}^{+\infty} \frac{d\omega' d\omega''}{(2\pi)^2}\langle\delta x(\omega')\delta x(\omega'')\rangle e^{-i\omega' t_1 - i\omega'' t_2}. \tag{1.3.5}$$

Because the correlation function of a stationary random process depends only on the difference $t_1 - t_2$, the mean value $\langle\delta x(\omega')\delta x(\omega'')\rangle$ should be nonzero only for $\omega' + \omega'' = 0$. In other words, it is proportional to the delta-function $\delta(\omega' + \omega'')$. The coefficient at this delta-function is obviously

$2\pi\psi_x(\omega')$, where $\psi_x(\omega)$ is the Fourier transform of the correlation function $\psi_x(t_1 - t_2)$. Thus,

$$\langle \delta x(\omega')\delta x(\omega'')\rangle = 2\pi\psi_x(\omega')\delta(\omega' + \omega''). \tag{1.3.6}$$

Upon substituting Eq. (1.3.6) into Eq. (1.3.4) one obtains:

$$\overline{[\delta x(t|\bar{f},\Delta f)]^2} = 2\int_{\bar{f}-\Delta f/2}^{\bar{f}+\Delta f/2} df\,\psi_x(\omega) \approx S_x(\bar{f})\Delta f. \tag{1.3.7}$$

The last equation means that the mean squared signal transmitted through a filter with a sufficiently narrow passband is proportional to the bandwidth Δf. The quantity $S_x(f)$ introduced in the r.h.s. of Eq. (1.3.7), i.e., the mean random signal squared (noise power) per unit frequency band is called the spectral density of noise (or of fluctuations), or simply the spectrum of fluctuations. It follows from Eq. (1.3.7) that:

$$S_x(f) = 2\int_{-\infty}^{+\infty} d(t_1 - t_2)e^{i\omega(t_1 - t_2)}\psi_x(t_1 - t_2) \equiv 2\psi_x(\omega), \tag{1.3.8}$$

i.e., the spectral density is twice the Fourier transform of the correlation function. This relationship is called the Wiener–Khintchine theorem (Wiener, 1930; Khintchine, 1934).

The equation for the spectral density may be represented in a different but equivalent form. Let $\delta x(t)$ be a record of the fluctuations measured during a sufficiently long time interval t_m, say, from $t = -t_m/2$ to $t = t_m/2$. Then:

$$S_x(f) = 2\lim_{t_m\to\infty}\frac{1}{t_m}\left| \int_{-t_m/2}^{t_m/2} dt\, e^{i\omega t}\delta x(t)\right|^2. \tag{1.3.9}$$

In the case of several coupled, i.e., correlated, fluctuating quantities x_α, where $\alpha = 1,\ldots,M$, the spectral density is obviously a matrix, $S_{\alpha\beta}(f)$, which may be defined in two ways. The first is a generalization of Eq. (1.3.9). Let $\delta x_\alpha(t)$ and $\delta x_\beta(t)$ be the records of two different random quantities during the time t_m. Then

$$S_{\alpha\beta}(f) = 2\lim_{t_m\to\infty}\frac{1}{t_m}\int_{-t_m/2}^{t_m/2} dt_1 e^{i\omega t_1}\delta x_\alpha(t_1) \times \int_{-t_m/2}^{t_m/2} dt_2 e^{-i\omega t_2}\delta x_\beta(t_2)$$
$$= 2\psi_{\alpha\beta}(\omega). \tag{1.3.10}$$

By this definition, the spectral density matrix is twice the Fourier transform of the correlation matrix, and its properties are, obviously, the same as those of $\psi_{\alpha\beta}(\omega)$ (Eqs. (1.2.20–22)). For the reader's convenience, these relationships

are reproduced here:

$$S_{\alpha\beta}^{*}(f) = S_{\alpha\beta}(-f), \tag{1.3.11}$$

$$S_{\alpha\beta}(f) = S_{\beta\alpha}^{*}(f), \tag{1.3.12}$$

$$S_{\alpha\beta}(f;\mathbf{B}) = \pm S_{\alpha\beta}^{*}(f;-\mathbf{B}). \tag{1.3.13}$$

These spectral densities are defined on the entire frequency axis, i.e., both for positive and negative frequencies. The nondiagonal elements of $S_{\alpha\beta}(f)$ are, in general, complex quantities.

The spectral density may be defined also in a way analogous to Eq. (1.3.7) as the spectral density of the mean product of filtered signals $\delta x_{\alpha}(\bar{f}, \Delta f)$ and $\delta x_{\beta}(\bar{f}, \Delta f)$:

$$\bar{S}_{\alpha\beta}(f) = \lim_{\Delta f \to 0} (\Delta f)^{-1} \overline{\delta x_{\alpha}(t|\bar{f}, \Delta f)\delta x_{\beta}(t|\bar{f}, \Delta f)}. \tag{1.3.14}$$

One can easily prove that $\bar{S}_{\alpha\beta}(f)$ is a symmetrical combination:

$$\begin{aligned}
\bar{S}_{\alpha\beta}(f) &= \psi_{\alpha\beta}(\omega) + \psi_{\beta\alpha}(\omega) = \frac{1}{2}\left[S_{\alpha\beta}(f) + S_{\beta\alpha}(f)\right] \\
&= \frac{1}{2}\left[S_{\alpha\beta}(f) + S_{\alpha\beta}(-f)\right] = \frac{1}{2}\left[S_{\alpha\beta}(f) + S_{\alpha\beta}^{*}(f)\right].
\end{aligned} \tag{1.3.15}$$

Although this spectral density, $\bar{S}_{\alpha\beta}(f)$, is defined for positive frequencies, it may be considered as an even function of f (Eq. (1.3.15)). It is always, even in a magnetic field $\mathbf{B} \neq 0$, a real symmetric matrix.

It follows from Eq. (1.3.7) that in the limit of very wide bandwidth which encompasses all significant frequencies, the mean squared signal is just the variance of the fluctuations:

$$\langle (\delta x)^2 \rangle = \psi_x(t_1 - t_2 = 0) = \int_0^{+\infty} df\, S_x(f) \tag{1.3.16}$$

It means that the integral of the spectral density over all positive frequencies is exactly the variance of the noise.

The correlation function $\psi_x(t_1 - t_2)$ and the corresponding spectral density $S_x(f)$ are of the lowest order. The measurement of higher-order correlation functions and spectral densities may be useful as well, especially in the case of non-Gaussian random processes, when these functions can not be expressed in terms of $\psi_x(t_1 - t_2)$ and $S_x(f)$, respectively. As was mentioned above, the noise power, $[\delta x(t|\bar{f}, \Delta f)]^2$, of the noise signal, transmitted through a filter with bandwidth Δf and central frequency $\bar{f}$, randomly fluctuates in time around its mean value

$$\overline{[\delta x(t|\bar{f}, \Delta f)]^2} = \int_{\bar{f}-\Delta f/2}^{\bar{f}+\Delta f/2} df\, S_x(f). \tag{1.3.17}$$

The correlation function of these fluctuations may be called 'the second correlation function' $\psi_x^{(2)}(t_1, t_2|\bar{f}, \Delta f)$ (the term 'second' may be used only for simplicity, this function is expressed in terms of a product of *four* fluctuations $\delta x(t)$). The corresponding spectral density is called 'the second spectral density':

$$S_x^{(2)}(f|\bar{f}, \Delta f)$$
$$= 2 \lim_{t_m \to \infty} \frac{1}{t_m} \Big| \int_{-t_m/2}^{t_m/2} dt\, e^{i\omega t} \Big[[\delta x(t|\bar{f}, \Delta f)]^2 - \overline{[\delta x(t|\bar{f}, \Delta f)]^2} \Big] \Big|^2. \quad (1.3.18)$$

This spectral density can be represented in terms of the fourth-order correlation function $\psi_x^{(4)}$ as follows:

$$S_x^{(2)}(f|\bar{f}, \Delta f) = 16\pi \int_{\bar{\omega}-\Delta\omega/2}^{\bar{\omega}+\Delta\omega/2} \frac{d\omega_1 d\omega_2}{(2\pi)^3} \Big[\psi_x^{(4)}(\omega, \omega_1, \omega_2) \\ - 2\pi\delta(\omega)\psi_x(\omega_1)\psi_x(\omega_2) \Big]. \quad (1.3.19)$$

Here

$$\psi_x^{(4)}(\omega, \omega_1, \omega_2) = \int_{-\infty}^{+\infty} d\theta d\theta_1 d\theta_2 e^{i[\omega\theta + (\omega_1 - \omega/2)\theta_1 + (\omega_2 + \omega/2)\theta_2]} \\ \times \langle \delta x(t_1)\delta x(t_1')\delta x(t_2)\delta x(t_2') \rangle, \quad (1.3.20)$$

and $\theta = (t_1 + t_1' - t_2 - t_2')/2$, $\theta_1 = t_1 - t_1'$, $\theta_2 = t_2 - t_2'$.

If the random process is Gaussian, the fourth-order correlation function can be expressed in terms of the lowest-order (ordinary) one by decoupling the mean product of four fluctuations (Eq. (1.3.20)) in all possible mean products of pairs of fluctuations (Eq. (1.2.25)). For Gaussian processes the correlation function of $[\delta x(t|\bar{f}, \Delta f)]^2$ and the spectral density $S^{(2)}$ equal, respectively,

$$\big[\psi_x^{(2)}(t_1, t_2|\bar{f}, \Delta f)\big]_{\text{Gauss}} = 2\Big[\overline{\delta x(t_1|\bar{f}, \Delta f)\delta x(t_2|\bar{f}, \Delta f)}\Big]^2,$$
$$\big[S_x^{(2)}(f|\bar{f}, \Delta f)\big]_{\text{Gauss}} = 2 \int_{\bar{f}-\Delta f/2+f/2}^{\bar{f}+\Delta f/2-f/2} dv S_x(v + f/2)S_x(v - f/2). \quad (1.3.21)$$

By the very meaning of the second spectrum, its frequencies are much smaller than the lowest limit of the frequency range of the first spectrum, i.e., $f/2 \ll f_1 = \bar{f} - \Delta f/2$. Hence, as follows from Eq. (1.3.21), the second spectral density of a Gaussian random process is independent of frequency f, i.e., its second spectrum is always 'white'. Thus, if $S^{(2)}(f)$ depends on f, this dependence is directly connected with deviations of the noise statistics from purely Gaussian (Weissman, 1988, 1993).

The difference

$$\tilde{S}_x^{(2)}(f|\bar{f}, \Delta f) - 1 \equiv \frac{S_x^{(2)}(f|\bar{f}, \Delta f)}{[S_x^{(2)}(f|\bar{f}, \Delta f)]_{\text{Gauss}}} - 1, \qquad (1.3.22)$$

is a measure of the deviation of the random process from a Gaussian one. The measurements of the second spectral density have been successfully employed to find out the statistical properties of $1/f$ noise (Weissman, 1988, 1993 and Secs. 8.2.2 and 8.6).

1.4 Ergodicity and nonergodicity of random processes

Each random quantity and each random process are characterized by some mean (average), nonrandom quantities. They include the moments (mean value, variance, etc.), the distribution functions, the correlation functions and so on. As we have already mentioned above in Secs. 1.1 and 1.2, two methods of averaging random quantities and obtaining these characteristics are possible. According to the first method, the data (records) of the random quantity measurement during a sufficiently long time t_m are averaged. Just this method is usually used in experiments. The equations for the mean value $\bar{x}$ and for the correlation function found by this method are given by Eq. (1.1.18) and (1.2.2), respectively.

The second method is averaging over an ensemble of identical systems, that is, systems having identical conditions for the random process considered and identical methods of its measurement. The equations for the mean value $\langle x \rangle$ and the correlation function $\psi_x(t_1, t_2)$ using this method of averaging are Eqs. (1.1.19) and (1.2.1), respectively.

Does each of these procedures of averaging yield a definite mean value of the random quantity being averaged? Are the results obtained by the two methods identical? The systems and random processes for which the answers are positive are called ergodic. Otherwise the system or the random process is nonergodic. It is important to understand what physical phenomena underlie the property of ergodicity of many physical systems, and why some other systems, or even the same systems under different conditions, are nonergodic.

The thermal motion of any statistical system may be viewed as incessant transitions between the microstates of this system. If the interaction of the system with surrounding bodies is sufficiently weak, each such microstate is an eigenstate of the Hamiltonian of the system itself. Owing just to the weak interaction of the system with other bodies, having often much greater thermal capacity ('thermal bath'), the energies E_n of the states which the

system happens to visit in the course of its thermal motion are, in general, different.

Let us consider the simplest case of an ensemble of identical equilibrium systems each of which is interacting with a 'thermal bath' with absolute temperature T. The probability that a system belonging to this ensemble happens to be in its m-th state is given by the well known Gibbs distribution:

$$w_m = \frac{\exp(-E_m/k_B T)}{\sum_n \exp(-E_n/k_B T)} \tag{1.4.1}$$

Here k_B is the Boltzmann constant. The mean value of a random quantity x is equal to:

$$\langle x \rangle = \sum_m w_m x_m, \tag{1.4.2}$$

where x_m is the value of x in the m-th microstate of the system, the sum is over all such microstates. Obviously, the probability of finding the system in states, the energies E_m of which significantly differ from the mean energy $\langle E \rangle$ of the system by an amount which exceeds the fluctuations of energy, is very small, and these microstates practically do not contribute to the mean value $\langle x \rangle$. If the fluctuations of system's energy are neglected, its motion can be considered as confined to a surface of constant energy in the space of microstates (this 'space' is called phase space). The averaging in Eq. (1.4.2) is then over microstates belonging to this surface only (over so called microcanonical ensemble).

The existence of a definite single mean value of averaging and the equivalence of the results of averaging over time and over an ensemble ($\bar{x} = \langle x \rangle$), are based on a definite assumption. It is assumed that during the time t_m of measurement the system visits, if not all the allowed microstates (it is, perhaps, not necessary), a representative enough sample of microstates which make up the major contribution to the mean value. According to this assumption, the measurement during such a finite time t_m gives with sufficient accuracy the same result as for much longer times of measurement, even $t_m \to \infty$. This assumption is really valid if the time t_m is much greater than the characteristic times τ of passing through all groups of representative states. The inequality

$$t_m \gg \tau \tag{1.4.3}$$

is just the condition of ergodicity of the system.

However, there are many systems for which the surface of allowed states in the phase space consists of several or even a great number of disconnected parts. These parts are separated by some barriers (in particular, high energy

barriers) which cannot be surmounted during the given time t_m of experiment and, moreover, during any reasonable time of experiment. The parts of the energy surface and the states close to this surface (within thermal fluctuations of the energy) are called 'valleys', or 'components' (Palmer, 1982). If the barriers separating the valleys are sufficiently high, the time of intervalley transitions τ_{interv} is much greater than any possible time t_m of the experiment. In this case the system is confined within only one valley in which it happened to be 'placed' by the procedure of the system's preparation. The measured mean value $\bar{x}$ of the random quantity is a characteristic of this valley only. The result of averaging may be different if the measurements are performed on a system which happened to be in a different valley. Let the quantity x be averaged over an ensemble which includes systems in different valleys, for instance, over the Gibbs ensemble, Eq. (1.4.1), which takes into account all valleys. If the measured mean quantities $\bar{x}$ corresponding to different valleys are different, the ensemble mean value $\langle x \rangle$ obviously differs from any specific measured $\bar{x}$, and the ergodicity is broken.

A simple example is a ferromagnet (Palmer, 1982). It is well known to be in a paramagnetic state at temperatures T higher than the critical Curie temperature T_c . The relaxation ('intermixing') times in this state are sufficiently small, therefore the ergodicity condition is easily satisfied. However, below T_c a spontaneous magnetic moment **M** is established. In the simplest case, the ferromagnet can be in two states with identical free energies and absolute values |**M**| but with opposite directions of the magnetic moment **M**. The transition from one state to another can be achieved if a domain wall crosses the magnet. The free energy of formation of the domain wall is proportional to its surface. The energy which is required to overturn one elementary spin magnetic moment is of the order of $k_B T_c \sim 10^{-2}$ eV. If the domain surface contains N spins, the energy of domain wall formation is $\sim N k_B T_c$. The probability of such fluctuation is proportional to $\exp(-N k_B T_c / k_B T)$, where the actual temperature $T < T_c$. Even for a magnet of microscopic dimensions ($N \sim 10^3$) the time τ_{interv} for domain wall formation and for crossing the barrier between the two states of the magnet becomes extremely long and may be greater by many orders than even the age of the Universe. The magnetic moment measured in a sample of a ferromagnet in any accessible time t_m will be $\mathbf{M} \neq 0$. But the mean value $\langle \mathbf{M} \rangle$ obtained by averaging over the full Gibbs ensemble, Eq. (1.4.1), in which systems with **M** of opposite signs, having identical energies, are equally probable, is zero (this ensemble corresponds to $t_m \to \infty$). Strictly speaking, this system is nonergodic. However, such kind of nonergodicity is easily removed by averaging over an ensemble of

$$\vdash\!\!-\!\!-\!\!-\!\!-\!\!\blacksquare\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!-\!\!\dashv\!\!-\!\!-\!\!-\!\!-$$

0            $\tau_{\min}$ t_{m} $\tau_{\max}$

Fig. 1.2 Schematic diagram of a continuous spectrum of relaxation times τ. $\tau_{\min}$ and $\tau_{\max}$ are the minimal and maximum relaxation times in the system. t_m is the time of noise measurement (averaging).

systems with only one direction of **M**, i.e., over a 'restricted ensemble' (Palmer, 1982).

Another kind of nonergodicity is encountered in disordered macroscopic systems. A known example is glass, e.g., amorphous SiO_2. It is in a metastable state (according to Eq. (1.4.1), at low temperatures it should exist in a crystalline form). The relaxation time of its transition to a stable equilibrium crystalline state is finite but huge. However, for many kinds of measurement, a macroscopic sample of glass may be considered as an ergodic system.

Another example is spin glasses (Sec. 8.6 and references to Ch. 8). Their study has shown that the phase space of these systems consists of a large or even a huge number of valleys. In disordered systems these valleys are separated by energy barriers of various heights: from small microscopic barriers up to barriers the heights of which are macroscopic energies, i.e., a hierarchy of barriers exists. As the time of intervalley transition τ_{interv} depends on the height of the barrier exponentially, there is an extremely wide ('exponentially wide') continuous spectrum of the times τ_{interv} from some minimal $\tau_{\min}$ up to some $\tau_{\max}$ which, as in the above mentioned ferromagnets, may exceed the age of the Universe by many orders of magnitude (Fig. 1.2). The condition $t_m \gg \tau_{\min}$ is usually easily satisfied. However, it is absolutely impossible to satisfy the condition $t_m \gg \tau_{\max}$. At any accessible time of the experiment, t_m is within the continuous spectrum of relaxation times. The system is an equilibrium and ergodic one for those states, the transitions between which require time $\tau < t_m$, but is nonergodic for states with $\tau \gg t_m$. Since there are such groups of states between which equilibrium is not reached, these systems are, in principle, not only nonergodic but also nonequilibrium. Some of the measured quantities depend therefore on the duration t_m of the measurement procedure and on the 'waiting' time t_w between establishing the conditions of the experiment and the start of the measurements.

The problem of $1/f$ (flicker) noise observed in many systems is connected with the problem of nonergodicity of these systems (Ch. 8).
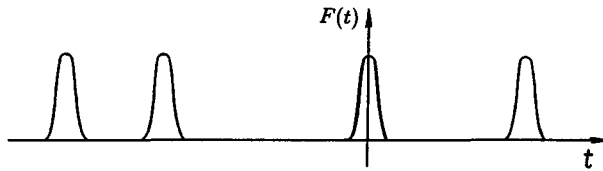
Fig. 1.3 Sequence of random pulses. $F(t)$ is the shape of a pulse.

1.5 Random pulses and shot noise

Random sequences of pulses, or pulsed random processes, are encountered very often. In the simplest case, the pulses have identical form, amplitude, and duration, and the only random quantity is the instant at which a pulse starts (Fig. 1.3). One well known example is the current in the anode circuit of the thermionic tube. Each electron emitted from the cathode and moving in the vacuum space between cathode and anode creates in the anode circuit a pulse $i(t)$, where t is time. The integral of the function $i(t)$ over time is equal to the electron charge e. Other examples are the impacts of gas molecules on the wall of the container, or the impacts of the molecules of the liquid on the macroscopic (Brownian) particle (Secs. 1.1, 1.9, and the books by Rytov, 1976 and Rytov *et al.*, 1987, 1988).

Let the index k number the pulses, and t_k be the instant corresponding to the start of the k-th pulse (or to its maximum). Each pulse can be described by some function $F(t-t_k;\mathbf{a}_k)$ which, in general, depends, if the pulses are not identical, on a finite set of random parameters $\mathbf{a}_k = (a_{k1},\ldots,a_{km})$ considered as a m-component vector. These parameters include the amplitude of the pulse, its duration, etc. Let the measured random quantity be the sum:

$$x(t) = \sum_k F(t - t_k;\mathbf{a_k}). \qquad (1.5.1)$$

The random quantities t_k and $\mathbf{a}_k$ are usually assumed to have the following statistical properties:

1. The random quantities t_k and $\mathbf{a}_k$ are statistically independent for different k and their distribution functions do not depend on k.
2. The probability of the time t_k being in the range between t and $t + dt$ is proportional to dt, independent of t, and equal to $n_1 dt$, where n_1 is the constant mean frequency of pulses.

Suppose that the measurement of the pulsed random process lasts a time t_m. The number n of pulses in this time interval is a random quantity fluctuating around the mean value $\bar{n} = n_1 t_m$. Let us split the entire ensemble

of identical systems exhibiting the pulsed random process into subensembles, each with a definite number $n = 0, 1, \ldots$ of pulses. As the pulses are independent, the fraction (probability) of systems with definite n is given by the Poisson distribution (Eq. (1.1.17)). Even at a fixed number n of pulses, when the sum in Eq. (1.5.1) contains exactly n terms, the variable x is a random quantity because the sequences of pulses are random. According to Eq. (1.1.11), its characteristic function is a product of n characteristic functions, each of which equals:

$$\phi_{x_k}(u) = \overline{\exp(iux_k)} = \int da\mathbf{w_a}(\mathbf{a}) \int_{-t_m/2}^{t_m/2} \frac{dt_k}{t_m} \exp[iuF(t - t_k; \mathbf{a})]. \qquad (1.5.2)$$

Here $w_\mathbf{a}(\mathbf{a})$ is the probability density function of the parameters $\mathbf{a}$ of the pulses. The second integral in Eq. (1.5.2) is over instants t_k at which the pulses start. The value of the integral is independent of both the number k and the time t.

Both the characteristic function $\phi_x(u|n) = [\phi_{x_k}(u)]^n$ and the corresponding probability density $w(x|n)$ are *conditional* ones. They each relate to a subensemble (see above) with a definite number n of pulses in the time interval between $-t_m/2$ and $t_m/2$. The total, i.e., unconditional characteristic function is obtained by averaging over all subensembles by using the Poisson probability (Eq. (1.1.17)) for the number of pulses that occur in the time t_m. The result reads:

$$\phi_x(u) = \exp\left\{ n_1 \int da\mathbf{w_a}(\mathbf{a}) \int_{-\infty}^{+\infty} d\theta \left[\exp(iuF(\theta, \mathbf{a})) - 1 \right] \right\}. \qquad (1.5.3)$$

Here $\theta = t - t'$. The replacement of the limits $t \pm t_m/2$ by $\pm\infty$ uses the steep enough drop of the pulse $F(t - t_k; \mathbf{a})$ at its edges.

The mean value and the variance can be expressed in terms of the characteristic function (Eq. (1.5.3)):

$$\langle x \rangle = \int_{-\infty}^{+\infty} dx \, xw(x) = \left(\frac{\partial\phi_x(u)}{\partial iu} \right)_{u=0} = n_1 \int da\mathbf{w_a}(\mathbf{a}) \int_{-\infty}^{+\infty} d\theta F(\theta, \mathbf{a}),$$

$$\langle (\delta x)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2 = \left(\frac{\partial^2 \ln\phi_x(u)}{\partial(iu)^2} \right)_{u=0} \qquad (1.5.4)$$

$$= n_1 \int da\mathbf{w_a}(\mathbf{a}) \int_{-\infty}^{+\infty} d\theta F^2(\theta, \mathbf{a}).$$

When all pulses are identical there is no integration over $\mathbf{a}$. For this case Eq. (1.5.4) has been derived by Campbell (1909).

The correlation function of the stationary pulsed random process can be found in the same way as the characteristic function, Eq. (1.5.3):

$$\psi_x(t_1 - t_2) = \langle \delta x(t_1) \delta x(t_2) \rangle = n_1 \int da w_{\mathbf{a}}(\mathbf{a}) \int_{-\infty}^{+\infty} d\theta F(\theta;\mathbf{a}) F(\theta + t_1 - t_2;\mathbf{a}).$$
(1.5.5)

Obviously, this equation at $t_1 \to t_2$ reduces to Eq. (1.5.4) for the noise variance.

According to the Wiener–Khintchine theorem (Eq. (1.3.8)), the corresponding spectral density is

$$S_x(f) = 2n_1 \int da w_{\mathbf{a}}(\mathbf{a}) |F(\omega;\mathbf{a})|^2,$$
(1.5.6)

where $F(\omega;\mathbf{a})$ is the Fourier transform of $F(t;\mathbf{a})$. For identical pulses, when no integration over $\mathbf{a}$ has to be performed, this equation was derived by Carson (1931).

The Fourier transform $F(\omega)$ at $\omega = 0$ is equal to the integral over the pulse. It may be denoted by q because in the case of current pulses it is exactly the charge of a particle. Hence the spectral density at low frequencies, many times lower than the inverse duration of each pulse, equals:

$$S_x(0) = 2q^2 n_1. \qquad q = \int_{-\infty}^{+\infty} dt F(t).$$
(1.5.7)

As was mentioned above, the current in the anode circuit of a thermionic tube is a pulsed random process. Schottky (1918), who was the first to examine it, called it shot noise. The pulses are uncorrelated because so are the events of electron emission from the cathode. The integral over a single pulse is equal to the absolute value of the electron charge e. The frequency of the pulses n_1 is the number flow, J, of particles per unit time, $J = I/e$, where I is the mean current. In the case of shot noise, the spectral density at low frequencies equals (Schottky, 1918):

$$S_I(0) = 2eI = 2e^2 J.$$
(1.5.8)

Measuring $S_I(0)$ and the mean current I one is able to determine the electron charge e. A number of measurements of e by this method are known.

Let us write down the equation for the spectral density in the case when the pulses are identical and rectangular. If the duration of a pulse is τ, its height is q/τ. In this case

$$F(\omega) = \frac{q}{\tau} \int_{-\tau/2}^{\tau/2} dt e^{i\omega t} = q \frac{\sin(\omega\tau/2)}{\omega\tau/2}.$$
(1.5.9)

According to Eq. (1.5.7),

$$S_x(f) = 2n_1 q^2 \left[\frac{\sin(\omega\tau/2)}{\omega\tau/2} \right]^2 \tag{1.5.10}$$

The spectral density significantly deviates from its low-frequency value at frequencies $f \simeq 1/\tau$. Obviously, the same is true even if the form of a single pulse deviates from the rectangular one: the characteristic duration of a pulse then plays the role of τ.

1.6 Markov processes; general theory

The notion of Markov, or Markovian, random processes, as well as the Markov approach in the theory of random processes are widely used (including this book) and applied to many systems. This notion and this approach are based on an approximation. As any well substantiated physical approximation it is, in its turn, based on definite properties of random processes in real physical systems (for comprehensive accounts of Markov processes for physicists see Lax, 1960; Rytov, 1976; Rytov *et al.*, 1987, 1988).

Let us examine, for instance, the temperature fluctuations in a macroscopic specimen due to heat exchange with the surrounding bodies. Since the temperature determines the distribution in energies of the specimen's states (through the Gibbs distribution), its fluctuation is accompanied by a change of a huge number of microscopic parameters characterizing the specimen: parameters of the electron system, phonons, and so on. Therefore the characteristic times of temperature measurement and monitoring (the duration of one measurement, the intervals between consecutive measurements) must be and, in fact, are many times greater than the times that are necessary for the microscopic parameters to adjust the change of the temperature. As on the time scale of temperature measurements the microscopic parameters can be considered as completely adjusted to the instantaneous value of the temperature, the conditional probability of the temperature fluctuation at instant t_2 taking a value δT_2, if at the preceding measurement at instant t_1 it happened to be δT_1, depends only on this last quantity and does not depend on the previous history of temperature fluctuations before t_1, which is 'forgotten'.

Some defects and impurities in solids are known to be mobile. They are moving by hops. Each such hop usually requires surmounting of an activation barrier. Therefore, the rate of hops of any individual defect is many times smaller than the frequency v of attempts to surmount the barrier ($v \sim 10^{11} - 10^{14}$ s^{-1}). After each hop, the crystal and the defect in its new position release the activation energy to the lattice vibrations (phonons) and

reach an equilibrium state. The equilibration time, which can be considered as a 'microscopic' one, is usually smaller by many orders of magnitude than the time between two consequent hops. That is why the defect 'forgets' its previous history long before the next hop occurs, the probability of which depends therefore on its present position only. Thus, the hopping motion of defects in a solid is usually a Markov random process.

Let us assume that the random function of time $x(t)$ is known (e.g., as a result of measurements) at some n instants $t_1 < t_2 < \ldots < t_n$, and the time intervals between these instants are many times greater than the times that can be considered as microscopic for the random process $x(t)$. A random process $x(t)$ can be characterized by the probability $w_n(x_1, t_1; \ldots; x_n, t_n) \times dx_1 \cdots dx_n$ that the random quantity $x(t)$ is at time t_1 in the range $(x_1, x_1 + dx_1)$, at time t_2 in the range $(x_2, x_2 + dx_2)$, and so on (Eq. (1.1.3)). The probability density functions w_n and w_{n-1} may be related by an equation which is, in fact, a definition of the conditional probability function P_n (Eq. (1.2.4)):

$$
\begin{aligned}
&w_n(x_1, t_1; \ldots; x_n, t_n) \\
&= w_{n-1}(x_1, t_1; \ldots; x_{n-1}, t_{n-1}) P_n(x_n, t_n | x_1, t_1; \ldots; x_{n-1}, t_{n-1}).
\end{aligned}
\tag{1.6.1}
$$

The quantity $P_n dx_n$ is the conditional probability that $x(t_n)$ is in the range $(x_n, x_n + dx_n)$ *if* at preceding instants t_i $(i = 1, \ldots, n-1)$ the random quantity took the given values x_i. Thus, in general, the conditional probability P_n of the random quantity to be in the state x_n at instant t_n depends on the entire preceding history of the process (on the entire path it has run), that is, on a great number of its previous states.

However, real random processes often possess a fundamental property: the conditional probability P_n at any number n depends not on all values $x_1, \ldots, x_{n-1}$ at all previous times $t_1, \ldots, t_{n-1}$ but only on the value x_{n-1} at the last instant t_{n-1} which precedes t_n:

$$
\begin{aligned}
&P_n(x_n, t_n | x_1, t_1; \ldots; x_{n-1}, t_{n-1}) \\
&= P_2(x_n, t_n | x_{n-1}, t_{n-1}) \equiv P(x_n, t_n | x_{n-1}, t_{n-1}).
\end{aligned}
\tag{1.6.2}
$$

This equation means that for those methods of monitoring the random quantity which are used in the experiment and are appropriate to this random quantity, the system which determines the randomness of $x(t)$ has enough time to forget the previous history of this random quantity. In other words, the system has no long-term memory. The processes, for which the conditional probability P_n satisfies the condition Eq. (1.6.2), have been first investigated by A.A. Markov, and are called Markov (or Markovian) random processes, or Markov chains, or processes without after-effect. Such processes have been illustrated above.

The conditional probability function $P(x, t|x_0, t_0)$ is called the transition probability. In the case of a homogeneous Markov process it depends only on the difference $t - t_0$. The homogeneous process may be nonstationary, but for stationary processes the homogeneity condition is always satisfied. Let us restrict ourselves to stationary processes only.

The definition of a Markov random process may be easily generalized: there may be several, say M, random quantities in the system. This multitude of random quantities may be denoted as an M-component vector $\mathbf{x}(t)$. In this case the transition probability can be written as $P(\mathbf{x}, t|\mathbf{x}_0, t_0)$.

The transition probability $P(\mathbf{x}, t|\mathbf{x}_0, t_0)$ at $t > t_0$ must satisfy a definite consistency condition. Let t' be any intermediate point of time between t_0 and t, i.e., $t_0 < t' < t$. The probability of the transition $\mathbf{x}_0 \rightarrow \mathbf{x}$ is obviously a sum of probabilities of such transitions through all possible values of $\mathbf{x}'$ at the instant t'. Therefore the following equation holds:

$$P(\mathbf{x}, t|\mathbf{x}_0, t_0) = \int d\mathbf{x}' P(\mathbf{x}, t|\mathbf{x}', t') P(\mathbf{x}', t'|\mathbf{x}_0, t_0), \qquad (1.6.3)$$

where $d\mathbf{x}' = dx_1' \cdots dx_M'$. It is called the Smoluchowski equation.

This equation imposes a substantial restriction on the possible form of the transition probability: integration over $\mathbf{x}'$ of the product of two such probabilities has to result also in a transition probability, and this result is independent of the intermediate time t'.

As the total probability of transition to *all* states of the system is equal to unity, the transition probability must satisfy the condition:

$$\int d\mathbf{x} P(\mathbf{x}, t|\mathbf{x}_0, t_0) = 1. \qquad (1.6.4)$$

Let $w(\mathbf{x})$ be the stationary probability density function. If the transition probability $P(\mathbf{x}, t|\mathbf{x}_0, t_0)$ is averaged over all initial values $\mathbf{x}_0$ with the density function $w(\mathbf{x}_0)$, the result must be the value of the probability density function $w(\mathbf{x})$:

$$\int d\mathbf{x}_0 P(\mathbf{x}, t|\mathbf{x}_0, t_0) w(\mathbf{x}_0) = w(\mathbf{x}). \qquad (1.6.5)$$

The physical meaning of $P(\mathbf{x}, t|\mathbf{x}_0, t_0)$ implies that if the difference $t - t_0$ is small as compared with the characteristic time of the random process, the random quantity $\mathbf{x}(t)$ has no time to deviate from the initial value $\mathbf{x}_0$. Therefore

$$\lim_{(t-t_0)\to 0} P(\mathbf{x}, t|\mathbf{x}_0, t_0) = \delta(\mathbf{x} - \mathbf{x}_0). \qquad (1.6.6)$$

Here $\delta(\mathbf{x} - \mathbf{x}_0)$ is the pulse (delta-) function which is nonzero only at $\mathbf{x} = \mathbf{x}_0$.

If, on the contrary, the time $t - t_0$ is much longer than the relaxation time of the random process, the system has enough time to 'forget' the initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$, and, irrespective of the value of $\mathbf{x}_0$, the transition probability tends to the probability density at $\mathbf{x}$:

$$\lim_{(t-t_0)\to\infty} P(\mathbf{x}, t|\mathbf{x}_0, t_0) = w(\mathbf{x}). \tag{1.6.7}$$

This property of P makes it more convenient to use instead of P another function:

$$p(\mathbf{x}, t|\mathbf{x}_0, t_0) = P(\mathbf{x}, t|\mathbf{x}_0, t_0) - w(\mathbf{x}), \tag{1.6.8}$$

which falls off to *zero* as $(t - t_0) \to \infty$. It follows from Eqs. (1.6.4) and (1.6.5) that this function satisfies the relations:

$$\int d\mathbf{x}\, p(\mathbf{x}, t|\mathbf{x}_0, t_0) = 0, \qquad \int d\mathbf{x}_0\, p(\mathbf{x}, t|\mathbf{x}_0, t_0)w(\mathbf{x}_0) = 0. \tag{1.6.9}$$

The Smoluchowski equation for p has the same form as the one for P (Eq. (1.6.3)).

Let us find the differential equations which govern the kinetics of the transition probabilities as functions of time (Kolmogorov equations). In the Smoluchowski equation for p we take the maximum time equal to $t + \Delta t$, the intermediate time $t' = t$:

$$p(\mathbf{x}, t + \Delta t|\mathbf{x}_0, t_0) = \int d\mathbf{x}'\, p(\mathbf{x}, t + \Delta t|\mathbf{x}', t)p(\mathbf{x}', t|\mathbf{x}_0, t_0). \tag{1.6.10}$$

Taking into account Eqs. (1.6.6) and (1.6.8) one may, at $\Delta t \to 0$, expand $p(\mathbf{x}, t + \Delta t|\mathbf{x}', t)$ and represent it in the form:

$$p(\mathbf{x}, t + \Delta t|\mathbf{x}', t) = \delta(\mathbf{x} - \mathbf{x}') - w(\mathbf{x}) - \lambda(\mathbf{x}, \mathbf{x}')\Delta t. \tag{1.6.11}$$

The meaning of $\lambda(\mathbf{x}, \mathbf{x}')$ will become obvious below.

Let us substitute the last equation into Eq. (1.6.10). According to the first Eq. (1.6.9), the term with $w(\mathbf{x})$ yields zero. Therefore

$$\frac{p(\mathbf{x}, t + \Delta t|\mathbf{x}_0, t_0) - p(\mathbf{x}, t|\mathbf{x}_0, t_0)}{\Delta t} = -\int d\mathbf{x}'\, \lambda(\mathbf{x}, \mathbf{x}')p(\mathbf{x}', t|\mathbf{x}_0, t_0). \tag{1.6.12}$$

In the limit $\Delta t \to 0$ Eq. (1.6.12) yields the Kolmogorov equation:

$$\frac{d}{dt}p(\mathbf{x}, t|\mathbf{x}_0, t_0) = -\int d\mathbf{x}'\, \lambda(\mathbf{x}, \mathbf{x}')p(\mathbf{x}', t|\mathbf{x}_0, t_0). \tag{1.6.13}$$

The quantities $-\lambda(\mathbf{x}, \mathbf{x}')$ are the derivatives of $p(\mathbf{x}, t|\mathbf{x}', t')$ with respect to the first time argument. At $\mathbf{x} \neq \mathbf{x}'$ they may be interpreted, with some

reservations, as probabilities of transition per unit time from the state $\mathbf{x}'$ to the state $\mathbf{x}$. By virtue of Eqs. (1.6.9) and (1.6.11),

$$\int d\mathbf{x}\lambda(\mathbf{x},\mathbf{x}') = 0, \qquad \int d\mathbf{x}'\lambda(\mathbf{x},\mathbf{x}')w(\mathbf{x}') = 0. \qquad (1.6.14)$$

Our next goal is to derive equations for the correlation functions of the fluctuations, $\psi_{\alpha\beta}(t_1 - t_2)$, given, in general, by Eq. (1.2.8). These equations should express these functions in terms of one-time correlation functions $\psi_{\alpha\beta}(0)$ (variances). However, at first one has to find, using the same methods, how a small perturbation, given at an initial instant, relaxes in time. Let, at the initial instant t_0, the quantities $\mathbf{x}(t_0) = \mathbf{x}_0$. The expected conditional mean value of a component x_α at the instant $t > t_0$ equals:

$$\langle x_\alpha(t)|\mathbf{x}_0,t_0\rangle = \int d\mathbf{x}\, x_\alpha P(\mathbf{x},t|\mathbf{x}_0,t_0). \qquad (1.6.15)$$

Since the mean value of this component $\langle x_\alpha\rangle = \int d\mathbf{x}\, x_\alpha w(\mathbf{x})$, the expected deviation of the random quantity x_α from its mean value equals:

$$\langle \Delta x_\alpha(t)|\mathbf{x}_0,t_0\rangle = \int d\mathbf{x}\, x_\alpha p(\mathbf{x},t|\mathbf{x}_0,t_0) = \int d\mathbf{x}\Delta x_\alpha p(\mathbf{x},t|\mathbf{x}_0,t_0). \qquad (1.6.16)$$

According to the Kolmogorov equation, the time derivative of this quantity equals:

$$\frac{d}{dt}\langle \Delta x_\alpha(t)|\mathbf{x}_0,t_0\rangle = -\int d\mathbf{x}'d\mathbf{x}\Delta x_\alpha\lambda(\mathbf{x},\mathbf{x}')p(\mathbf{x}',t|\mathbf{x}_0,t_0). \qquad (1.6.17)$$

The deviations $\Delta\mathbf{x}'$ from the mean values $\langle\mathbf{x}'\rangle$ may also be considered as small. Therefore the integral $\int d\mathbf{x}\,\Delta x_\alpha\lambda(\mathbf{x},\mathbf{x}')$, which depends on $\mathbf{x}'$, may be represented as a linear function of $\Delta\mathbf{x}'$:

$$\int d\mathbf{x}\Delta x_\alpha\, \lambda(\mathbf{x},\mathbf{x}') = \Lambda_{\alpha\beta}\Delta x'_\beta. \qquad (1.6.18)$$

Here $\Lambda_{\alpha\beta}$ is a matrix, the eigenvalues of which have the meaning of inverse relaxation times of the system, and summation over repeated indices is implied. From the last two equations one immediately obtains:

$$\frac{d}{dt}\langle \Delta x_\alpha(t)|\mathbf{x}_0,t_0\rangle = -\Lambda_{\alpha\beta}\langle \Delta x_\beta(t)|\mathbf{x}_0,t_0\rangle. \qquad (1.6.19)$$

The correlation function is easily expressed in terms of the conditional value of the fluctuation (Eq. (1.6.16)). Let us first represent $\psi_{\alpha\beta}(t_1 - t_2)$ as a sum of 'one-sided' parts:

$$\psi_{\alpha\beta}(t_1 - t_2) = \psi^+_{\alpha\beta}(t_1 - t_2) + \psi^-_{\alpha\beta}(t_1 - t_2), \qquad (1.6.20)$$

where

$$\psi_{\alpha\beta}^{+}(t_1 - t_2) = \Theta(t_1 - t_2)\langle\delta x_\alpha(t_1)\delta x_\beta(t_2)\rangle$$
$$\psi_{\alpha\beta}^{-}(t_1 - t_2) = \Theta(t_2 - t_1)\langle\delta x_\beta(t_2)\delta x_\alpha(t_1)\rangle. \tag{1.6.21}$$

Here $\Theta(t)$ is the known Heaviside step function:

$$\Theta(t) = \begin{cases} 1, & \text{if } t \geq 0; \\ 0, & \text{if } t < 0. \end{cases} \tag{1.6.22}$$

According to Eqs. (1.2.4), (1.6.8), and (1.6.16):

$$\begin{aligned}
\psi_{\alpha\beta}^{+}(t) &= \Theta(t) \int d\mathbf{x}\,d\mathbf{x}'\, x_\alpha p(\mathbf{x}, t|\mathbf{x}', 0) x_\beta' w(\mathbf{x}') \\
&= \Theta(t) \int d\mathbf{x}\,d\mathbf{x}'\, \delta x_\alpha p(\mathbf{x}, t|\mathbf{x}', 0) \delta x_\beta' w(\mathbf{x}') \\
&= \Theta(t) \int d\mathbf{x}'\, \langle\delta x_\alpha(t)|\mathbf{x}_0, 0\rangle \delta x_\beta' w(\mathbf{x}'),
\end{aligned} \tag{1.6.23}$$
$$\psi_{\alpha\beta}^{-}(t) = \psi_{\beta\alpha}^{+}(|t|).$$

The time derivative of $\psi_{\alpha\beta}^{+}(t)$, where $t = t_1 - t_2$, consists of two parts. The derivative of the step function is $\Theta'(t) = \delta(t)$. The corresponding coefficient is the integral in Eq. (1.6.23), taken at $t_1 = t_2$, i.e., $\psi_{\alpha\beta}(0)$. The time-derivative of $p(\mathbf{x}, t_1|\mathbf{x}', t_2)$ is found using the Kolmogorov equation (1.6.13). One of the integrals is given by Eq. (1.6.18). The result is a system of equations for $\psi_{\alpha\beta}^{+}$:

$$\frac{\partial}{\partial t}\psi_{\alpha\beta}^{+}(t) + \Lambda_{\alpha\gamma}\psi_{\gamma\beta}^{+}(t) = \psi_{\alpha\beta}(0)\delta(t). \tag{1.6.24}$$

Perhaps, the most convenient way to solve this system is to diagonalize the matrix $\Lambda_{\alpha\beta}$ (we assume that it is possible). Let $\lambda_m\,(m = 1,\ldots,M)$ and $\chi_\alpha^{(m)}$ be, respectively, the eigenvalues and eigenvectors of the $M \times M$ matrix $\Lambda_{\alpha\beta}$. Similarly, $\phi_\alpha^{(m)}$ is the eigenfunction of the conjugated matrix $\Lambda_{\alpha\beta}^{+} = \Lambda_{\beta\alpha}^{*}$ which corresponds to its eigenvalue λ_m^{*} of the last matrix:

$$\Lambda_{\alpha\beta}\chi_\beta^{(m)} = \lambda_m\chi_\alpha^{(m)}, \qquad \Lambda_{\alpha\beta}^{+}\phi_\beta^{(m)} = \lambda_m^{*}\phi_\alpha^{(m)}. \tag{1.6.25}$$

Let us assume also that all eigenvalues λ_m are different. The normalized eigenfunctions $\chi_\alpha^{(m)}$ and $\phi_\alpha^{(m)}$ satisfy the following orthogonality conditions (Gantmacher, 1977; Pease, 1965):

$$\sum_\alpha \chi_\alpha^{(m)}\phi_\alpha^{(n)*} = \delta_{nm}. \tag{1.6.26}$$

We can now introduce M linear combinations of fluctuations δx_α, the correlation functions of which decay in time exponentially, each with a

definite relaxation time λ_m^{-1}:

$$\Phi_{mn}^+(t) = \sum_{\alpha\beta} \phi_\alpha^{(m)*} \psi_{\alpha\beta}^+(t) \chi_\beta^{(n)}. \tag{1.6.27}$$

It follows from Eqs. (1.6.24)–(1.6.27) that $\Phi_{mn}(t)$ satisfy the equations:

$$\frac{d\Phi_{mn}^+(t)}{dt} + \lambda_m \Phi_{mn}^+(t) = \Phi_{mn}(0)\delta(t). \tag{1.6.28}$$

The solutions to these equations are:

$$\Phi_{mn}^+(t) = \Theta(t)\Phi_{mn}(0)e^{-\lambda_m t}. \tag{1.6.29}$$

Using Eqs. (1.6.26) and (1.6.27) one can find the original correlation functions:

$$\psi_{\alpha\beta}^+(t) = \sum_{mn} \chi_\alpha^{(m)} \Phi_{mn}^+(t) \phi_\beta^{(n)*} = \Theta(t) \sum_{m\gamma} e^{-\lambda_m t} \chi_\alpha^{(m)} \phi_\gamma^{(m)*} \psi_{\gamma\beta}(0). \tag{1.6.30}$$

The matrix of spectral densities can be found from Eqs. (1.3.8) and (1.6.21):

$$S_{\alpha\beta}(f) = 2\sum_{m\gamma} \left\{ \frac{\chi_\alpha^{(m)} \psi_{\gamma\beta}(0) \phi_\gamma^{(m)*}}{\lambda_m - i\omega} + \frac{\chi_\beta^{(m)} \psi_{\gamma\alpha}(0) \phi_\gamma^{(m)*}}{\lambda_m + i\omega} \right\}. \tag{1.6.31}$$

If there is only one quantity $\lambda = \tau^{-1}$ (τ is the relaxation time), the spectral density has a simple Lorentzian form:

$$S_x(f) = 4\langle(\delta x)^2\rangle\tau(1 + \omega^2\tau^2)^{-1}. \tag{1.6.32}$$

According to Eqs. (1.6.21) and (1.6.30) the correlation functions decay at $(t_1 - t_2) > 0$ and $(t_1 - t_2) < 0$ as linear superpositions of exponential functions $\exp(-\lambda_m|t_1 - t_2|)$. It means that in the Markov approximation the slopes of the correlation functions at $(t_1 - t_2) \to +0$ and at $(t_1 - t_2) \to -0$ are obviously different and are not zero, in contradiction to the general condition Eq. (1.2.17). According to this condition the correlation function is an even function of time with zero derivative at $t = 0$. This inconsistency of the Markov approximation has been very clearly explained by Lax (1960). This approximation is valid only at times much greater than the microscopic ones. The true time-dependence of the correlation function satisfies Eq. (1.2.17). The difference between the true behaviour of the correlation function and its approximate, Markovian, dependence is significant only in a very small time interval $|t| \leq \sim \tau_c \ll \tau$, where τ is the relaxation time of the fluctuations considered, and τ_c is a microscopic forgetting time of the nonmacroscopic variables which has different meaning in different systems. For instance, in a gas it is the duration of the particles' collisions that, at low density of the particles, is much smaller than the time between successive collisions of a
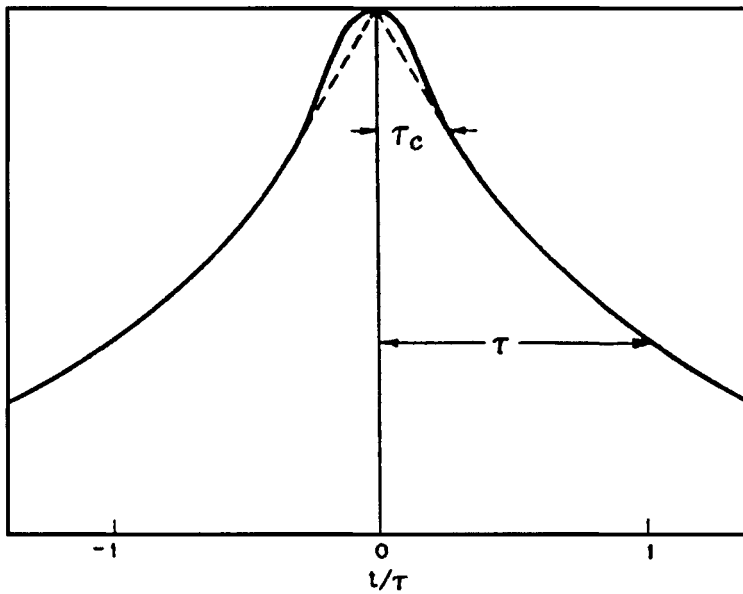
Fig. 1.4 Correlation function $\psi(t)$ of a random process with a single relaxation time τ. Solid curve: real decay of the correlation function. Dashed curve: purely exponential correlation function $\exp(-|t|/\tau)$ obtained in the Markov approximation. It deviates from the real one at small times $|t| < \tau_c$ which in Markov approximation are 'microscopic' times. From Lax (1960).

particle. The true correlation function and its Markov approximation are schematically shown in Fig. 1.4 (Lax, 1960).

1.7 Discrete Markov processes; random telegraph noise

A very important special case, especially in solid state physics, is random processes in which the random quantity takes only discrete values $x_i(i = 1, \ldots, N)$ and randomly switches between these values. It may be a defect, the transitions of which between two states modulate the resistance of the conductor with this defect. The noise generated by these transitions is illustrated in Fig. 1.5. It is called random telegraph noise (RTN), or random telegraph signals. The time between two successive transitions is usually many times greater than the time of relaxation (equilibration) of the crystal after each such transition. Then this process is a typical Markov one: the probability of transition depends only on the system's present state and does not depend on its history.

In the case of a discrete random quantity the unconditional probability w_i of finding the system in its i-th 'state' depends on the discrete index i. The

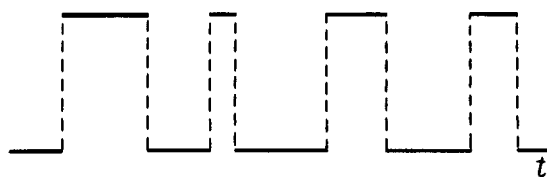Fig. 1.5 Random telegraph noise. The intervals between consequent pulses and their durations are random.

transition probability functions $P_{ii'}(t), p_{ii'}(t) = P_{ii'}(t) - w_i$, and $\lambda_{ii'}$ (analogue of $\lambda(\mathbf{x}, \mathbf{x}')$) are $N \times N$ matrices. They satisfy conditions analogous to Eqs. (1.6.4), (1.6.9), and (1.6.14):

$$\sum_i P_{ii'} = 1, \quad \sum_i p_{ii'} = 0, \quad \sum_{i'} p_{ii'} w_{i'} = 0, \tag{1.7.1}$$

$$\sum_i \lambda_{ii'} = 0, \quad \sum_{i'} \lambda_{ii'} w_{i'} = 0. \tag{1.7.2}$$

The nondiagonal elements of the matrix $-\lambda_{ij}$ have the direct meaning of transition probabilities per unit time from j to i and, for clarity, can be denoted as $W(j \to i)$. The diagonal elements can be expressed in terms of the sum of nondiagonal elements using the first Eq. (1.7.2):

$$\lambda_{ii} = -\sum_{j(\neq i)} \lambda_{ji}. \tag{1.7.3}$$

The Kolmogorov equations for a discrete Markov process take the form:

$$\frac{d}{dt} p_{ij}(t) = -\sum_k \lambda_{ik} p_{kj}(t). \tag{1.7.4}$$

The initial conditions are: $p_{ij}(0) = \delta_{ij} - w_i$.

The solutions to the Kolmogorov equations are exponential functions of the type $\exp(-Wt)$, where W are the eigenvalues of the matrix λ_{ij}, i.e., the roots of the equation of degree N:

$$\det \|\lambda_{ij} - W \delta_{ij}\| = 0. \tag{1.7.5}$$

In a stable system, the roots are either positive real or form complex pairs with a positive real part.

As in the general case, the correlation function of a discrete Markov random process can be represented in the form of Eq. (1.6.20). The equation for $\psi^+(t)$ follows from Eq. (1.6.23):

$$\psi^+(t) = \Theta(t) \sum_{ij} x_i p_{ij}(t) x_j w_j. \tag{1.7.6}$$

In order to find the explicit form of $\psi(t)$ at given transition rates, one has to solve the Kolmogorov equations (1.7.4) for $p_{ij}(t)$ and then find the probabilities w_i by solving the last Eq. (1.7.2) which, actually, is a system of equations in w_i. In the simplest case when $N = 2$ (the quantity x randomly switches between two states x_1 and x_2) the only nonzero eigenvalue is $W \equiv \tau^{-1} = W(1 \to 2) + W(2 \to 1)$, i.e., is equal to the sum of the rates of transitions back and forth. The probabilities of the two states are:

$$w_1 = \frac{W(2 \to 1)}{W(1 \to 2) + W(2 \to 1)}, \quad w_2 = \frac{W(1 \to 2)}{W(1 \to 2) + W(2 \to 1)}, \quad (1.7.7)$$

$w_1 + w_2 = 1$.

The matrix $p_{ij}(t)$ in the case of $N = 2$ equals:

$$\hat{p}(t) = \begin{pmatrix} w_2 & -w_1 \\ -w_2 & w_1 \end{pmatrix} e^{-Wt}. \quad (1.7.8)$$

The one-sided correlation functions

$$\psi^+(t) = \Theta(t)w_1 w_2 (x_1 - x_2)^2 e^{-Wt}, \quad \psi^-(t) = \psi^+(|t|). \quad (1.7.9)$$

Thus, according to Eq. (1.3.8), the spectral density equals:

$$S(f) = 4 w_1 w_2 (x_1 - x_2)^2 \tau / (1 + \omega^2 \tau^2). \quad (1.7.10)$$

The spectrum of the simplest telegraph noise is Lorentzian (Machlup, 1954).

1.8 Quasi-continuous (diffusion-like) Markov processes

In some Markov processes, the random quantities $\mathbf{x}(t)$ vary in time quasi-continuously, that is, by relatively very small steps which, moreover, are randomly positive and negative. Any considerable change of the random quantity is in this case a result of a very great number of such steps. Such processes are called continuous or, quite naturally, diffusion-like Markov processes.

Let us find the equation for the conditional probability distribution $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$ for this kind of random process. The starting point is the Smoluchowski equation (1.6.3) in which the maximum time is taken to be $t + \Delta t$, and the intermediate time equals t (as in Eq. (1.6.10)). Let us multiply both sides of the equation by an arbitrary function $h(\mathbf{x})$, which is assumed to be zero at the boundaries of the interval of $\mathbf{x}$ variation, and integrate over $\mathbf{x}$. Upon interchanging $\mathbf{x}$ and $\mathbf{x}'$ in the integrand one obtains (Wang & Uhlenbeck, 1945; Rytov, 1976; Rytov *et al.*, 1987, 1988):

$$\int d\mathbf{x}\, h(\mathbf{x}) P(\mathbf{x}, t + \Delta t | \mathbf{x}_0, t_0) = \int d\mathbf{x}\, d\mathbf{x}'\, h(\mathbf{x}') P(\mathbf{x}', t + \Delta t | \mathbf{x}, t) P(\mathbf{x}, t | \mathbf{x}_0, t_0). \quad (1.8.1)$$

At small values of Δt the quantity $P(\mathbf{x}', t + \Delta t | \mathbf{x}, t)$ is significant only for small differences $|\mathbf{x}' - \mathbf{x}|$ (quasi-continuity). Since $\mathbf{x}$ is quasi-continuous and $\Delta t \to 0$, only terms up to the second order in $\mathbf{x}' - \mathbf{x}$ should be retained. Taking into account the normalization condition for P, Eq. (1.6.4), one obtains:

$$\int d\mathbf{x} h(\mathbf{x})[P(\mathbf{x}, t + \Delta t | \mathbf{x}_0, t_0) - P(\mathbf{x}, t | \mathbf{x}_0, t_0)]$$

$$= \int d\mathbf{x} h'_\alpha(\mathbf{x}) P(\mathbf{x}, t | \mathbf{x}_0, t_0) \int d\mathbf{x}' (x'_\alpha - x_\alpha) P(\mathbf{x}', t + \Delta t | \mathbf{x}, t) \qquad (1.8.2)$$

$$+ \frac{1}{2} \int d\mathbf{x} h''_{\alpha\beta}(\mathbf{x}) P(\mathbf{x}, t | \mathbf{x}_0, t_0) \int d\mathbf{x}' (x'_\alpha - x_\alpha)(x'_\beta - x_\beta) P(\mathbf{x}', t + \Delta t | \mathbf{x}, t).$$

Here $h' \equiv \partial h / \partial x_\alpha$, $\quad h''_{\alpha\beta} \equiv \partial^2 h / \partial x_\alpha \partial x_\beta$.

In this equation, $P(\mathbf{x}, t + \Delta t | \mathbf{x}, t)$ can be expanded in powers of Δt. The zero-th order term, i.e., $\delta(\mathbf{x}' - \mathbf{x})$ (Eq. (1.6.6)) yields a zero contribution to the integrals over $\mathbf{x}'$. In the r.h.s. only terms linear in Δt remain. By dividing both sides by Δt one obtains in the l.h.s. a time derivative. Integration over $\mathbf{x}$ by parts, taking into account that $h(\mathbf{x}) = 0$ at the lower and upper limits, yields an equation in integrals over $\mathbf{x}$ with the function $h(\mathbf{x})$ as a factor in the integrand. Since the equation is valid at arbitrary $h(\mathbf{x})$, one obtains the equation which is called by physicists the Fokker–Planck (Fokker, 1914; Planck, 1917), or Einstein–Fokker–Planck, equation, and in mathematical text-books it is called the second (or forward) Kolmogorov equation:

$$\frac{\partial P(\mathbf{x}, t | \mathbf{x}_0, t_0)}{\partial t} + \frac{\partial}{\partial x_\alpha} \left[A_\alpha(\mathbf{x}, t) P(\mathbf{x}, t | \mathbf{x}_0, t_0) \right]$$

$$- \frac{\partial^2}{\partial x_\alpha \partial x_\beta} \left[D_{\alpha\beta}(\mathbf{x}, t) P(\mathbf{x}, t | \mathbf{x}_0, t_0) \right] = 0. \qquad (1.8.3)$$

Here summation over repeated indices is implied and the following notations have been introduced:

$$A_\alpha(\mathbf{x}, t) = \int d\mathbf{x}' (x'_\alpha - x_\alpha) \dot{P}(\mathbf{x}', t | \mathbf{x}, t),$$

$$D_{\alpha\beta}(\mathbf{x}, t) = \frac{1}{2} \int d\mathbf{x}' (x'_\alpha - x_\alpha)(x'_\beta - x_\beta) \dot{P}(\mathbf{x}', t | \mathbf{x}, t), \qquad (1.8.4)$$

In the last equations

$$\dot{P}(\mathbf{x}', t | \mathbf{x}, t) = \left[\frac{\partial P(\mathbf{x}', t' | \mathbf{x}, t)}{\partial t'} \right]_{t'=t}.$$

The solution to Eq. (1.8.3) must be nonnegative and must satisfy the normalization condition (1.6.4) and the initial condition (1.6.6).

The Fokker–Planck equation has the form of a continuity equation in which the role of the flow in x-space is played by the quantity:

$$J_\alpha(\mathbf{x}, t) = A_\alpha(\mathbf{x}, t)P(\mathbf{x}, t|\mathbf{x}_0, t_0) - \frac{\partial}{\partial x_\beta}\left[D_{\alpha\beta}(\mathbf{x}, t)P(\mathbf{x}, t|\mathbf{x}_0, t_0)\right]. \qquad (1.8.5)$$

The first part of J_α has the meaning of a drift flow, the second one, proportional to the gradient, is a diffusion flow. In fact, the quantity $\mathbf{A}(\mathbf{x}, t)$ is nonzero only if there is an asymmetry of transitions, those with $x' - x > 0$ and with $x' - x < 0$. The quantity $D_{\alpha\beta}$ has an obvious meaning of the tensor of diffusion coefficients. Depending on the meaning of the variables $\mathbf{x}$, it may be not only the diffusion in the configuration (common) space but also diffusion in the space of velocities, diffusion of energy, diffusion of the phase of oscillations and so on. The Fokker–Planck equation is widely used in the theory of random processes.

If at the initial time t_0, not a definite state $\mathbf{x}_0$ of the system is given but a probability density function $w(\mathbf{x}_0 t_0)$, it is more appropriate to use an equation for the probability density function $w(\mathbf{x}t)$. One has to take into account Eq. (1.2.4) for the two-dimensional density function,

$$w_2(\mathbf{x}t; \mathbf{x}_0 t_0) = P(\mathbf{x}t|\mathbf{x}_0 t_0)w(\mathbf{x}_0 t_0),$$

and Eq. (1.1.4), which expresses the function $w(\mathbf{x}t)$ in terms of $w_2(\mathbf{x}t; \mathbf{x}_0 t_0)$. Then

$$w(\mathbf{x}t) = \int d\mathbf{x}_0 P(\mathbf{x}t|\mathbf{x}_0 t_0)w(\mathbf{x}_0 t_0).$$

Multiplying the Fokker–Planck Eq. (1.8.3) by $w(\mathbf{x}_0 t_0)$ and integrating over $\mathbf{x}_0$ one obtains the equation for $w(\mathbf{x}t)$:

$$\frac{\partial w(\mathbf{x}t)}{\partial t} + \frac{\partial}{\partial x_\alpha}\left[A_\alpha(\mathbf{x}t)w(\mathbf{x}t)\right] - \frac{\partial^2}{\partial x_\alpha \partial x_\beta}\left[D_{\alpha\beta}(\mathbf{x}t)w(\mathbf{x}t)\right] = 0. \qquad (1.8.6)$$

Summation over repeated indices is implied.

To have some idea of how the Fokker-Planck equation 'works', it is instructive to discuss its application to a simple problem. For randomly moving Brownian particles (Sec. 1.9) the vectors $\mathbf{J_r}(\mathbf{r}\mathbf{v}t)$ and $\mathbf{J_v}(\mathbf{r}\mathbf{v}t)$ are flow densities in common space ($\mathbf{r}$ is the radius-vector) and in the space of the particles' velocities $\mathbf{v}$, respectively, at definite $\mathbf{r}$, $\mathbf{v}$, and time t. According to the definition, Eq. (1.8.4), $\mathbf{A_r}$ is the speed of the radius-vector's change,

$$\mathbf{A_r} = \lim_{\Delta t \to 0} \int d\mathbf{r}' d\mathbf{v}' \frac{(\mathbf{r}' - \mathbf{r})}{\Delta t}\left[P(\mathbf{r}'\mathbf{v}', t + \Delta t|\mathbf{r}\mathbf{v}t) - P(\mathbf{r}'\mathbf{v}', t|\mathbf{r}\mathbf{v}t)\right] = \mathbf{v}, \qquad (1.8.7)$$

that is, the velocity of a particle. $D_{r_\alpha r_\beta}$ is the speed of growth of the product

of the particle's displacements along the directions α and β, i.e., the tensor of diffusion coefficients.

According to the same definition, Eq. (1.8.4),

$$\mathbf{A_v(v)} = \lim_{\Delta t \to 0} \int d\mathbf{r}' d\mathbf{v}' \frac{\mathbf{v}' - \mathbf{v}}{\Delta t} [P(\mathbf{r}'\mathbf{v}', t + \Delta t | \mathbf{r}\mathbf{v}t) - P(\mathbf{r}'\mathbf{v}', t | \mathbf{r}\mathbf{v}t)] \qquad (1.8.8)$$

is the acceleration of the particle. Similarly, $D_{v_\alpha v_\beta}$ determines the spreading of particles' velocities, i.e., diffusion in the space of velocities.

The acceleration $\mathbf{A_v(v)}$ consists of two parts. If a force $-\nabla U$, where U is the potential energy, is acting on each particle it yields an acceleration $(-\nabla U)/M$, where M is the mass of a particle. The second part is due to the hindering of the particle's motion by collisions with the molecules of the medium. If the particle's velocity v_α is at instant t equal to zero, it can change (due to collisions) by the instant $t + \Delta t$ to $v'_\alpha > 0$ and $v'_\alpha < 0$ with equal probability. Integration over $\mathbf{v}'$ in Eq. (1.8.8) yields zero. Assume that the component $v_\alpha > 0$. Then the relative velocity of the particle and the molecules moving in the opposite direction is higher, and that of the particle and molecules moving in the same direction is lower than in the case $v_\alpha = 0$. The impacts of the former are on average stronger, and the probability of a negative change of velocity $v'_\alpha - v_\alpha < 0$ is higher than that of a positive change. It means that at small velocities $|\mathbf{v}|$ this second part of $\mathbf{A_v}$ may be represented by $-\mathbf{v}/\tau$. The time τ is the relaxation time of the directional velocity of a particle. Therefore

$$\mathbf{A_v} = \frac{1}{M}(-\nabla U) - \frac{\mathbf{v}}{\tau}. \qquad (1.8.9)$$

The quantities

$$n(\mathbf{r}, t) = \int d\mathbf{v}\, w(\mathbf{r}\mathbf{v}t), \quad \mathbf{j}(\mathbf{r}, t) = \int d\mathbf{v}\, \mathbf{J_r}\, w(\mathbf{r}\mathbf{v}t) \qquad (1.8.10)$$

are the total density and flow density of particles, respectively, (with any velocity) at the point $\mathbf{r}$ and at instant t. After integration of Eq. (1.8.6) over $\mathbf{v}$, the term with divergence of $\mathbf{J_v}$ drops out. One obtains the continuity equation:

$$\frac{\partial n}{\partial t} + \operatorname{div}\mathbf{j} = 0. \qquad (1.8.11)$$

The flow density is a sum of drift and diffusion flows. In an isotropic medium the diffusion coefficient $D_\mathbf{r}$ is isotropic:

$$\mathbf{j} = \langle \mathbf{v} \rangle n(\mathbf{r}t) - D\nabla n(\mathbf{r}t), \qquad \langle \mathbf{v} \rangle n(\mathbf{r}t) = \int d\mathbf{v}\, \mathbf{v}\, w(\mathbf{r}\mathbf{v}t). \qquad (1.8.12)$$

These relations are used, in particular, in the theory of Brownian motion.

1.9 Brownian motion

In 1827 the English botanist Robert Brown discovered and thoroughly studied a phenomenon called after him 'Brownian motion'. Fine particles, in particular, spores of mushrooms or pollen suspended in a liquid, are moving quickly, randomly, and incessantly (Brown, 1828). As the dimensions of the particles are of the order of 1 μm, they are visible only by microscope. R. Brown was not the first to observe this motion (see the early history in the book by Brush, 1976).

'Brown's discovery, therefore, was not his observation of the motion of microscopic particles in fluids; that observation had been made many times before; instead, it was his emancipation from the previously current notion that such movements had a specifically organic character. What Brown showed was that almost any kind of matter, organic or inorganic, can be broken into fine particles that exhibit the same kind of dancing motion; thus he removed the subject from the realm of biology into the realm of physics' (Brush, 1976).

His experiments also proved that this motion is not a result of effects which may be related to the imperfectness of the experiment, like convection under inhomogeneous heating of the liquid, but a fundamental physical phenomenon.

Brownian motion was properly explained only in the first decade of the twentieth century when Einstein (1905, 1906, 1907) and Smoluchowski (1906a, 1906b, 1915) developed the theory of this phenomenon. Most likely, the reader knows that the Brownian particles move because they participate in the thermal motion of the molecules of the liquid. Brownian motion may be also viewed as being caused by uncompensated impacts of the liquid's molecules.

According to statistical mechanics, the mean random velocity squared of a particle with mass M in a medium with absolute temperature T is equal to $\langle v^2 \rangle = 3k_B T/M$, where k_B is the Boltzmann constant. For a particle with radius $a \sim 1$ μm and with a mass density of the order of the water density at $T = 300$ K, the mean velocity is ~ 0.1 cm/s, that is, of the order of thousands of radii of the particle per second. The velocities seen under microscope are smaller by many times. The theory of Brownian motion developed by A. Einstein and M. von Smoluchowski had not only qualitatively explained the phenomenon and eliminated the paradoxes, in particular, the one just mentioned, but also established its quantitative laws. The experimental verification of these laws became one of the most important arguments in favor of the molecular structure of matter and kinetic nature of thermal phenomena.

Over the decades after the publication of the theory by A. Einstein and M. von Smoluchowski, many fluctuation phenomena similar to Brownian motion have been discovered. Therefore at present this term is understood in a broad sense. Besides the motion of fine particles suspended in a liquid, the rotational Brownian motion of molecules, random (Brownian) motion of charge carriers in conductors that produces electrical noise, Brownian motion of domain walls in ferromagnets and ferroelectrics, random changes of the phase of an oscillator, and many other similar phenomena are known.

The presentation of the fundamental properties of Brownian motion given below is based on the theory of diffusion-like Markov processes (Sec. 1.8). The Langevin approach to Brownian motion is presented in the next section.

The Brownian particle is moving randomly under the impacts of the surrounding molecules. As the mass m of the molecules is many times smaller than the mass M of the Brownian particle ($m \ll M$), the change $|\Delta\mathbf{P}|$ of the momentum of the latter is relatively very small. In fact, $|\Delta\mathbf{P}| \leq 2p$, where $\mathbf{p}$ is the momentum of the incident molecule. The equality $|\Delta\mathbf{P}| = 2p$ corresponds to the motion of the molecule after the elastic collision exactly in the backward direction. As $p \sim \sqrt{2mk_B T}$ and $P \sim \sqrt{2Mk_B T}$, the relative changes of the momentum and velocity v of the Brownian particle are

$$|\Delta\mathbf{P}|/P = |\Delta\mathbf{v}|/v \leq 2\sqrt{m/M} \ll 1. \qquad (1.9.1)$$

The duration of one impact may be estimated, assuming that the forces acting between the molecule and the surface of the Brownian particle in the process of collision are atomic forces. Therefore, the duration of the interaction is expected to be of the order of a period of short-wavelength vibrations of atoms in condensed matter, i.e., $\sim 10^{-14} - 10^{-12}$ s. According to the theory of random pulsed processes (Sec. 1.5) the correlation time τ_c of such a process is of the order of the duration of one pulse, that is,

$$\tau_c \sim 10^{-14} - 10^{-12} \text{ s.} \qquad (1.9.2)$$

By this time, the Brownian particle moving with thermal velocity $v_T = \sqrt{3k_B T/M}$ traverses a very small path $\sim v_T \tau_c$.

These estimates show that the changes of both the position $\mathbf{r}$ and velocity $\mathbf{v}$ of the Brownian particle occur by very small steps, or in other words, considerable changes of $\mathbf{r}$ and $\mathbf{v}$ occur as a result of a huge number of collisions with the molecules of the liquid. It allows us to consider Brownian motion as a quasi-continuous (diffusion-like) Markov random process which is governed by the Fokker–Planck equation (1.8.3) (or Eq. (1.8.6)).

Let us derive the fundamental relations of the theory of Brownian motion. Suppose that a force $-\nabla U$ is acting on each Brownian particle, where U is

the potential energy of a particle. It may be either the force of gravity $M\mathbf{g}$, where $\mathbf{g}$ is the free fall acceleration, or the force $q\mathbf{F}$ of an electric field $\mathbf{F}$ if each particle has a charge q. Consider the case of a spatially uniform system of particles with probability density $w(\mathbf{v}, t)$ independent of $\mathbf{r}$. Multiplying Eq. (1.8.6) by $\mathbf{v}$ and integrating over $\mathbf{v}$, one obtains the equation for the mean (drift) velocity $\langle \mathbf{v} \rangle$ of the particles. With account of Eq. (1.8.9) it takes the form:

$$\frac{\partial}{\partial t}\langle \mathbf{v} \rangle = \frac{1}{M}(-\nabla U) - \frac{1}{\tau}\langle \mathbf{v} \rangle. \tag{1.9.3}$$

Under stationary conditions

$$\langle \mathbf{v} \rangle = \frac{\tau}{M}(-\nabla U). \tag{1.9.4}$$

In the case of an electric field $-\nabla U = q\mathbf{F}$, the coefficient $\mu = q\tau/M$ is the mobility of the particles. The quantity $k_f = M/\tau$ may be considered as the friction coefficient of the Brownian particle. In a simple model of a spherical particle of radius a, according to the Stokes equation, $k_f = 6\pi\eta a$, where η is the dynamic viscosity of the liquid.

Let us examine two problems. The first is the distribution of the particles in space. According to Eqs. (1.8.12) and (1.9.4) the density of the particles' flow is at each point given by

$$\mathbf{J} = \frac{\tau}{M}n(\mathbf{r}t)(-\nabla U) - D\nabla n(\mathbf{r}, t). \tag{1.9.5}$$

In equilibrium, that is, in the absence of a net flow of particles (in the field of gravity the fall of particles is prevented by a reflecting wall or Earth's surface)

$$\frac{1}{n}\nabla n = \frac{\tau}{MD}(-\nabla U) = \frac{\mu}{qD}(-\nabla U). \tag{1.9.6}$$

As is known from statistical mechanics, the equilibrium distribution of the concentration in a field is the Boltzmann distribution:

$$n(\mathbf{r}) = n(0)\exp\left[-\frac{U(\mathbf{r}) - U(0)}{k_B T}\right] \tag{1.9.7}$$

Comparing Eqs. (1.9.6) and (1.9.7), A. Einstein found the fundamental relation between the diffusion coefficient and the mobility:

$$D = \frac{k_B T}{q}\mu, \qquad \text{or} \qquad D = \frac{k_B T\tau}{M}. \tag{1.9.8}$$

In the absence of any external forces, the density of the Brownian particles obeys the diffusion equation :

$$\frac{\partial n}{\partial t} - D\Delta n = 0. \tag{1.9.9}$$

Suppose the particles at the initial instant $t_0 = 0$ are confined to the point $\mathbf{r}_0 = 0$, that is, their density is $n(\mathbf{r}, 0) = \delta(\mathbf{r})$. The solution to the diffusion equation which satisfies this condition reads:

$$n(\mathbf{r}, t) = [4\pi Dt]^{-d/2} \exp\left[-\frac{r^2}{4Dt}\right]. \tag{1.9.10}$$

Here d is the Euclidean dimension of the space in which the spreading of the initially confined particles is observed. The mean displacement squared of the particles from their initial position in any direction α is equal to

$$\langle x_\alpha^2 \rangle = \int d\mathbf{r} x_\alpha^2 n(\mathbf{r}, t) = 2Dt. \tag{1.9.11}$$

It is the famous Einstein's equation for the variance of the displacement of diffusing particles.

The second problem is the relaxation of the Brownian particle's velocity and the correlation function of the particle's velocities. According to Eq. (1.9.3), in the absence of external forces the mean velocity at instant t, under the condition that at time t_0 it equals $\mathbf{v}_0$, is given by the following equation:

$$\langle \mathbf{v}(t) \rangle_{\mathbf{v}_0} = \mathbf{v}_0 \exp(-t/\tau). \tag{1.9.12}$$

It means that τ is, in fact, the relaxation time of the directional velocity of a particle.

Let us find the correlation function $\psi_{v\alpha\beta}(t_1 - t_2)$ of the components $v_\alpha(t_1)$ and $v_\beta(t_2)$ of a particle's velocity. According to the general theory of Markov processes (Sec. 1.6), one has to find first the one-sided correlation function:

$$\begin{aligned}
\psi_{v\alpha\beta}^+(t_1 - t_2) &= \Theta(t_1 - t_2)\psi_{v\alpha\beta}(t_1 - t_2) \\
&= \Theta(t_1 - t_2)\int d\mathbf{r}_1 d\mathbf{v}_1 d\mathbf{v}_2 v_{1\alpha} P(\mathbf{r}_1 \mathbf{v}_1 t_1 | \mathbf{r}_2 \mathbf{v}_2 t_2) v_{2\beta} w(\mathbf{v}_2).
\end{aligned} \tag{1.9.13}$$

Here $w(\mathbf{v})$ is the probability density function of the velocities.

Multiply the Fokker–Planck equation (1.8.3) by $v_{1\alpha} v_{2\beta} w(\mathbf{v}_2)$ and integrate all terms over $\mathbf{r}$ and velocities. Using Eq. (1.8.9) for $\mathbf{A}_v$ one obtains a simple equation:

$$\frac{\partial}{\partial t}\psi_{v\alpha\beta}^+(t) + \frac{1}{\tau}\psi_{v\alpha\beta}^+(t) = \delta_{\alpha\beta}\langle v_\alpha^2 \rangle \delta(t)$$

Using also the equation for $\psi_{\mathrm{v}\alpha\beta}^{-}(t)$ one obtains:

$$\psi_{\mathrm{v}\alpha\beta}(t_1 - t_2) = \delta_{\alpha\beta}\langle v_\alpha^2 \rangle \exp[-|t_1 - t_2|/\tau]. \qquad (1.9.14)$$

According to the Wiener–Khintchine relation (1.3.8), the corresponding spectral density is

$$S_{\mathrm{v}\alpha\beta}(f) = \delta_{\alpha\beta}4\langle v_\alpha^2 \rangle \tau(1 + \omega^2\tau^2)^{-1}. \qquad (1.9.15)$$

Using Eq. (1.9.8) for the diffusion coefficient of Brownian particles, one can relate the low-frequency ($\omega \ll \tau^{-1}$) spectral density and the diffusion coefficient:

$$S_{\mathrm{v}\alpha\beta}(0) = 4D_{\alpha\beta} = 4D\delta_{\alpha\beta}. \qquad (1.9.16)$$

This relation, which may be called the fluctuation–diffusion relation, means that the low-frequency spectral density of the particles' random velocities differs from their diffusion coefficient only by a constant factor 4. It is important that the temperature T of the system does not appear in this relation. A similar fluctuation–diffusion relation is valid even in some nonequilibrium systems in which the distribution of particles in energy strongly differs from an equilibrium one. An example is a gas of hot electrons in semiconductors when the interelectron interaction is negligible (Sec. 3.5).

Let Brownian particles with radii $a = 1$ μm move in water at $T = 300$ K and their mass density ρ be close to the density of water. The dynamic viscosity of water equals $\eta \approx 10^{-3}$ Pa·s. The velocity relaxation time is then

$$\tau = 2\rho a^2/9\eta \sim 2 \cdot 10^{-7}\mathrm{s}. \qquad (1.9.17)$$

The trajectory of a Brownian particle under visual observation is, in fact, averaged over time intervals of the order of the time of the eye's persistence of vision t_v. In other words, of all frequencies f of particle's velocity fluctuations only a low-frequency band with $f \le t_v^{-1}$ is visually recorded. The mean square velocity measured by visual observations is of the order of $S_v(0)\Delta f$. As $t_v \gg \tau$ the measured path-length is smaller than the real one and the measured velocity is smaller than the thermal velocity $\sqrt{3k_B T/M}$ by a factor τ/t_v. The difference between the real path and the visible, smoothed, one is schematically shown in Fig. 1.6.

1.10 The Langevin approach to the kinetics of fluctuations

In this section a very general approach to the calculation of correlation properties of random processes is presented. Its idea was suggested by
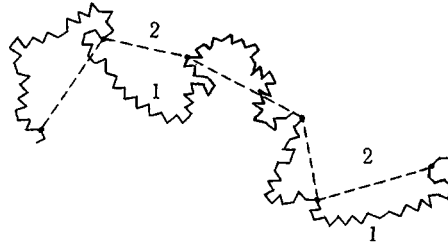
Fig. 1.6 Schematic diagram of a trajectory of a Brownian particle. Dashed curve: visible trajectory smoothed over time intervals equal to the time of the eye's persistence of vision, t_v. Solid curve: trajectory closer to the real one.

Langevin (1908) in his theory of Brownian motion. The principal difference between the Langevin approach and the Markov one (Secs. 1.6–1.8) is the following. The correlation functions $\psi_{\alpha\beta}(t_1, t_2)$ and the transition probabilities $P(\mathbf{x}, t | \mathbf{x}_0, t_0)$ are definite, not random, functions, and the equations for the one-sided correlation function $\psi_{\alpha\beta}^+$ and for P in the Markov approach are equations for definite, not random, quantities which are just the sought-for characteristics of the random process. In contrast with this approach, in the Langevin method equations are derived and solved for the actual random variables $\mathbf{x}(t)$ or their fluctuations $\delta\mathbf{x}(t) = \mathbf{x}(t) - \langle\mathbf{x}\rangle$, and the correlation function is then calculated by averaging the product of fluctuations $\delta\mathbf{x}(t)$ taken at different instants of time (Eqs. (1.2.1) and (1.2.8)):

$$\psi_{\alpha\beta}(t_1, t_2) = \langle \delta x_\alpha(t_1) \delta x_\beta(t_2) \rangle.$$

Let us examine the correlation properties of the Brownian particle's velocity $\mathbf{v}(t)$ using the Langevin approach. As was shown in Sec. 1.9, the relaxation time of the velocity, $\tau = M/k_f \gg \tau_c$, where τ_c is the correlation time of random forces acting on the Brownian particle and is of the order of the duration of an impact of a molecule onto the particle. Using this inequality, $\mathbf{v}(t)$ can be smoothed over time intervals Δt, which satisfy two inequalities: $\tau_c \ll \Delta t \ll \tau$, i.e., they must be much longer than τ_c but much shorter than τ. Let us derive the equation of motion for this smoothed $\mathbf{v}(t)$. In the case of Brownian motion this equation is Newton's equation.

The force acting on the Brownian particle is set up by the impacts of surrounding molecules. This force falls, however, into two parts, and this is the main Langevin's idea. The first part is explicitly expressed in terms of the random velocity $\mathbf{v}$ itself, which, as was mentioned above, is smoothed (averaged) over time intervals Δt. The latter must be long enough so that during this time the particle experiences a huge number of impacts. Hence, this part of the force can be treated macroscopically. The time intervals

must be also so short that the particle's velocity remains almost unchanged during Δt. At small $|\mathbf{v}|$ this 'friction' force equals $\mathbf{F}_f = -k_f \mathbf{v} = -(M/\tau)\mathbf{v}$ (Eq. (1.9.3)). It results from the fact that the impacts of molecules moving in the opposite direction are on average stronger than the impacts of the ones moving in the direction of the particle's motion.

However, at each instant t there is also another part of the force created by the random impacts of the liquid's molecules. This part of the force is nonzero even at those instants at which $\mathbf{v} = 0$. Its variation in time is very frequent, and its correlation time is smaller by many orders of magnitude than the relaxation time τ of the velocity $\mathbf{v}(t)$. This force divided by mass M is denoted here by $\mathbf{y}(t)$. Its mean value $\langle \mathbf{y} \rangle = 0$. According to the theory of random pulses (Sec. 1.5), the correlation function of $\mathbf{y}(t)$ is a narrow peak with a width of the order of τ_c. It means that, when only processes smoothed over time intervals $\Delta t \gg \tau_c$ are considered, the correlation function of $\mathbf{y}(t)$ may be taken as proportional to the delta-function of the time difference:

$$\psi_{\mathbf{y}}(t_1 - t_2) = \frac{1}{2} S_{\mathbf{y}}(0)\delta(t_1 - t_2). \tag{1.10.1}$$

Here the coefficient at $\delta(t_1 - t_2)$ is expressed, according to the Wiener–Khintchine relationship (1.3.8), in terms of the spectral density at zero frequency:

$$S_{\mathbf{y}}(f) = S_{\mathbf{y}}(0) = \text{const.}, \qquad f \ll \tau_c^{-1}. \tag{1.10.2}$$

Thus, the equation of motion of the Brownian particle takes the form suggested by Langevin (1908):

$$\frac{d\mathbf{v}}{dt} + \frac{\mathbf{v}}{\tau} = \mathbf{y}(t). \tag{1.10.3}$$

The r.h.s. can not be expressed in terms of $\mathbf{v}(t)$. It plays the role of the source of the Brownian particle's random motion. Such a source is called a Langevin source, and Eq. (1.10.2) is called the Langevin equation.

Turning to Fourier transforms $\mathbf{v}(\omega)$ and $\mathbf{y}(\omega)$, one obtains from Eq. (1.10.3):

$$\mathbf{v}(\omega) = (-i\omega + \tau^{-1})^{-1}\mathbf{y}(\omega) \tag{1.10.4}$$

As was shown above (Sec. 1.3), the spectral density of a random quantity $x(t)$ is proportional to $\langle x(\omega)x^*(\omega) \rangle$. Therefore,

$$S_{v_\alpha}(f) = (-i\omega + \tau^{-1})^{-1}(i\omega + \tau^{-1})^{-1}S_{y_\alpha}(f) = (\omega^2 + \tau^{-2})^{-1}S_{y_\alpha}(0). \tag{1.10.5}$$

Thus, in the Langevin approach the spectral density of the fluctuations is expressed in terms of the spectral density of the fluctuations' source, which

is independent of frequency, i.e., is a constant, at all frequencies $f \ll \tau_c^{-1}$. This constant has to be found.

The last problem is especially simple in equilibrium systems. The variance of the fluctuations (Sec. 1.2 and Eq. (1.3.16)), i.e.,

$$\psi_x(0) = \langle (\delta x)^2 \rangle = \int_0^\infty df \, S_x(f), \tag{1.10.6}$$

is a thermodynamic, not a kinetic, quantity (it does not depend on the kinetics of the system, in particular, on its relaxation times), and is known from statistical mechanics or thermodynamics. In the case of Brownian particles

$$\langle v_\alpha^2 \rangle = k_B T / M. \tag{1.10.7}$$

On the other hand,

$$\langle v_\alpha^2 \rangle = \int_0^\infty df \, S_{v_\alpha}(f) = S_{y_\alpha} \int_0^\infty df (\omega^2 + \tau^{-2})^{-1} = \frac{\tau}{4} S_{y_\alpha}. \tag{1.10.8}$$

Comparing Eqs. (1.10.8) and (1.10.7) one obtains:

$$S_{y_\alpha} = \frac{4 \, k_B T}{\tau \, M}. \tag{1.10.9}$$

Upon substitution of this expression into Eq. (1.10.5) one obtains an equation for the spectral density of random velocity which coincides with Eq. (1.9.15) derived by a different (Markov, Fokker–Plank) method.

Equation (1.7.10) for the spectral density $S_x(f)$ of a random variable $x(t)$, which is randomly switching between two values x_1 and x_2, has been derived using the Markov method. It is instructive to derive the same equation by using the Langevin approach.

The fluctuations $\delta x(t)$ can be viewed as originating from the fluctuations of the occupancy numbers, $n_1(t)$ and $n_2(t) = 1 - n_1(t)$, of the two states of the system around their mean values $\langle n_1 \rangle = w_1$ and $\langle n_2 \rangle = w_2 = 1 - w_1$. Then

$$\delta x(t) = x_1 \delta n_1(t) + x_2 \delta n_2(t) = (x_1 - x_2)\delta n_1(t). \tag{1.10.10}$$

The Langevin kinetic equation for $\delta n_1(t)$ reads:

$$\frac{\partial \delta n_1}{\partial t} + W(1 \to 2)\delta n_1(t) - W(2 \to 1)\delta n_2(t) = \delta J_{2 \to 1}(t) - \delta J_{1 \to 2}(t). \tag{1.10.11}$$

Here $W(1 \to 2)$ and $W(2 \to 1)$ are, as in Sec. 1.7, the probabilities, per unit time, of transitions between the two states, $\delta J_{2 \to 1}(t)$ and $\delta J_{1 \to 2}(t)$ are

the Langevin sources (flows) corresponding to these transitions. Using the equation $\delta n_2 = -\delta n_1$, one obtains the final form of the Langevin equation:

$$\frac{\partial \delta n_1}{\partial t} + \frac{1}{\tau}\delta n_1(t) = \delta J_{2\to1}(t) - \delta J_{1\to2}(t), \qquad \tau^{-1} = W(1 \to 2) + W(2 \to 1).$$
$$(1.10.12)$$

The random process is assumed to be a Markov one: any transition of the system depends only on the its present state, all previous transitions are 'forgotten'. It implies that different transitions are uncorrelated. Therefore the correlation functions of the Langevin sources are proportional to the mean rates of the transitions (compare with Eq. (1.5.8) for the shot noise):

$$\langle \delta J_{1\to2}(t_1)\delta J_{1\to2}(t_2)\rangle = \delta(t_1 - t_2)J_{1\to2}, \qquad S_{J_{1\to2}}(f) = 2J_{1\to2},$$
$$\langle \delta J_{2\to1}(t_1)\delta J_{2\to1}(t_2)\rangle = \delta(t_1 - t_2)J_{2\to1}, \qquad S_{J_{2\to1}}(f) = 2J_{2\to1}, \qquad (1.10.13)$$
$$\langle \delta J_{1\to2}(t_1)\delta J_{2\to1}(t_2)\rangle = 0.$$

Under stationary conditions, the mean rates are equal: $J_{1\to2} = J_{2\to1} = W(1 \to 2)w_1 = W(2 \to 1)w_2$.

The Fourier transform of $\delta n_1(t)$ is found from Eq. (1.10.12):

$$\delta n_1(\omega) = (-i\omega + \tau^{-1})^{-1}[\delta J_{2\to1}(\omega) - \delta J_{1\to2}(\omega)]. \qquad (1.10.14)$$

The spectral density of a random variable $x(t)$ is proportional to $\langle x(\omega) \times x^*(\omega)\rangle$. Therefore

$$S_{n_1}(f) = \frac{S_{J_{1\to2}} + S_{J_{2\to1}}}{\omega^2 + \tau^{-2}} = \frac{2(J_{2\to1} + J_{1\to2})}{\omega^2 + \tau^{-2}} = \frac{4J_{1\to2}\tau^2}{1 + \omega^2\tau^2}. \qquad (1.10.15)$$

Since $S_x(f) = (x_1 - x_2)^2 S_{n_1}(f)$ and $W(1 \to 2)\tau = w_2$, the equation for $S_x(f)$ exactly coincides with Eq. (1.7.10).

The Langevin equations for fluctuations in other systems are derived in the same way. They have the form:

$$\hat{L}(t, x)\delta x(t) = y(t). \qquad (1.10.16)$$

Here $\hat{L}(t, x)$ is a linear operator acting on the fluctuation $\delta x(t)$. The r.h.s. $y(t)$, as in Eq. (1.10.3), is the Langevin source of fluctuations. Depending on the physical meaning of both the fluctuating quantity $x(t)$ and Eq. (1.10.16), the Langevin source may be a random force, random flux of particles, random current, etc.

It is assumed that the fluctuation $\delta x(t)$ is smoothed (averaged) over some time intervals Δt that satisfy two conditions (see below). The kinetics of such a smoothed fluctuation is the same as the kinetics of a small (linear) deviation $\langle \Delta x(t)\rangle$ of the same quantity from its stationary (e.g., equilibrium) value under a given, nonrandom perturbation. That is why the l.h.s. of the

Langevin equation for $\delta x(t)$ coincides with the linearized equation for the linear response $\langle \Delta x(t) \rangle$ to a small perturbation.

On one hand, the intervals of time, Δt, should be sufficiently short. Specifically, the expected change of the random variable, $\delta x(t + \Delta t) - \delta x(t)$, must be so small that time-derivatives can be used in the operator $\hat{L}(t, x)$. On the other hand, Δt should be much greater than the correlation time of the Langevin source, so that the correlation between $y(t + \Delta t)$ and $y(t)$ is negligible: the Langevin source randomly changes its sign many times during Δt. Thus, the Langevin approach, like the Markov approach (Sec. 1.6), is based on the existence of the same large gap between the relaxation time of the random quantity under consideration and the much smaller relaxation times, τ_c, of other random variables in the same system which can be considered, comparatively, as 'microscopic' ones. The meaning of the term 'microscopic' used here may differ from the meaning of this term in atomic physics: the quantity $x(t)$, from the standpoint of atomic physics, may also be considered as a microscopic quantity.

The Langevin source of fluctuations can be considered as delta-correlated, that is, its correlation function can be approximated by $\delta(t_1 - t_2)$ times a constant coefficient. If the system is in an equilibrium state, this coefficient is found, as in the example of Brownian motion, by equating the variance of the fluctuations (integral of the spectral density over all positive frequencies) to the corresponding expression known from statistical mechanics. In many nonequilibrium systems the source of fluctuations is similar to a sequence of uncorrelated random events like uncorrelated pulses (see Eq. (1.10.13)). Then Eq. (1.5.6) for the spectral density of this random process can be used.

Several advantages of the Langevin approach can be stated:

1. If the system is out of equilibrium, the one-time correlations (variances and covariances) $\psi_{\alpha\beta}(0)$, in terms of which the correlation functions $\psi_{\alpha\beta}^{+}(t)$ are expressed if the Markov approach is used (Sec. 1.6), can not be calculated from statistical-mechanical equations. In this case one has to derive and solve the equations for one-time correlation functions $\psi_{\alpha\beta}(0)$ apart from equations for $\psi_{\alpha\beta}^{+}(t)$. At the same time the Langevin source, as was mentioned above, is often similar to a sequence of uncorrelated random events. Its low-frequency spectral density, which has to be known if the Langevin approach is used, is proportional to the mean rate of these events (Eq. (1.5.6)). This spectral density can be found for all those systems, including nonequilibrium ones, for which the problem of mean quantities can be solved. Thus, for nonequilibrium systems the Langevin approach may be much simpler.

2. The Langevin approach can be applied to nonstationary systems if the char-

acteristic time of the system's variation is much larger than the correlation time τ_c of the Langevin sources.

3. If there are N coupled random variables x_α ($\alpha = 1, \ldots, N$), one has to solve N^2 equations for the functions $\psi^+_{\alpha\beta}(t)$ (Eq. (1.6.23)), and, in the case of nonequilibrium systems, also the equations for $\psi_{\alpha\beta}(0)$. Using the Langevin method, one has to solve only N inhomogeneous equations for N fluctuations $\delta x_\alpha(t)$ (Lax, 1966).

4. Consider two coupled systems in which the fluctuations are caused by internal Langevin sources. If the coupling between these systems is varied, the correlation functions of the fluctuations may change drastically. However, the spectral densities of the Langevin sources may remain unchanged (Lax, 1966).

5. If the random variable $x(\mathbf{r}t)$ depends not only on the time t but also on the radius-vector $\mathbf{r}$, its correlation function depends on the boundary conditions for the fluctuations. However, because the correlation function of the bulk Langevin sources, $\langle y(\mathbf{r}_1 t_1) y(\mathbf{r}_2 t_2) \rangle$, is proportional to $\delta(\mathbf{r}_1 - \mathbf{r}_2)$, it is independent of the boundary conditions. One and the same expression for the correlation function of the bulk Langevin sources can be used in problems with different boundary conditions, and, consequently, quite different correlation functions $\langle x(\mathbf{r}_1 t_1) x(\mathbf{r}_2 t_2) \rangle$.

In this book the Langevin method is widely used.