1. Brownian Motion

In 1827 the botanist *Brown* discovered under his microscope vigorous irregular motion of small particles originating from pollen floating on water [1.1]. He also observed that very fine particles of minerals undergo similar incessant motion as if they were living objects. This discovery must have been a great wonder at that time. The idea of combining such a motion – Brownian motion – with molecular motion became fairly widespread in the latter half of the nineteenth century when atomism had not yet been fully recognized as reality. It was the celebrated work of *Einstein*, which appeared in 1905, that gave the first clear theoretical explanation of such a phenomenon which could be directly verified quantitatively by experiments and thus established the very basic foundation of the atomic theory of matter [1.2]. Einstein did not known that Brownian motion had actually been observed many years before when he first came upon this idea to verify the reality of the atomic concept. At any rate, Einstein's theory had a great impact at that time, finally convincing people of the theory of heat as molecular motion, and so paved the way to modern physics of the twentieth century. It also greatly influenced pure mathematics, that is, the theory of stochastic processes.

The theory of stochastic processes, called Wiener processes, was initiated by N. Wiener as a mathematical model of Brownian motion. Some years later this was combined with Feynman's path integral formulation of quantum mechanics. R. P. Feynman did not know of Wiener's work when he devised this method independently. It is very instructive that such unconscious coincidences often arise at very decisive moments in the progress of science in seemingly far separated disciplines.

The theory of Brownian motion was further developed by P. Langevin, M. Smoluchowski, G. E. Uhlenbeck, L. S. Ornstein, and many others. The classical theory was excellently reviewed by *Wang* and *Uhlenbeck* [1.3]. The present treatment owes a great deal to this review, which still remains a standard reference. But our intention is to extend the theory a bit further and to base on it the developments of nonequilibrium statistical mechanics, treated in some detail in the following chapters. Grasping physical phenomena as stochastic processes is one of the very fundamental methods of this approach. Brownian motion is an excellent example of this.

Einstein proved that the diffusion constant D of a Brownian particle is related to its mobility μ by

$$D = \mu k T .$$

This relationship, which is called the *Einstein relation*, provides us with a very good basis of experimental verification that Brownian motion is in fact related to the thermal motion of molecules. It is very important to realize that this has an even deeper meaning. Namely, it is the first example of one of the most general theorems of statistical physics, called the *fluctuation-dissipation theorem* [1.4]. This theorem is, as shown in Sect. 1.6, one of the most fundamental cornerstones supporting statistical mechanics of nonequilibrium states.

Fig. 1.1. Brownian motion caused by incessant impacts of gas molecules against a mirror suspended in a dilute gas [1.5]

It turns out that Brownian motion is not only limited to the motion of small particles, but is actually very common. For example, irregular oscillatory motion of a small mirror suspended in a dilute gas is caused by incessant impacts of gas molecules (Fig. 1.1). In electrical circuits, thermal motion of electrons in conductors gives rise to fluctuations of electric currents as well as potential differences between different components. When suitably amplified, such fluctuations can be heard by the ear as so-called thermal noise. More generally, every physical quantity we observe is accompanied by similar fluctuations due to thermal motion of microscopic degrees of freedom in matter. In a great many cases, such fluctuations are small in comparison with the average values of the quantity under observation and can generally be ignored. However, such fluctuations reflect the microscopic motions in the system under study, so analyzing them provides very important keys for studying the system. In this chapter, the basic concepts of stochastic processes are discussed, taking mainly the Brownian motion of Brownian particles as the simplest example. Even though other examples are not mentioned explicitly, the reader should keep in mind that these concepts and methods are not confined to the simplest model but are general and applicable to a wide class of physical phenomena.

1.1 Brownian Motion as a Stochastic Process

Suppose that we observe a Brownian particle under a microscope over a time interval $0 \leqq t \leqq T$ and obtain a record of its position $x(t)$ as a function

of time. For simplicity, we consider in the following only the projection onto the x axis and treat the problem as one-dimensional motion, but the essentials are the same for three-dimensional motion. The observations are repeated in tme to get N readings of the particle position

$$x_1(t), x_2(t), \ldots, x_N(t) . \tag{1.1.1}$$

These readings are all different, that is, the motion of the Brownian particle is not reproducible.

Then we ask, "what can physics predict about Brownian motion?" Obviously, unlike in mechanics, we are not able to make deterministic predictions: we must rather take a probabilistic outlook. The value of the displacement $x(t)$ of the Brownian particle at time t is probabilistic and each of the observed series $x_j(t)$ is a sample from a statistical ensemble. If we repeat the observations a great many times to make N very large, we should be able to find empirically the distribution law obeyed by $x(t)$.

The stochastic variable (random variable) is $\boldsymbol{x}(t)$[1]. This is a series of random variables having t as a parameter. Such a time series of random variables is generally called a *stochastic process*. If a continuous observation is made, a function $x(t)$ with a continuous parameter t is obtained as a sample of the process. If observations are made at discrete times

$$0 < t_1 < t_2 \ldots < t_n < T, \tag{1.1.2}$$

then a set of n real numbers

$$x(t_1), x(t_2), \ldots, x(t_n)$$

is a sample obtained by the observations. If we regard the set as a vector, then an n-dimensional real space R^n is the sample space of the process $\boldsymbol{x}(t)$ for the selected time points (1.1.2). An element of the sample space may also be represented by a zig-zag path (Fig. 1.2). One may consider the limit of very large n and vanishing lengths of time segments to attain a path with a continuous time. This is an intuitive conclusion, not easily made rigorous mathematically. In fact, the introduction of a proper measure in the space consisting of all possible paths $x(t)$ $(0 \leqq t \leqq T)$ requires considerable mathematics, into which we shall not go here since the physical aspects of the problem are emphasized instead.

In understanding Brownian motion as such a stochastic process, how can $\boldsymbol{x}(t)$ be described in terms of probability theory? Firstly, what is the

---

1 A sample of the observed values of random variable $\boldsymbol{x}(t)$ is $x(t)$. This is similar to the relationship between an observable (dynamical quantity) and its observed value in quantum mechanics.

A bold-face italic letter is used for a random variable to distinguish it from its sample value. However, this special lettering will be dispensed with when there is no fear of confusion between these two concepts.

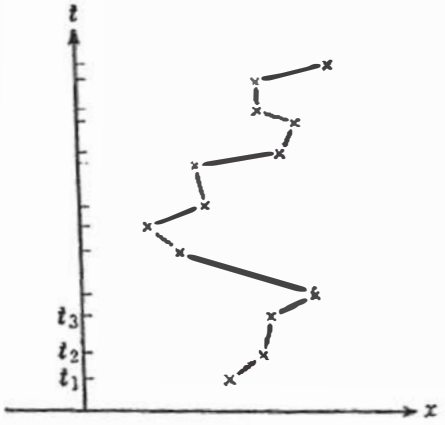


Fig. 1.2. A sample reading of the process $x(t)$ by intermittent observation

probability of finding an observed value $x(t)$ of $x(t)$ at time t in the interval between x and $x+dx$? If its probability density is W_1, then the probability is

$$W_1(x,t)\,dx = \Pr\{x < x(t) \leqq x+dx\}\,. \tag{1.1.3}$$

On the right-hand side, Pr {...} means the probability of the event in the curly bracket occurring. Next, what is the probability that two observed values $x(t_1)$ and $x(t_2)$ at times t_1 and t_2 are found in the intervals (x_1, x_1+dx_1) and (x_2, x_2+dx_2), respectively? For this, the probability density W_2 is defined by

$$\begin{aligned} &W_2(x_1,t_1;x_2,t_2)\,dx_1\,dx_2 \\ &\quad = \Pr\{x_1 < x(t_1) \leqq x_1+dx_1, x_2 < x(t_2) \leqq x_2+dx_2\}\,. \end{aligned} \tag{1.1.4}$$

More generally, for a set of observed values $x(t_1), x(t_2), \ldots x(t_n)$ at $t_1, t_2, \ldots$ and t_n

$$\begin{aligned} &W_n(x_1,t_1;x_2,t_2;\ldots;x_n,t_n)\,dx_1\,dx_2\ldots dx_n \\ &\quad = \Pr\{x_j < x(t_j) \leqq x_j+dx_j, j=1,2,\ldots,n\}\,. \end{aligned} \tag{1.1.5}$$

This is the joint probability distribution for n random variables, $x(t_1)$, $x(t_2), \ldots$, and $x(t_n)$. The stochastic process $x(t)$ is defined when such probabilities are given for any set of $n(n=1,2,\ldots\infty)$ time points. In other words, each possible path of the Brownian motion $x(t)$ has a probability assigned to it so that the probability (1.1.5) is defined as the sum of these probabilities for all possible paths going through the gates $dx_1, dx_2, \ldots dx_n$ set at the selected time points. This is the basic idea of constructing a path integral, but we shall not go into this topic here.

Many kinds of probabilities can be derived from definitions (1.1.3–5). Particularly important is the transition probability, which is defined as the probability that the Brownian particle is found at time t_1 between x_1 and

$x_1 + dx_1$ when it was at x_0 at time t_0:

$$P(x_0, t_0 | x_1, t_1)\, dx_1 = \frac{W_2(x_0, t_0; x_1, t_1)\, dx_1}{W_1(x_0, t_0)}. \tag{1.1.6}$$

Transition probabilities for two time points are most commonly used, but a more general definition of transition probabilities is

$$P(x_0, t_0 | x_1, t_1; \ldots; x_n, t_n)\, dx_1 \ldots dx_n = \frac{W_{n+1}(x_0, t_0; x_1, t_1; \ldots; x_n, t_n)\, dx_1 \ldots dx_n}{W_1(x_0, t_0)} \tag{1.1.7}$$

for n observations at n time points when the initial state x_0 is precisely defined at time t_0.

Brownian motions in reality are complex in many respects, so that idealization and abstraction are necessary to formulate them in physical or mathematical terms. There are many different levels of such idealization, each corresponding to a stage of our understanding of their physical nature. The primary purpose of the discussions in this chapter is to make the meaning of these levels as clear as possible. Let us now start from the most simplified model of Brownian motion.

Consider a medium which contains a large number of Brownian particles and define the particle density as $n(x,t)$. Brownian motion of particles makes the distribution of particles tend toward uniformity. This process is called *diffusion*. Corresponding to the gradient of the density distribution, a flow is produced

$$j_d = -D \frac{\partial n}{\partial x}, \tag{1.1.8}$$

which then induces a change of the density according to

$$\frac{\partial n(x, t)}{\partial t} = -\frac{\partial j_d}{\partial x} = D \frac{\partial^2 n}{\partial x^2}. \tag{1.1.9}$$

This is the *diffusion equation*. When a uniform force field such as gravitation exists, a uniform flow is produced with the terminal velocity u_0 determined by the balance of the driving force K and the frictional force from the surrounding fluid acting on a particle. This flow is denoted by j_K and is given by

$$j_K = n u_0 = \frac{nK}{m\gamma}, \tag{1.1.10}$$

where m is the mass of the particle and $m\gamma$ is the friction constant. Therefore the total flow is

$$j = j_K + j_d = \frac{nK}{m\gamma} - D \frac{\partial n}{\partial x}. \tag{1.1.11}$$

In the presence of the external field K, the diffusion equation becomes

$$\frac{\partial n(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left(\frac{nK}{m\gamma}\right) + D\frac{\partial^2 n}{\partial x^2} \tag{1.1.12}$$

instead of (1.1.9).

Starting from any distribution of density, the particles attain a final equilibrium distribution after a sufficiently long time. A uniform distribution of density arises when there is no external field of force, but in a gravitational field, sedimentation equilibrium arises, represented by

$$n(x) = n(x_0)\exp\left(\frac{K(x-x_0)}{kT}\right) \tag{1.1.13}$$

at temperature T of the fluid. This equilibrium is the balance of two flows j_d and j_K, a simple example of *detailed balance*. This means that the distribution (1.1.13) makes the flow j in (1.1.11) vanish. Therefore

$$\frac{D}{kT} = \frac{1}{m\gamma} \quad \text{or}$$

$$D = \mu kT \tag{1.1.14}$$

must hold. Here $\mu = 1/m\gamma$ is the mobility, which is the ratio of u_0 to the force K. Equation (1.1.14) is nothing but the *Einstein relation* given in the introductory section. This relationship is further discussed below; but note here that it is obtained from a very simple idea.

As long as the particle density is not too high, the interactions between Brownian particles can be ignored so that the diffusion described by (1.1.9 or 12) is the result of independent particle motion. Namely, the density $n(x,t)$ at time t and the spatial point x is

$$n(x,t) = \int n(x_0,t_0)\,dx_0\,P(x_0,t_0|x,t)\,, \tag{1.1.15}$$

where $n(x_0,t_0)$ is the density at t_0 and x_0. The transition probability $P(x_0,t_0|x,t)$ satisfies the diffusion equation (we consider Brownian motion in the absence of an external field)

$$\frac{\partial}{\partial t}P(x_0,t_0|x,t) = D\frac{\partial^2}{\partial x^2}P(x_0,t_0|x,t) \tag{1.1.16}$$

because (1.1.9) must be satisfied by $n(x,t)$ given by (1.1.15) for an arbitrary initial condition $n(x_0,t_0)$. Then (1.1.16) simply becomes

$$\frac{\partial}{\partial t}P(x,t) = D\frac{\partial^2}{\partial x^2}P(x,t)\,. \tag{1.1.17}$$

The transition probability $P(x_0, t_0|x, t)$ is the fundamental solution of (1.1.17) for the initial condition

$$P(x_0, t_0|x, t) = \delta(x - x_0) \tag{1.1.18}$$

and is given by

$$P(x_0, t_0|x, t) = \frac{1}{\sqrt{4\pi D(t-t_0)}} \exp\left(-\frac{(x-x_0)^2}{4D(t-t_0)}\right). \tag{1.1.19}$$

If boundaries or sources exist, some appropriate boundary conditions must be imposed.

This is the simplest possible idealization of Brownian motion. The probability of finding a particle at x at time t when it was certainly located at x_0 at t_0 is independent of the knowledge of where the particle was before t_0. Its history previous to time t_0 is summarized, so to say, in the information that the particle was located at x_0 at time t_0, expressed by

$$P(x', t'; x_0, t_0|x, t) = P(x_0, t_0|x, t) \quad (t' < t_0) \tag{1.1.20}$$

and hence

$$P(x_0, t_0|x_1, t_1; x_2, t_2)$$

$$= P(x_0, t_0|x_1, t_1)\, P(x_1, t_1|x_2, t_2) \quad (t_0 < t_1 < t_2). \tag{1.1.21}$$

Namely, the evolution of the process in the time interval (t_0, t_2) can be constructed by evolution in the two intervals (t_0, t_1) and (t_1, t_2), where t_1 is an arbitrary time point between t_0 and t_2. Therefore, integrating over all possible values of x_1 at t_1 gives

$$P(x_0, t_0 \mid x_2, t_2) = \int P(x_0, t_0 \mid x_1, t_1)\, dx_1\, P(x_1, t_1 \mid x_2, t_2)\,. \tag{1.1.22}$$

Generally, a stochastic process $x(t)$ is called *Markovian* if it satisfies conditions (1.1.21, 22). That the Brownian motion defined by (1.1.16 or 17) is Markovian is a consequence of the fact that these equations are first order with respect to t. It can be easily proved directly that the transition probability (1.1.19) satisfies (1.1.22).

In the presence of an external force field, (1.1.17) is replaced by

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial x}(\mu K P) + D\frac{\partial^2}{\partial x^2} P\,,$$

but everything stated above remains true, with only minor modifications.

1.2 The Central Limit Theorem and Brownian Motion

The probability distribution of the displacement $X = x - x_0$ over the time interval $(0, t)$ (for simplicity, $t_0 = 0$) (1.1.19) is a *normal distribution*, or a *Gaussian distribution*, and its variance grows in proportion to time:

$$\langle X^2 \rangle = 2Dt . \tag{1.2.1}$$

Now the time interval $(0, t)$ is divided into n ($\geqslant 1$) segments Δt_i ($i = 1, 2, \ldots, n$) and displacements in each segment are denoted by ΔX_i. Then naturally

$$X = \sum_{i=1}^{n} \Delta X_i \quad \text{and} \tag{1.2.2}$$

$$\langle \Delta X_i \rangle = 0 . \tag{1.2.3}$$

The expectation of the total displacement is zero

$$\langle X \rangle = 0 .$$

Further, displacements in different time segments are statistically independent, as implied by (1.1.21):

$$\langle \Delta X_i \Delta X_j \rangle = 0 \qquad (i \neq j) . \tag{1.2.4}$$

Therefore it follows from (1.2.2) that

$$\langle X^2 \rangle = \sum_{i=1}^{n} \langle \Delta X_i^2 \rangle . \tag{1.2.5}$$

Taking, for simplicity, equal lengths for the time segments, $\langle \Delta X_i^2 \rangle$ are then all equal so that

$$\langle X^2 \rangle = n \langle \Delta X^2 \rangle = t \frac{\langle \Delta X^2 \rangle}{\Delta t} .$$

Comparing this with (1.2.1) gives for the diffusion constant

$$D = \frac{\langle \Delta X^2 \rangle}{2\Delta t} . \tag{1.2.6}$$

As long as the diffusion model of Sect. 1.1 is true for the displacement over each time interval Δt, the result (1.2.6) is simply a repetition of (1.2.1). However, the above consideration has a deeper meaning.

The well-known Gaussian law of errors teaches us that an observation error X follows a normal distribution if the error is an accumulation of a large number of small errors. The displacement X of a Brownian particle is also a sum of a large number of successive small displacements ΔX_i. Therefore, we should expect that the distribution law of displacement X over a sufficiently long time interval t is normal even if the diffusion equation (1.1.17) loses its validity for displacements in too short a time interval Δt [in

fact, Sect. 1.6 shows that (1.1.17, 19) are not valid in general for short time intervals], and that the variance $\langle X^2 \rangle$ is given by (1.2.1) with the diffusion constant D defined by (1.1.14). Then the transition probability $P(x_0, t_0 \mid x, t)$ has the form (1.1.19) and satisfies the diffusion equation (1.1.16).

The Gaussian law of errors is contained in a very general theorem of probability theory called the *central limit theorem*, which is of fundamental importance in statistical physics. Therefore let us discuss it with some generality. In the same way as for (1.2.2), consider a sum of $n (\gg 1)$ independent random variables $\Delta X_1, \Delta X_2, \ldots \Delta X_n$ and set

$$X_n = \Delta X_1 + \Delta X_2 + \ldots + \Delta X_n . \tag{1.2.7}$$

Here $\Delta X_1, \Delta X_2, \ldots$ and ΔX_n are assumed to have zero expectation values as (1.2.3) and the variances

$$\langle \Delta X_j^2 \rangle = \sigma_j^2 .$$

We set

$$s_n^2 = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2 . \tag{1.2.8}$$

The central limit theorem now states that if a certain appropriate condition is met by the random variables $\Delta X_1, \Delta X_2, \ldots \Delta X_n$, the probability distribution of the random variable

$$Y_n = \frac{X_n}{s_n} \tag{1.2.9}$$

approaches a normal distribution with variance equal to 1; namely, its distribution density $f_n(Y_n)$ tends to $f(Y)$ as

$$f_n(Y) \to \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} Y^2\right) \tag{1.2.10}$$

asymptotically as n increases to infinity. Therefore the probability distribution density $P(X_n)$ of X_n has the property

$$P(X_n) \approx \frac{1}{\sqrt{2\pi s_n}} \exp\left(-\frac{X_n^2}{2s_n^2}\right) \quad (n \gg 1) . \tag{1.2.11}$$

This is the Gaussian law of errors already mentioned.

The essential point of the conditions for the validity of the central limit theorem is that n random variables $\Delta X_1, \Delta X_2, \ldots$ and ΔX_n are all alike and there are no eminent few that dominate the others. When expressed mathematically, this qualitative condition is formulated as various kinds of sufficient conditions. There are different forms of the central limit theorem, such as Lindeberg's or Ljapunov's, for which the reader is referred to mathematical textbooks [1.6]. Here our treatment is greatly simplified by introducing rather restricting conditions.

For a general treatment of such problems, it is most convenient to use the *characteristic function*. The characteristic function for a random variable x is defined by

$$\Phi(\xi) = \langle e^{i\xi x} \rangle . \tag{1.2.12}$$

In particular, if the probability distribution density $f(x)$ exists, it is expressed by

$$\Phi(\xi) = \int_{-\infty}^{\infty} e^{i\xi x} f(x)\, dx, \tag{1.2.13}$$

which is nothing but the Fourier transform of $f(x)$. Then $f(x)$ is obtained from $\Phi(\xi)$ as its inverse Fourier transform. But it is not necessary that the density function $f(x)$ exist. General theorems of probability theory state that the characteristic function $\Phi(\xi)$ exists even if the density function f does not, and that the probability distribution of x is uniquely determined from the knowledge of $\Phi(\xi)$ [1.7].

If two random variables x and y are independent, then obviously

$$\langle e^{i\xi(x+y)} \rangle = \langle e^{i\xi x} \rangle \langle e^{i\xi y} \rangle . \tag{1.2.14}$$

More generally, the characteristic function of a sum of an arbitrary number of independent random variables is equal to the product of characteristic functions of the respective random variables. This is one of the basic properties of the characteristic function. The partition function introduced in [Ref. 1.8, Chap. 2] as the fundamental function in equilibrium statistical mechanics is a kind of characteristic function for an unnormalized probability distribution of microscopic variables (where a real parameter $-\beta$ was used instead of imaginary $i\xi$).

If the moments

$$\langle x^n \rangle = \int_{-\infty}^{\infty} x^n f(x)\, dx \qquad (n = 0, 1, 2, \ldots) \tag{1.2.15}$$

exist for all n's, the characteristic function $\Phi(\xi)$ is analytic in the neighborhood of $\xi = 0$ and is expanded as

$$\Phi(\xi) = \sum_{n=0}^{\infty} \frac{(i\xi)^n}{n!} \langle x^n \rangle . \tag{1.2.16}$$

Conversely, the moment $\langle x^n \rangle$ is then obtained from $\Phi(\xi)$ as

$$\frac{1}{i^n} \left[\left(\frac{d}{d\xi} \right)^n \Phi(\xi) \right]_{\xi=0} = \langle x^n \rangle . \tag{1.2.17}$$

However, it should be remembered that the moments do not necessarily exist. For example, for the Cauchy distribution

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2} , \tag{1.2.18}$$

the second and higher moments are all divergent. Correspondingly, the characteristic function is not analytic at $\xi = 0$ as is clear from

$$\Phi(\xi) = e^{-|\xi|}. \tag{1.2.19}$$

For such an expansion as (1.2.16) to be possible, it is necessary that the distribution function $f(x)$ tends to zero sufficiently fast as $x \to \pm\infty$.

For example, the characteristic function for the normal distribution

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \tag{1.2.20}$$

is calculated as follows:

$$\begin{aligned}\Phi(\xi) &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-m)^2}{2\sigma^2} + ix\xi\right) dx \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp\left[im\xi - \frac{\sigma^2}{2}\xi^2 - \frac{1}{2\sigma^2}(x - m - i\sigma^2\xi)^2\right] dx .\end{aligned}$$

Here note the equality

$$\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(y-a)^2\right] dy = \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2}\right) dy = \sqrt{2\pi} \tag{1.2.21}$$

obtained by shifting the path of integration as shown in Fig. 1.3 from the path AB on the real axis to CD parallel to AB through point a. Since the function $\exp(-y^2/2)$ is analytic everywhere on the complex plane, the difference of the integrations is due only to integrations along AC and BD, but these vanish as A and B are pushed to $-\infty$ and ∞ respectively. Therefore

$$\Phi(\xi) = \exp\left(im\xi - \frac{\sigma^2}{2}\xi^2\right) \tag{1.2.22}$$

is the characteristic function for the normal distribution. This simple result is worth remembering.

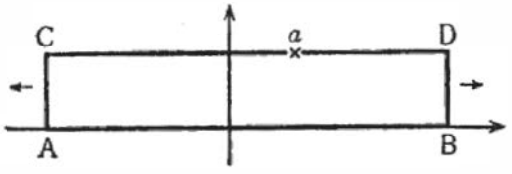


Fig. 1.3. The integration path to prove (1.2.21)

A cumulant function $\Psi(\xi)$ is defined by

$$\Phi(\xi) = e^{\Psi(\xi)}. \tag{1.2.23}$$

This corresponds to the thermodynamic characteristic function (free energy divided by kT) in statistical mechanics. It is written as

$$\Psi(\xi) = \ln \Phi(\xi). \tag{1.2.24}$$

If the expansion (1.2.16) is possible, this may be expanded to

$$\Psi(\xi) = \sum_{n=1}^{\infty} \frac{(i\xi)^n}{n!} \langle x^n \rangle_c , \tag{1.2.25}$$

where the expansion coefficient $\langle x^n \rangle_c$ is called the nth cumulant and is related to the moments $\langle x^m \rangle$ $m \leqq n$ by (1.2.23 or 24). Explicit relations are

$$\begin{aligned} &\langle x \rangle_c = \langle x \rangle , \\ &\langle x^2 \rangle_c = \langle x^2 \rangle - \langle x \rangle^2 , \quad \langle x^2 \rangle = \langle x^2 \rangle_c + \langle x \rangle_c^2 \\ &\langle x^3 \rangle_c = \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2\langle x \rangle^3 \\ &\langle x^3 \rangle = \langle x^3 \rangle_c + 3\langle x \rangle_c \langle x^2 \rangle_c + \langle x \rangle_c^3 \end{aligned} \tag{1.2.26}$$

for $n \leqq 3$. Generally, the nth cumulant is expressed in terms of moments not higher than the nth. Conversely, the nth moment is expressed in terms of cumulants not higher than the nth. In particular, $\langle x \rangle_c$ is the expectation and $\langle x^2 \rangle_c$ is the variance.

For normal distribution (1.2.20), from (1.2.22)

$$\langle x \rangle_c = m , \quad \langle x^2 \rangle_c = \sigma^2 , \quad \langle x^n \rangle_c = 0 \qquad (n \geqq 3) . \tag{1.2.27}$$

A characteristic feature of the normal distribution is that cumulants higher than the third are all zero.

Now the characteristic function of X_n, (1.2.7), is

$$\Phi(\xi) = \langle e^{i\xi X_n} \rangle = \prod_{j=1}^{n} \langle e^{i\xi \Delta X_j} \rangle . \tag{1.2.28}$$

Denoting the cumulant function by $\psi_j(\xi)$ for ΔX_j $(j = 1, 2, \ldots)$ and that by $\Psi_n(\xi)$ for X_n, then from (1.2.28)

$$\Psi_n(\xi) = \sum_{j=1}^{n} \psi_j(\xi) .$$

If the expansions

$$\psi_j(\xi) = i\xi \langle \Delta X_j \rangle_c - \frac{\xi^2}{2} \langle \Delta X_j^2 \rangle_c + \frac{(i\xi)^3}{3!} \langle \Delta X_j^3 \rangle_c + \ldots \tag{1.2.29}$$

are all possible, then

$$\Psi_n(\xi) = -\frac{\xi^2}{2} s_n^2 + \frac{(i\xi)^3}{3!} \sum_{j=1}^{n} \langle \Delta X_j^3 \rangle_c + \ldots \tag{1.2.30}$$

from (1.2.7, 8) by assuming $\langle \Delta X_j \rangle = 0$. The characteristic function for Y_n, (1.2.9),

$$\langle e^{i\eta Y_n} \rangle = \langle e^{i\eta X_n / s_n} \rangle ,$$

is

$$\langle e^{i\eta Y_n} \rangle = \exp\left(-\frac{1}{2}\eta^2 + \frac{(i\eta)^3}{3!} \sum_{j=1}^{n} \frac{\langle \Delta X_j^3 \rangle_c}{s_n^3} + \ldots \right) \tag{1.2.31}$$

obtained by replacing ξ in (1.2.30) by η/s_n. Assume that the mth moments of ΔX_j are all finite and of the same order of magnitude. Then s_n^2 increases, according to (1.2.8), in the order of n with increasing n. Then the mth cumulant in (1.2.31) tends to zero

$$\frac{O(n)}{O(n^{m/2})} \to 0$$

from $m \geqq 3$. Therefore

$$\langle e^{i\eta Y_n} \rangle \to e^{-\eta^2/2} \tag{1.2.32}$$

This shows, as noted previously, that Y_n approaches a normal distribution with the variance equal to 1.

In the above we have imposed a very strict condition, namely the existence of moments at all orders, which is not in fact necessary for proving the central limit theorem. However, this is not unreasonable to assume for many physical processes. Whether or not this is allowed in reality, the central limit theorem indicates that the probabilistic motion of a Brownian particle is, for a sufficiently long time, described very well by a diffusion process defined by (1.1.19). For shorter times, there is no reason to expect the diffusion process to be valid for a physical process of particle motion. If a particle moves with a velocity u at a certain time t, we have to wait a finite time before we find different velocities. This time τ_c is the *correlation time* of the velocity. However, when the time segment Δt is much longer than the correlation time τ_c, displacements in each time segment are regarded as independent of each other. So, if t is so long that $n = t/\Delta t$ is much larger than 1, the total displacement X in (1.2.2) must have a normal distribution (1.1.19) with the variance

$$\langle X^2 \rangle = 2Dt\,. \tag{1.2.33}$$

This is a consequence of the central limit theorem.

The random walk problem is often considered as a model of Brownian motion. The simplest model is random walk with steps $\pm a$ to the right or to the left randomly at every τ. After n steps the displacement $x = ma$ has the binomial distribution

$$P_n(m) = \frac{n!}{2^n}\left(\frac{n+m}{2}!\,\frac{n-m}{2}!\right)^{-1} \tag{1.2.34}$$

When n is large, this is approximated by a normal distribution. From (1.2.6) the diffusion constant is

$$D = \frac{a^2}{2\tau}\,. \tag{1.2.35}$$

This is also easily seen by using the Stirling formula valid for (1.2.34). The relationship (1.2.35) is, however, more general and not limited to any

particular random walk model. In three dimensions, it becomes

$$D=\frac{l^2}{6\tau}, \tag{1.2.36}$$

where l is the mean free path and τ the mean free time. Equation (1.2.35) for a one-dimensional case is recovered by setting $l^2=3a^2$. Depending on the way in which definitions are made, extra factors may be introduced into these equations. Most important is the physical meaning of the diffusion constant implied by these equations.

1.3 The Langevin Equation and Harmonic Analysis

So far we have concentrated on the displacement of Brownian particles. However, the physical model should start from the motion itself. The equation of motion of a particle is, of course,

$$m\frac{du}{dt}=F, \tag{1.3.1}$$

where u is the velocity and F the force acting on the particle from molecules of the fluid surrounding the Brownian particle. As before, we treat one-dimensional cases, but three-dimensional cases are the same.

The force F may be divided into two parts. The first part is the *frictional force* and is proportional to the velocity of the particle. If the frictional coefficient is denoted by $m\gamma$ as in (1.1.10), the frictional force is assumed to be

$$F_u=-m\gamma u. \tag{1.3.2}$$

If the Stokes law is assumed for a spherical particle, for example, the frictional coefficient is

$$m\gamma=6\pi a\eta, \tag{1.3.3}$$

where a is the radius of the particle and η the viscosity of the fluid.

The second part of the force is the remainder of the force, F_u subtracted from F, and is regarded as random, independent of the motion of the particle. This part is called the *random force* and is hereafter denoted as $R(t)$.

Then (1.3.1) is written as

$$m\frac{du}{dt}=-m\gamma u+R(t). \tag{1.3.4}$$

as a stochastic equation.

The same consideration can be applied to Brownian motion in the presence of a force field, for example, the gravitational field or a harmonic force binding the particle elastically to the origin. If the potential of the force is denoted by V, the equation of motion becomes

$$\frac{d\boldsymbol{p}}{dt} = -\frac{\partial V}{\partial x} - \gamma \boldsymbol{p} + \boldsymbol{R}(t), \tag{1.3.5}$$

$$\frac{dx}{dt} = \frac{\boldsymbol{p}}{m} \equiv \boldsymbol{u}(t). \tag{1.3.6}$$

A set of equations of motion containing a random force, like (1.3.4) or (1.3.5, 6), is called a Langevin equation.

The random force $\boldsymbol{R}(t)$ is a stochastic process randomly changing in time. Brownian motion, $\boldsymbol{u}(t)$ [or $\boldsymbol{p}(t)$] and $x(t)$ are also stochastic processes driven by [generated by] $\boldsymbol{R}(t)$. They are related to $\boldsymbol{R}(t)$ by (1.3.4 or 5, 6). If we consider the force as causing the motion, then the random force $\boldsymbol{R}(t)$ produces Brownian motion. Thus our problem is to determine the stochastic processes $\boldsymbol{u}(t)$ and $x(t)$ from knowing $\boldsymbol{R}(t)$. This is what is meant by solving stochastic equations like (1.3.4) or (1.3.5, 6).

A standard method of solving a stochastic equation is harmonic analysis. This expresses a motion by a superposition of oscillating functions, which is an orthodox method for linear systems. Since the environment is considered to be in a stationary condition with constant temperature and pressure, the Brownian motion must also be stationary if the particle is aged (kept for a sufficiently long time) in the environment. In other words, the probability such as (1.1.5) must be invariant with respect to a shift in time:

$$W_n(x_1, t_1; x_2, t_2; \ldots; x_n, t_n) = W_n(x_1, t_1 + \tau; x_2, t_2 + \tau; \ldots; x_n, t_n + \tau). \tag{1.3.7}$$

The processes $\boldsymbol{R}(t)$, $\boldsymbol{u}(t)$ and $x(t)$ are stationary in this sense.

Let us consider in general a stationary process $z(t)$, a sample of which is $z(t)$ observed over a time interval $0 \leqq t \leqq T$. This function $z(t)$ is expanded in a Fourier series as

$$z(t) = \sum_{n=-\infty}^{\infty} a_n e^{i\omega_n t}, \tag{1.3.8}$$

where the frequencies are

$$\omega_n = \frac{2\pi n}{T} \quad (n = 0, \pm 1, \pm 2, \ldots), \tag{1.3.9}$$

corresponding to the interval T. The function $z(t)$ is considered real so that the Fourier coefficient a_n has the form

$$a_n = a_n' + i a_n'', \quad a_{-n} = a_n^* = a_n' - i a_n''. \tag{1.3.10}$$

The stochastic process $z(t)$ is now expanded as

$$z(t) = \sum_{n=-\infty}^{\infty} a_n e^{i\omega_n t} \quad (0 \leqq t \leqq T). \tag{1.3.11}$$

Then the Fourier coefficient a_n in (1.3.8) is a sample of the random variable a_n defined by

$$a_n = \frac{1}{T}\int_0^T z(t)\, e^{-i\omega_n t}\, dt. \tag{1.3.12}$$

Equation (1.3.11) shows that the stochastic process $z(t)$ is expressed by a set of a countably infinite number of random variables $\{a_n\}$. For the process $z(t)$, the probability is defined as explained in Sect. 1.2 and so the probabilities and expectations are also defined for $\{a_n\}$. The expectation value of a_n is given by

$$\langle a_n \rangle = \frac{1}{T}\int_0^T \langle z(t) \rangle\, e^{-i\omega_n t}\, dt.$$

It follows that

$$\langle a_n \rangle = 0 \quad (n \neq 0), \tag{1.3.13}$$

since $\langle z(t) \rangle$ is a constant as it is stationary. For $n = 0$, obviously

$$\langle a_0 \rangle = \frac{1}{T}\int_0^T \langle z(t) \rangle\, dt = \langle z \rangle. \tag{1.3.14}$$

A sample of a_0 is the time average of a sample $z(t)$ of the process $z(t)$ over the interval $(0, T)$:

$$a_0 = \overline{z(t)}^T \equiv \frac{1}{T}\int_0^T z(t)\, dt. \tag{1.3.15}$$

Generally, this need not be equal to $\langle z \rangle$. A process $z(t)$ is called *ergodic* if

$$\lim_{T\to\infty} \overline{z(t)}^T = \langle z \rangle \tag{1.3.16}$$

holds.

On the other hand, the expectation $\langle z \rangle$ is the average of $z(t)$ over the set of *all* possible values of z. This set is not necessarily identical with the set of values of $z(t)$ for $0 \leqq t \leqq T$. It may well be that different subsets are reached from a different starting state $z(0)$, in which case the time average becomes different for different $z(0)$. For example, a combination of two distinct processes is obviously nonergodic. A nonergodic process is generally decomposed into ergodic processes. In most cases, a stationary stochastic process is already reduced to such a simple one, so that we may hereafter assume ergodicity. In this sense, $a_0 = \langle z \rangle = \text{const}$ and we consider $z(t) - \langle z \rangle$ to

assume that

$$\langle a_n \rangle = 0 \quad (n = 0, \pm 1, \ldots, \pm \infty) \tag{1.3.17}$$

without losing generality. The averaged strength of the Fourier component a_n is defined by

$$\langle |a_n|^2 \rangle = \langle |a_n'|^2 \rangle + \langle |a_n''|^2 \rangle. \tag{1.3.18}$$

The right-hand side of this expression is the sum of mean square averages of the real and imaginary parts of the amplitude a_n. When a suitable filter is used to select only the frequencies lying in the interval $\Delta\omega$, the strengths of these Fourier components are observable. The average intensity $I(\omega)$ is

$$I(\omega)\Delta\omega = \sum_{\omega_n \text{ in } \Delta\omega} \langle |a_n|^2 \rangle. \tag{1.3.19}$$

The right-hand side of this expression is a sum over all frequencies contained in the frequency band. The number of such modes is

$$\frac{\Delta\omega}{2\pi/T} = \frac{T}{2\pi}\Delta\omega$$

because the interval $\delta\omega$ of successive frequencies in (1.3.9) is $2\pi/T$. We may suppose $\langle a_n^2 \rangle$ to be continuous in the frequency ω_n and write (1.3.19) as

$$I(\omega) = \lim_{T\to\infty} \frac{T}{2\pi} \langle |a_n|^2 \rangle \tag{1.3.20}$$

to define the intensity spectrum of the process $z(t)$ at frequency ω. In fact, if $z(t)$ is the noise voltage produced between two terminals of an electrical network, $I(\omega)$ is the intensity of the noise heard by filtering frequencies to a narrow bandwidth $\Delta\omega$ around ω. In this sense, $I(\omega)$ defined by (1.3.20) is called the power spectrum of the process $z(t)$.

The power spectrum is obtained by application of the well-known *Wiener-Khintchine* theorem [1.9]. The correlation function for the process $z(t)$ is

$$\phi(t) = \langle z(t_0)\, z(t_0 + t) \rangle. \tag{1.3.21}$$

This represents the correlation between the observed values of $z(t)$ at two points in time t_0 and $t_0 + t$ and is independent of t_0 since $z(t)$ is stationary. The Wiener-Khintchine theorem asserts that

$$I(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(t)\, e^{-i\omega t}\, dt \tag{1.3.22}$$

holds. Conversely, it holds that

$$\phi(t) = \int_{-\infty}^{\infty} I(\omega)\, e^{i\omega t}\, d\omega\,. \tag{1.3.23}$$

This theorem is proved as follows. From (1.3.12)

$$\langle |a_n|^2 \rangle = \frac{1}{T^2} \int_0^T dt_1 \int_0^T dt_2 \langle z(t_1)\, z(t_2) \rangle \exp\left[-i\omega_n (t_1 - t_2)\right]. \tag{1.3.24}$$

In the integrand, the correlation function is dependent only on the time difference $t_1 - t_2$. The integration over t_1 and t_2 is carried out over the square $0 \leqq t_1 \leqq T$, $0 \leqq t_2 \leqq T$ as shown in Fig. 1.4. The integration is divided into two parts. For $t_1 > t_2$ the integration variables are changed to $t_1 - t_2 = t$ and t_2, for which the Jacobian $\partial(t_1, t_2)/\partial(t, t_2) = 1$. Integration over t_2 is carried out from 0 to $T - t$, resulting in

$$\int_0^T (T - t)\, \phi(t)\, e^{-i\omega_n t}\, dt, \tag{1.3.25}$$

since the integrand does not depend on t_2. Similarly, integration for the part $t_1 < t_2$ yields

$$\int_0^T (T - t)\, \phi(-t)\, e^{i\omega_n t}\, dt. \tag{1.3.26}$$

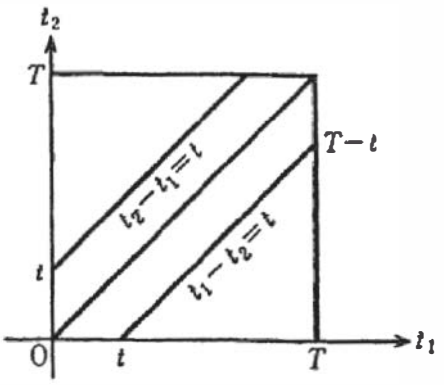


Fig. 1.4. Changing the integration variables in (1.3.24)

These results are inserted into (1.3.24) and T is set to infinity. As long as the integrals

$$\int_0^\infty \phi(\pm t)\, e^{\mp i\omega_n t}\, dt, \quad \int_0^\infty t\phi(\pm t)\, e^{\mp i\omega_n t}\, dt \tag{1.3.27}$$

are convergent, then

$$I(\omega) = \frac{1}{2\pi} \left[\int_0^\infty \phi(t)\, e^{-i\omega t}\, dt + \int_0^\infty \phi(-t)\, e^{i\omega t}\, dt \right]. \tag{1.3.28}$$

Converting the second term to an integral over $(-\infty, 0)$ gives (1.3.22).

Instead of Fourier-analyzing over a finite time, we may Fourier-analyze over an infinite time interval and write

$$z(t) = \int_{-\infty}^{\infty} a(\omega)\, e^{i\omega t}\, d\omega \tag{1.3.29}$$

$$a(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} z(t)\, e^{-i\omega t}\, dt \tag{1.3.30}$$

as Fourier integrals. If $z(t)$ is real,

$$a(-\omega) = a^*(\omega)$$

is the complex conjugate of $a(\omega)$. Furthermore, if $z(t)$ is stationary, it holds that

$$\langle a(\omega)\, a(\omega')\rangle = I(\omega)\, \delta(\omega + \omega') \tag{1.3.31}$$

or

$$\langle a(\omega)\, a^*(\omega')\rangle = I(\omega)\, \delta(\omega - \omega')\,, \tag{1.3.32}$$

where $I(\omega)$ is the power spectrum given by (1.3.20). In fact, the left-hand side of (1.3.31) is calculated as

$$\begin{aligned}\langle a(\omega_1)\, a(\omega_2)\rangle &= \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^{\infty} dt_2\, \phi(t_2 - t_1) \exp(-i\omega_1 t_1 - i\omega_2 t_2) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \phi(t) \exp(-i\omega_1 t)\, dt \cdot \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp[-i(\omega_1 + \omega_2)\tau]\, d\tau .\end{aligned}$$

With (1.3.22) we obtain (1.3.31). It is obvious that (1.3.23) follows conversely from (1.3.29, 31).

Harmonic analysis is particularly useful for linear stochastic equations like (1.3.4). We Fourier-expand the random force as a stationary stochastic process,

$$\boldsymbol{R}(t) = \sum_{n=-\infty}^{\infty} \boldsymbol{R}_n\, e^{i\omega_n t}. \tag{1.3.33}$$

In the same way the velocity $\boldsymbol{u}(t)$ of a Brownian particle is expanded as

$$\boldsymbol{u}(t) = \sum_{n=-\infty}^{\infty} \boldsymbol{u}_n e^{i\omega_n t}. \tag{1.3.34}$$

Then the stochastic differential equation (1.3.4)

$$m\dot{\boldsymbol{u}}(t) = -m\gamma \boldsymbol{u}(t) + \boldsymbol{R}(t) \tag{1.3.35}$$

is converted into the relation

$$\boldsymbol{u}_n = \frac{1}{i\omega_n + \gamma} \frac{\boldsymbol{R}_n}{m} \tag{1.3.36}$$

between the Fourier components. If we write the power spectra of $R(t)$ and $u(t)$ as I_R and I_u, respectively, we immediately obtain from (1.3.20, 36)

$$I_u(\omega) = \frac{1}{|i\omega + \gamma|^2} \frac{I_R(\omega)}{m^2} = \frac{1}{\omega^2 + \gamma^2} \frac{I_R(\omega)}{m^2} . \tag{1.3.37}$$

If the Brownian particle is bound to the origin by an elastic force, the Langevin equation is assumed to be

$$\frac{d^2x}{dt} + \gamma \frac{dx}{dt} + \omega_0^2 x = \frac{1}{m} R(t) , \tag{1.3.38}$$

where x is the displacement and ω_0 is the natural frequency of the elastic binding. Torsional oscillation of a small mirror suspended in a dilute gas is an example to which harmonic analysis is also directly applicable. The power spectrum I_x of the displacement is easily seen to be

$$\begin{aligned} I_x(\omega) &= \frac{1}{|\omega_0^2 - \omega^2 + i\gamma\omega|^2} \frac{I_R(\omega)}{m^2} \\ &= \frac{1}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \frac{I_R(\omega)}{m^2} \end{aligned} \tag{1.3.39}$$

in the same way as we derived (1.3.37).

The Wiener-Khintchine theorem (1.3.22, 23) shows that knowing the power spectrum is equivalent to knowing the correlation function. When $I_R(\omega)$ is known, (1.3.37 or 39) converts it into $I_u(\omega)$ or $I_x(\omega)$, so that this solves (1.3.35 or 38) to the same extent. As shown in Sect. 1.2, the solution is complete if $R(t)$ is a Gaussian process.

The nature of the random force $R(t)$ is considered in more detail in the following sections. Here we make the simplest possible assumption that the power spectrum I_R is independent of frequency

$$I_R(\omega) = I_R = \text{const} , \tag{1.3.40}$$

when the spectrum is said to be *white*. Obviously, from (1.3.23) it follows that the correlation function of a process having a white power spectrum has a vanishingly short correlation time,

$$\phi_R(t_1 - t_2) \equiv \langle R(t_1) R(t_2) \rangle = 2\pi I_R \delta(t_1 - t_2) . \tag{1.3.41}$$

If the random force can be described by (1.3.41), it follows from (1.3.37) with (1.3.23) that

$$\phi_u(t) = \int_{-\infty}^{\infty} \frac{e^{i\omega t} d\omega}{\omega^2 + \gamma^2} \frac{I_R}{m^2} , \tag{1.3.42}$$

namely,

$$\langle u(t_1)\, u(t_2)\rangle = \frac{\pi I_R}{m^2\gamma}\, e^{-\gamma|t_1-t_2|}. \tag{1.3.43}$$

Thus, the correlation function of velocity of a free Brownian particle decays exponentially in time with the decay constant γ. In particular, for $t_1 = t_2$, (1.3.43) reduces to

$$\langle u^2\rangle = \frac{\pi I_R}{m^2\gamma}. \tag{1.3.44}$$

If the Brownian particle has been kept for a sufficiently long time in the fluid at temperature T, the equipartition law

$$m\langle u^2\rangle = kT \tag{1.3.45}$$

must hold for the energy distribution. For (1.3.44) to be consistent with this,

$$I_R = \frac{m\gamma kT}{\pi} \tag{1.3.46}$$

must hold. In other words, the random force $\boldsymbol{R}$ (if it has a white spectrum) must have the intensity given by (1.3.46) so that the Langevin equation (1.3.35) represents free Brownian motion in thermal equilibrium at temperature T.

For harmonic Brownian motion, (1.3.39) leads to

$$\begin{aligned}\phi_x(t) &= \int_{-\infty}^{\infty} \frac{e^{i\omega t}\, d\omega}{(\omega_0^2-\omega^2)^2+\gamma^2\omega^2}\,\frac{I_R}{m^2} \\ &= \frac{\pi I_R}{m^2\gamma\omega_0^2}\left(\cos\omega_1 t + \frac{\gamma}{2\omega_1}\sin\omega_1 t\right)e^{-\gamma t/2} \quad (t>0),\end{aligned} \tag{1.3.47}$$

if I_R is assumed to be constant (white). For a derivation of this result, the residues at the poles

$$\omega = \pm\frac{i}{2}\gamma \pm \omega_1, \quad \text{where} \quad \omega_1 = \left(\omega_0^2 - \frac{\gamma^2}{4}\right)^{1/2},$$

of the integrand are calculated. In the limit $t \to 0$, this reduces to

$$\phi_x(0) = \langle x^2\rangle = \frac{\pi I_R}{m^2\gamma\omega_0^2} = \frac{kT}{m\omega_0^2}. \tag{1.3.48}$$

Thus (1.3.46) again guarantees the equipartition law of energy.

1.4 Gaussian Processes

A general stochastic process is defined by giving the probabilities (1.1.5) for all possible sets of $t_1, t_2, \ldots, t_n$ $(n = 1, 2, \ldots)$. Probabilities of lower hierarchy are derived from those of higher hierarchy, but the latter generally contain new information not contained in the former. The situation becomes simpler for Markovian processes, in which all higher probabilities are determined by the transition probability $P(x_1, t_1 | x_2, t_2)$. This kind of stochastic process is considered below. Here we take up another class of simple processes, namely the Gaussian processes. This is an extension of the normal distribution discussed in Sect. 1.2 to stochastic processes. Just as a normal distribution is defined by its second moment or the variance, a Gaussian process is completely defined by the correlation function (1.3.21).

A stochastic process $z(t)$ is Gaussian if the probability distribution of its observed values $z_1, z_2, \ldots, z_n$ at n time points $t_1, t_2, \ldots, t_n$ is an n-dimensional Gaussian (normal) distribution; namely, W_n in (1.1.5) has the form

$$W_n(z_1, t_1; z_2, t_2; \ldots; z_n, t_n) = C \exp\left[-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n} a_{jk}(z_j - m_j)(z_k - m_k)\right], \tag{1.4.1}$$

where

$$m_j = \langle z_j \rangle \equiv \langle z(t_j) \rangle \tag{1.4.2}$$

is the expectation value of $z(t)$ at time t_j and the matrix

$$(a_{jk}) \equiv A \tag{1.4.3}$$

is positive definite. The elements of its inverse matrix A^{-1} are the correlation functions of the process $z(t)$

$$\begin{aligned}(A^{-1})_{jk} &= \langle (z_j - m_j)(z_k - m_k) \rangle \\ &= \langle [z(t_j) - \langle z(t_j) \rangle][z(t_k) - \langle z(t_k) \rangle] \rangle .\end{aligned} \tag{1.4.4}$$

In order to see this, we use the characteristic function explained in Sect. 1.2 in a slightly generalized form. We introduce the parameters $\zeta_1, \zeta_2, \ldots, \zeta_n$ corresponding to the n random variables $z_1, z_2, \ldots, z_n$ and write the characteristic function of (1.4.1) as

$$\Phi(\zeta_1, \ldots, \zeta_n) = \int_{-\infty}^{\infty} dz_1 \ldots \int_{-\infty}^{\infty} dz_n \, W_n(z_1, t_1; \ldots; z_n, t_n) \exp\left(\mathrm{i}\sum_{j=1}^{n} \zeta_j z_j\right). \tag{1.4.5}$$

For brevity, we use the vector notations[2]

$$z = (z_1, z_2, \ldots, z_n), \quad \zeta = (\zeta_1, \zeta_2, \ldots, \zeta_n)$$

2 Note that bold-faced letters are not used for these vectors to avoid confusion with random variables

and carry out the integration in the following way. Inserting (1.4.1) into W_n, the exponential function in (1.4.5) is rewritten as

$$\begin{aligned}&\exp\left[-\tfrac{1}{2}(z-m)A(z-m)+i\zeta z\right]\\&\quad=\exp\left(i\zeta m-\tfrac{1}{2}yAy+i\zeta y\right)\\&\quad=\exp\left(i\zeta m-\tfrac{1}{2}uAu-iuAv+\tfrac{1}{2}vAv+i\zeta u-\zeta v\right),\end{aligned}$$

setting

$$m=(m_1, m_2, \ldots, m_n), \quad z-m=y=u+iv.$$

Now we choose the vector v by the condition

$$Av=\zeta, \quad \text{namely} \quad v=A^{-1}\zeta.$$

Then the first-order term of u vanishes, and the integral becomes

$$\Phi(\zeta)=\exp\left(im\zeta-\tfrac{1}{2}\zeta A^{-1}\zeta\right)\int_{-\infty}^{\infty}du_1\ldots\int_{-\infty}^{\infty}du_n\,C\exp\left(-\tfrac{1}{2}uAu\right). \tag{1.4.6}$$

Integration along the real axes of $z_1, z_2, \ldots, z_n$ was here transformed to that along the real axes of $u_1, u_2, \ldots, u_n$ just as for (1.2.21). The integral can be explicitly calculated by orthogonal transformation to diagonalize the quadratic form, uAu. But this is not necessary, because we should have $\Phi=1$ for $\zeta_1=\zeta_2=\ldots=\zeta_n=0$, as is seen by the fact that W_n is normalized by the constant C. Therefore,

$$\Phi(\zeta_1, \zeta_2, \ldots, \zeta_n)=\exp\left(i\sum_{j=1}^{n}m_j\zeta_j-\frac{1}{2}\sum_{j=1}^{n}\sum_{k=1}^{n}(A^{-1})_{jk}\zeta_j\zeta_k\right). \tag{1.4.7}$$

The moment and cumulant definitions introduced by (1.2.15, 16, 25) can easily be generalized to an n-dimensional random variable $(z_1, z_2, \ldots z_n)$. Namely, the $(r_1, r_2, \ldots, r_n)$th moment is

$$\langle z_1^{r_1}z_2^{r_2}\ldots z_n^{r_n}\rangle=\int dz_1\ldots\int dz_n\,W(z_1, \ldots, z_n)\,z_1^{r_1}\ldots z_n^{r_n}, \tag{1.4.8}$$

[$W(z_1, z_2, \ldots, z_n)$ is the joint distribution of $z_1, z_2, \ldots, z_n$] and the characteristic function (1.4.5) is expanded in a power series

$$\Phi(\zeta)=\sum_{r_1=0}^{\infty}\ldots\sum_{r_n=0}^{\infty}\frac{(i\zeta_1)^{r_1}\ldots(i\zeta_n)^{r_n}}{r_1!\ldots r_n!}\langle z_1^{r_1}\ldots z_n^{r_n}\rangle. \tag{1.4.9}$$

This gives all the moments, provided that such an expansion is possible. The cumulant function $\Psi(\zeta)$ is defined by

$$\Phi(\zeta)=\exp\Psi(\zeta), \quad \Psi(\zeta)=\ln\Phi(\zeta). \tag{1.4.10}$$

The cumulants are defined by

$$\Psi(\zeta) = \sum{}' \frac{(\mathrm{i}\zeta_1)^{r_1} \ldots (\mathrm{i}\zeta_n)^{r_n}}{r_1! \ldots r_n!} \langle z_1^{r_1} \ldots z_n^{r_n} \rangle_c \tag{1.4.11}$$

if the expansion is possible, where $\sum'$ means the omission of the term with $r_1 = r_2 = \ldots = r_n = 0$. Cumulants and moments are mutually related by (1.4.10), e.g.,

$$\begin{aligned} \langle z_1 z_2 \rangle &= \langle z_1 z_2 \rangle_c + \langle z_1 \rangle \langle z_2 \rangle, \\ \langle z_1 z_2 z_3 \rangle &= \langle z_1 z_2 z_3 \rangle_c + \langle z_1 \rangle \langle z_2 z_3 \rangle_c + \langle z_2 \rangle \langle z_1 z_3 \rangle_c \\ &\quad + \langle z_3 \rangle \langle z_1 z_2 \rangle_c + \langle z_1 \rangle \langle z_2 \rangle \langle z_3 \rangle. \end{aligned} \tag{1.4.12}$$

As evident in (1.2.27), all cumulants, $n \geqq 3$, are identically zero for a one-dimensional Gaussian distribution. This statement is generalized to an n-dimensional Gaussian distribution for which all cumulants vanish except the first and second, as is seen in (1.4.7) which contains terms only to second order in ζ. The coefficients of second-order terms are the variance matrix (1.4.4). Its element

$$\begin{aligned} \langle z(t_j)\, z(t_k) \rangle_c &= \langle z(t_j)\, z(t_k) \rangle - \langle z(t_j) \rangle \langle z(t_k) \rangle \\ &= \langle [z(t_j) - \langle z(t_j) \rangle] [z(t_k) - \langle z(t_k) \rangle] \rangle \end{aligned} \tag{1.4.13}$$

is the correlation function of $z(t)$. Therefore, (1.4.7) becomes

$$\Phi(\zeta_1, \ldots, \zeta_n) = \exp\left(\mathrm{i} \sum_{j=1}^{n} \zeta_j m(t_j) - \frac{1}{2} \sum_{j=1}^{n} \sum_{k=1}^{n} \phi(t_j, t_k)\, \zeta_j \zeta_k \right), \tag{1.4.14}$$

where

$$m(t_j) = \langle z(t_j) \rangle, \quad \phi(t_j, t_k) = \langle [z(t_j) - \langle z(t_j) \rangle] [z(t_k) - \langle z(t_k) \rangle] \rangle. \tag{1.4.15}$$

Thus, the process $z(t)$ is completely determined by the expectations and the correlation functions since the characteristic function is completely defined by these quantities.

Assume for simplicity that

$$m(t) = 0.$$

For an arbitrary set of n time points $(t_1, t_2, \ldots, t_n)$,

$$\langle z(t_1) \ldots z(t_n) \rangle = \begin{cases} 0 & \text{for odd } n, \\ \sum\limits_{\text{pairing pairs}} \prod \phi(t_j, t_k) & \text{for even } n, \end{cases} \tag{1.4.16}$$

holds. This is easily seen by comparing the power series expansion of (1.4.14) [setting $m(t_j) = 0$] in $\zeta_1, \zeta_2, \ldots \zeta_n$ and (1.4.9). In (1.4.16), the

summation means the following: we divide the set $t_1, t_2, \ldots, t_n$ (with an even n; any of these time points may coincide) into pairs and construct the product of $\phi(t_j, t_k)$ for this pairing and sum up such terms for all possible ways of pairing. For example, thus

$$\langle z(t_1)\, z(t_2)\, z(t_3)\, z(t_4)\rangle = \phi(t_1, t_2)\, \phi(t_3, t_4) + \phi(t_1, t_3)\, \phi(t_2, t_4) + \phi(t_1, t_4)\, \phi(t_2, t_3).$$

In the definition of the characteristic function (1.4.5), we set

$$z_j = z(t_j), \quad \zeta_j = \zeta(t_j)\, \Delta t_j \quad (j = 1, 2, \ldots, n)$$

and take the limit of $n \to \infty$ and $\Delta t_j \to 0$ for $t_0 < t_1 < t_2 \ldots < t_n < t$ to attain the limit

$$\sum_{j=1}^{n} \zeta_j z_j = \sum_{j=1}^{n} \zeta(t_j)\, z(t_j)\, \Delta t_j \to \int_{t_0}^{t} \zeta(t')\, z(t')\, dt'.$$

This defines

$$\Phi[\zeta(t)] = \left\langle \exp\left[i \int_{t_0}^{t} \zeta(t')\, z(t')\, dt' \right] \right\rangle. \tag{1.4.17}$$

This is the most general form of the characteristic function for the processes $z(t)$ and is called the *characteristic functional* because it contains an arbitrary function $\zeta(t)$.

In particular, if $z(t)$ is Gaussian, its characteristic functional is

$$\Phi[\zeta(t)] = \exp\left[i \int_{t_0}^{t} \zeta(t')\, m(t')\, dt' - \frac{1}{2} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\, \phi(t_1, t_2)\, \zeta(t_1)\, \zeta(t_2) \right] \tag{1.4.18}$$

corresponding to (1.4.14). In other words, the characteristic functional of a Gaussian process is completely defined in terms of the expectation $m(t)$ and the correlation function $\phi(t_1, t_2)$. If it is stationary, $m(t)$ is a constant so that it can be set equal to zero without losing generality. Furthermore, the correlation function $\phi(t_1, t_2)$ is a function of $t_1 - t_2$ only. Thus the characteristic function has the form

$$\Phi[\zeta(t)] = \exp\left[-\frac{1}{2} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\, \phi(t_1 - t_2)\, \zeta(t_1)\, \zeta(t_2) \right]. \tag{1.4.19}$$

If the characteristic functional $\Phi[\zeta(t)]$ is known, a suitably chosen $\zeta(t)$ gives a desired characteristic function. For example, setting

$$\zeta(t) = \sum_{j=1}^{n} \zeta_j\, \delta(t - t_j), \tag{1.4.20}$$

(1.4.17) is reduced to (1.4.5). By taking a functional derivative of the characteristic functional, we can calculate various sorts of expectations, provided that the derivative is analytic.

Gaussian processes are rather common in the real world and are regarded as standards. The reason for this seems to lie in the central limit theorem just in the same way as the Gaussian law of error. Assume that the random variable $z(t)$ is a sum of independent random variables

$$z(t) = \Delta z_1(t) + \Delta z_2(t) + \ldots + \Delta z_n(t)\,, \tag{1.4.21}$$

where each component is by itself a stochastic process. Then the characteristic functional of $z(t)$ has the form

$$\Phi[\zeta(t)] = \exp\left\{\sum_{j=1}^{n} \psi_j[\zeta(t)]\right\},$$

where $\psi_j[\zeta(t)]$ is the cumulant function for the characteristic functional of $\Delta z_j(t)$ and is assumed to be expanded as

$$\psi_j[\zeta(t)] = \mathrm{i}\int_{t_0}^{t} \zeta(t_1)\,\langle \Delta z_j(t_1)\rangle\, dt_1 - \frac{1}{2}\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\, \langle \Delta z_j(t_1)\,\Delta z_j(t_2)\rangle_c\, \zeta(t_1)\,\zeta(t_2) + \ldots .$$

If the sums of cumulants are all of the order of n like

$$\sum_{j=1}^{n} \langle \Delta z_j(t)\rangle = O(n), \quad \sum_{j=1}^{n} \langle \Delta z_j(t_1)\,\Delta z_j(t_2)\rangle_c = O(n), \ldots,$$

we change the variable to

$$y(t) = \frac{z(t)}{\sqrt{n}} \tag{1.4.22}$$

and obtain its characteristic functional as

$$\begin{aligned}\Phi[\eta(t)] &= \left\langle \exp\left[\mathrm{i}\int_{t_0}^{t} \eta(t')\,y(t')\,dt'\right]\right\rangle \\ &= \exp\left\{\sum_{j=1}^{n} \psi_j[n^{-1/2}\,\eta(t)]\right\} \\ &= \exp\left[\mathrm{i}\int_{t_0}^{t} \eta(t_1)\,\langle y(t_1)\rangle\, dt_1 - \frac{1}{2}\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2\, \langle y(t_1)\,y(t_2)\rangle\,\eta(t_1)\,\eta(t_2)\right. \\ &\qquad \left. + O(n^{-1/2}) + O(n^{-1}) + \ldots\right],\end{aligned} \tag{1.4.23}$$

which approaches the form (1.4.18) as $n \to \infty$. Here, $O(n^{-1/2})$ and $O(n^{-1})$ mean that the higher-order terms in the cumulant expansion become asymptotically small at these orders. Like the previous discussion of the central limit theorem in Sect. 1.2, the above argument is not quite satisfactory as a mathematical proof, but it indicates that we may generally expect a Gaussian property if a physical process results from a large number of independent random processes. More generally, even such a strict independence is not necessarily required for the Gaussian property to hold, provided that the interaction satisfies certain appropriate conditions not specified here. Furthermore, there are certainly important cases where the Gaussian properties no longer hold. Physical processes in the neighborhood of a phase change are interesting examples. The non-Gaussian nature of critical fluctuations has been a central topic in recent progress in statistical mechanics.

Gaussian distributions have remarkable stabilities. If the joint distribution of the random variables, $X_1, X_2, \ldots$ and X_n is Gaussian, any linear combination

$$Y_j = \sum_{k=1}^{n} C_{jk} X_k$$

of these variables also has a Gaussian distribution. Likewise, if $z(t)$ is a Gaussian process, any linear transformation of it

$$y(t) = \int_a^b C(t, t')\, z(t')\, dt', \tag{1.4.24}$$

is also Gaussian because it holds that

$$\langle y(t) \rangle = \int_a^b C(t, t') \langle z(t') \rangle\, dt', \tag{1.4.25}$$

$$\langle y(t_1)\, y(t_2) \rangle_c = \int_a^b dt_1' \int_a^b dt_2'\, C(t_1, t_1')\, C(t_2, t_2') \langle z(t_1')\, z(t_2') \rangle_c \tag{1.4.26}$$

and cumulants of $y(t)$ higher than the third vanish together with those of $z(t)$. In particular, the Fourier coefficients of $z(t)$ defined by (1.3.12) have a Gaussian distribution if $z(t)$ is a Gaussian process.

1.5 Brownian Motion Modeled by a Gaussian Process

The Langevin equation (1.3.35) depicts Brownian motion as driven by the random force $R(t)$. As an idealization of Brownian motion, $R(t)$ is assumed to satisfy the following conditions:

i) $R(t)$ is a Gaussian process,
ii) $R(t)$ has a white spectrum, namely, (1.3.40) holds.

The process $u(t)$ then represents the random change of velocity of a Brownian particle and is often called the *Ornstein-Uhlenbeck process* after the authors who treated the problem extensively for the first time [1.10].

In fact, these assumptions seem very plausible if Brownian particles are far larger than the molecules of the surrounding fluid. The force $R(t)$ acting on a Brownian particle results from a great many impacts from the fluid molecules, so that the Gaussian property is expected to hold by the central limit theorem. Secondly, the time constant of the motion of fluid molecules will be much shorter than the characteristic time of the Brownian particle if the mass of a Brownian particle is much larger than that of fluid molecules. (Rigorously speaking, this is not quite sufficient. As Sect. 1.6 shows, this idealization of $R(t)$ is legitimate only in the limit of very large mass density of the Brownian particle.) If that is the case, as an idealization the characteristic time of successive impacts from fluid molecules may be considered as infinitely short.

The Langevin equation

$$\dot{u}(t) = -\gamma u + \frac{R(t)}{m} \tag{1.5.1}$$

is easily integrated to give

$$u(t) = u(t_0)\, e^{-\gamma(t-t_0)} + \int_{t_0}^{t} e^{-\gamma(t-t')} \frac{R(t')}{m}\, dt'. \tag{1.5.2}$$

Obviously this has the form of (1.4.24) so that $u(t)$ must be a Gaussian process if $R(t)$ is Gaussian.

The power spectrum of $u(t)$ has already been obtained in (1.3.37) and its correlation function by (1.3.43) under assumption (ii). Therefore, the process $u(t)$ is completely defined. The transition probability $P(u_0, t_0 | u, t)$ to find the velocity u at time t when the velocity was u_0 at the initial time t_0 is derived from (1.5.2) as follows.

The characteristic functional of $u(t)$ is

$$\left\langle e^{i\xi u(t)} \right\rangle = \exp\left(i\xi u_0 \exp[-\gamma(t-t_0)] - \frac{\xi^2}{2} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \right.$$
$$\left. \times \exp[-\gamma(t-t_1) - \gamma(t-t_2)] \frac{\langle R(t_1) R(t_2) \rangle}{m^2} \right). \tag{1.5.3}$$

This is obtained by replacing $z(t)$ in (1.4.18) by $R(t)$ and $\zeta(t')$ by $\xi \exp[-\gamma(t-t')]/m$. However, this is rather obvious from the derivation of (1.4.18). If (1.3.41) is assumed further for $R(t)$, the second term in the

exponent in the integrand of (1.5.3) is calculated as

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \exp\left[-\gamma(t-t_1)-\gamma(t-t_2)\right] \frac{2\pi I_R}{m^2}\,\delta(t_1-t_2)$$

$$= \frac{2\pi I_R}{m^2} \int_{t_0}^{t} dt'\, \mathrm{e}^{-2\gamma(t-t')} = \frac{\pi I_R}{m^2} \frac{1-\mathrm{e}^{-2\gamma(t-t_0)}}{\gamma}.$$

With (1.3.46), then (1.5.3) becomes

$$\left\langle \mathrm{e}^{\mathrm{i}\xi u(t)} \right\rangle = \exp\left[\mathrm{i}\,\xi\, u_0\, \mathrm{e}^{-\gamma(t-t_0)} - \frac{kT}{2m}\left(1-\mathrm{e}^{-2\gamma(t-t_0)}\right)\xi^2 \right]. \tag{1.5.4}$$

This indicates that $u(t)$ has the Gaussian distribution

$$P(u_0, t_0 | u, t) = \left(\frac{m}{2\pi kT}\right)^{1/2} \frac{1}{\left(1-\mathrm{e}^{-2\gamma(t-t_0)}\right)^{1/2}} \exp\left(-\frac{m}{2kT} \frac{(u-u_0\,\mathrm{e}^{-\gamma(t-t_0)})^2}{1-\mathrm{e}^{-2\gamma(t-t_0)}} \right) \tag{1.5.5}$$

which is the transition probability for $(u_0, t_0) \to (u, t)$. The expectation of the velocity decays exponentially as

$$\langle u(t) \rangle = u_0\, \mathrm{e}^{-\gamma(t-t_0)}, \tag{1.5.6}$$

if the initial value was u_0 at time t_0. This is, of course, to be expected. The variance around the expectation grows in time as

$$\left\langle \left[u(t) - u_0\, \mathrm{e}^{-\gamma(t-t_0)}\right]^2 \right\rangle = \frac{kT}{m}\left(1-\mathrm{e}^{-2\gamma(t-t_0)}\right) \tag{1.5.7}$$

and approaches the Maxwellian value at $t \to \infty$. As shown below, the distribution (1.5.5) is the fundamental solution of the diffusion equation in the velocity space

$$\frac{\partial}{\partial t} P = \frac{\partial}{\partial u}\left(\gamma u + D_u \frac{\partial}{\partial u}\right) P, \quad D_u = \frac{\gamma kT}{m}. \tag{1.5.8}$$

The displacement $x(t)$ in the time interval $(0, t)$ is obtained by integrating (1.5.2) as

$$x(t) = \int_0^t u(t')\, dt = u_0 \frac{1-\mathrm{e}^{-\gamma t}}{\gamma} + \int_0^t dt_1 \int_0^{t_1} dt_2\, \mathrm{e}^{-\gamma(t_1-t_2)} \frac{R(t_2)}{m} \tag{1.5.9}$$

($t_0 = 0$, for brevity). Obviously this has a Gaussian distribution if u_0 and $R(t)$ are Gaussian. The second term is transformed into

$$\int_0^t dt_2 \int_{t_2}^t dt_1\, \mathrm{e}^{-\gamma(t_1-t_2)} \frac{R(t_2)}{m} = \int_0^t dt' \frac{1-\mathrm{e}^{-\gamma(t-t')}}{\gamma} \frac{R(t')}{m}.$$

The characteristic functional of $x(t)$,

$$\langle e^{i\xi x(t)}\rangle = \left\langle \exp\left(i\,\xi\, u_0 \frac{1-e^{-\gamma t}}{\gamma}\right)\right\rangle \left\langle \exp\left(i\,\xi \int_0^t dt' \frac{1-e^{-\gamma(t-t')}}{\gamma}\frac{R(t')}{m}\right)\right\rangle,$$

is calculated similarly as for (1.5.3). Assuming the distribution of u_0 to be Maxwellian and using (1.3.44), then

$$\langle e^{i\xi x(t)}\rangle = \exp\left[-\xi^2 \frac{\langle u^2\rangle}{\gamma}\left(t - \frac{1-e^{-\gamma t}}{\gamma}\right)\right]. \tag{1.5.10}$$

Further we set

$$\langle u^2\rangle = \frac{kT}{m} \quad \text{and} \quad D = \frac{\langle u^2\rangle}{\gamma} = \frac{kT}{m\gamma} \tag{1.5.11}$$

to rewrite (1.5.10) as

$$\langle e^{ix(t)}\rangle = \exp\left[-\frac{\xi^2}{2} 2D\left(t - \frac{1-e^{-\gamma t}}{\gamma}\right)\right] \tag{1.5.12}$$

This corresponds to the transition probability $P(0, 0|x, t)$ of a Brownian particle to arrive around x at time t when it was certainly at $x = 0$ at the initial time. The probability is Gaussian and is given by

$$P(0,0|x,t) = \left[4\pi D\left(t - \frac{1-e^{-\gamma t}}{\gamma}\right)\right]^{-1/2} \exp\left\{-x^2\left[4D\left(t - \frac{1-e^{-\gamma t}}{\gamma}\right)\right]^{-1}\right\}. \tag{1.5.13}$$

The mean square average (variance) of the displacement in the time interval $(0, t)$ is

$$\langle x(t)^2\rangle = 2D\left(t - \frac{1-e^{-\gamma t}}{\gamma}\right). \tag{1.5.14}$$

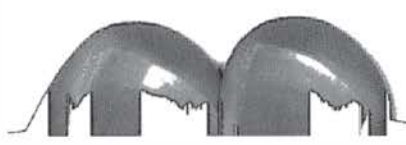
For a short time $t \ll 1/\gamma$ the distribution of $x(t)$ is Gaussian with the variance

$$\langle x^2\rangle \approx \langle u^2\rangle t^2 \tag{1.5.15}$$

as is seen from (1.5.11, 14). In such a short time, the Brownian particle still keeps its initial velocity. The result (1.5.15) is simply a reflection of the initial Maxwellian distribution of the velocity.

For longer times $t \gg 1/\gamma$, the Brownian particle repeatedly zig-zags and loses the memory of its initial velocity. It is natural to expect that displacements after such a long time become a diffusion process as discussed in Sect. 1.1. In fact, (1.5.13) is then approximated by

$$P(0,0|x,t) \approx \frac{1}{(4\pi Dt)^{1/2}} \exp\left(-\frac{x^2}{4Dt}\right) \quad (t \gg \gamma^{-1}) \tag{1.5.16}$$



in accordance with (1.1.19). Furthermore, (1.5.11) is identical with (1.1.14). Namely, the Einstein relation is obtained here again. The above reasoning is not quite the same as that discussed in Sect. 1.1, but the essential point lies in the assertion that Brownian motion in a medium in thermal equilibrium also tends to attain thermal equilibrium.

1.6 The Fluctuation-Dissipation Theorem

In the Langevin equation (1.3.4), the force on a Brownian particle was divided into the frictional force $-m\gamma u$ and the random force $R(t)$, between which a relationship like (1.3.46) exists, indicating that the power intensity of $R(t)$ is proportional to the friction coefficient and the thermal energy kT. We have also seen that the Einstein relation relates the diffusion constant to the friction coefficient. Both express that such a mechanism of energy dissipation is closely related to fluctuations in thermal equilibrium and they are simple examples of a more general principle called the fluctuation-dissipation theorem [1.4]. Chapter 4 deals with a quantum-statistical derivation of this theorem. Here we consider this from the viewpoint of the Brownian motion theory.

As already mentioned, Brownian motion is not limited to Brownian particles. It is, generally speaking, a fluctuating motion of a mode in a macroscopic dynamical system with a very large number of particles or a large number of degrees of freedom. It is particularly simple for a particle much heavier than the molecules in a medium or for a mirror in a gas, described well by the simple Langevin equation discussed in Sect. 1.3. However, various modifications are required for the Langevin equation to be applied to more general sorts of Brownian motion. One modification is to abandon the assumption of a white spectrum for the random force $R(t)$. This means, as seen in the following, that retarded friction is accounted for. This is very necessary for applying the theory to more realistic problems for which idealizations are not legitimate.

In the Langevin equation (1.3.4) the friction is assumed to be determined by the instantaneous velocity of the particle. However, in general, friction will be retarded so that the Langevin equation should be generalized to

$$\frac{d}{dt}u(t) = -\int_{-\infty}^{t} \gamma(t-t')\,u(t')\,dt' + \frac{1}{m}R(t) + \frac{1}{m}K(t), \tag{1.6.1}$$

where $\gamma(t)$ expresses the friction retardation. Equation (1.6.1) is called a generalized Langevin equation. On the right-hand side, $R(t)$ is the random force and $K(t)$ is an external force. The random force is zero on average,

satisfying the condition

$$\langle R(t) \rangle = 0 . \tag{1.6.2}$$

Suppose now that the external force is periodic as

$$K(t) = K_0 \cos \omega t = \mathrm{Re}\{K_0 \, \mathrm{e}^{\mathrm{i}\omega t}\} .$$

Then the average velocity induced by this force is

$$\langle u(t) \rangle = \mathrm{Re}\{\mu(\omega) \, K_0 \, \mathrm{e}^{\mathrm{i}\omega t}\} , \tag{1.6.3}$$

where $\mu(\omega)$ is the complex mobility for the frequency ω and is given by

$$\mu(\omega) = \frac{1}{m} \frac{1}{\mathrm{i}\omega + \gamma[\omega]} , \tag{1.6.4}$$

where

$$\gamma[\omega] = \int_0^\infty \gamma(t) \, \mathrm{e}^{-\mathrm{i}\omega t} \, dt \tag{1.6.5}$$

is the Fourier-Laplace transform[3] of the retardation function of friction. This is obtained from the averaged equation of (1.6.1)

$$\frac{d}{dt} \langle u(t) \rangle = - \int_{-\infty}^{t} \gamma(t - t') \langle u(t') \rangle \, dt' + \mathrm{Re}\left\{ \frac{1}{m} K_0 \, \mathrm{e}^{\mathrm{i}\omega t} \right\} .$$

If the particles are charged with e and the particle density is n, the current induced by an electric field E is

$$j(t) = en \langle u(t) \rangle = \mathrm{Re}\{e^2 n \mu(\omega) E_0 \, \mathrm{e}^{\mathrm{i}\omega t}\}$$

so that the complex conductivity $\sigma(\omega)$ is

$$\sigma(\omega) = e^2 n \mu(\omega) = \frac{e^2 n}{m} \frac{1}{\mathrm{i}\omega + \gamma[\omega]} . \tag{1.6.6}$$

In fact, if we write conductivity, or more generally a complex admittance in this form, $\gamma[\omega]$ is usually not constant but depends on the frequency ω. If we treat such a system from the viewpoint of Brownian motion theory, the retardation function $\gamma(t)$ must be introduced, given as the inverse of (1.6.5).

---

3 A Fourier-Laplace transform is defined for the integration range $(0, \infty)$ in contrast to an ordinary Fourier transform with $(-\infty, \infty)$. Conventional Laplace transforms use a complex parameters s instead of $\mathrm{i}\omega$

The generalized Langevin equation (1.6.1) is linear so that it can be treated by harmonic analysis, Sect. 1.3. The Brownian motion follows the equation

$$\frac{d}{dt}u(t) = -\int_{-\infty}^{t} \gamma(t-t')\,u(t')\,dt' + \frac{1}{m}R(t), \tag{1.6.7}$$

if the external force K is zero and the motion is driven only by the random force $R(t)$. Now we Fourier-analyze $R(t)$ and $u(t)$ as

$$R(t) = \int_{-\infty}^{\infty} R(\omega)\,e^{i\omega t}\,d\omega, \qquad u(t) = \int_{-\infty}^{\infty} u(\omega)\,e^{i\omega t}\,d\omega$$

to obtain

$$u(\omega) = \frac{1}{i\omega + \gamma[\omega]}\,\frac{R(\omega)}{m}$$

from (1.6.7). If $R(t)$ is stationary, $u(t)$ becomes stationary for large enough t. The power spectra of the two processes are related to each other by

$$I_u(\omega) = \frac{1}{m^2}\,\frac{I_R(\omega)}{|i\omega + \gamma[\omega]|^2}, \tag{1.6.8}$$

as seen from (1.3.20).

When the spectrum $I_R(\omega)$ for the random force $R(t)$ is given, (1.6.8) yields $I_u(\omega)$, from which the correlation function $\langle u(0)\,u(t)\rangle$ is obtained by the Wiener-Khintchine theorem. If it should represent the velocity distribution in thermal equilibrium, the spectrum $I_R(\omega)$ is required to fulfill a certain condition. The condition is a generalization of (1.3.46) and is given by

$$I_R(\omega) = \frac{mkT}{\pi}\,\mathrm{Re}\,\{\gamma[\omega]\} \quad \text{or} \tag{1.6.9}$$

$$\langle R(\omega)\,R^*(\omega')\rangle = \frac{mkT}{\pi}\,\mathrm{Re}\,\{\gamma[\omega]\,\delta(\omega-\omega')\}\,. \tag{1.6.10}$$

This means

$$\langle R(t_1)\,R(t_2)\rangle = mkT\gamma(t_1 - t_2) \tag{1.6.11}$$

for the correlation function of $R(t)$ as is seen in the following way. The function $\gamma(t)$ in (1.6.5) is defined only for $t > 0$ but is extended to $t < 0$ by assuming $\gamma(t) = \gamma(-t)$. Then

$$\mathrm{Re}\,\{\gamma[\omega]\} = \frac{1}{2}\int_{-\infty}^{\infty} \gamma(t)\,e^{-i\omega t}\,dt,$$

from (1.6.5). Equation (1.6.11) follows from (1.6.9, 1.3.23). In (1.6.9) it must be assumed that

$$\mathrm{Re}\,\{\gamma[\omega]\} \geqq 0 \quad (\text{for real } \omega), \tag{1.6.12}$$

since the power spectrum should never become negative.

If $\gamma^*[\omega]$ denotes the complex conjugate of $\gamma[\omega]$, from (1.6.8, 9)

$$\langle u(t_0)\, u(t_0+t)\rangle = \frac{kT}{2\pi m} \int_{-\infty}^{\infty} \left(\frac{1}{\mathrm{i}\omega + \gamma[\omega]} + \frac{1}{-\mathrm{i}\omega + \gamma^*[\omega]} \right) \mathrm{e}^{\mathrm{i}\omega t}\, d\omega\,. \tag{1.6.13}$$

We now show that the contribution from the second term of the integrand vanishes for $t > 0$. To see this, observe that the function $\gamma(\omega)$ defined by (1.6.5) is analytic in the lower half-plane $\mathrm{Im}\,\{\omega\} < 0$. For such a function the dispersion relation discussed in Sect. 3.6 holds. This gives

$$\gamma[\omega] = \frac{\mathrm{i}}{\pi} \int_{-\infty}^{\infty} \frac{\gamma'(\nu)}{\nu - \omega}\, d\nu\,.$$

Setting $\omega = \omega' - \mathrm{i}\,\omega''$, then

$$\mathrm{Re}\,\{\gamma[\omega]\} = \frac{1}{\pi} \int_{-\infty}^{\infty} d\nu\, \gamma'(\nu)\, \frac{\omega''}{(\nu-\omega')^2 + \omega''^2}\,.$$

It follows from (1.6.12) that

$$\mathrm{Re}\,\{\gamma[\omega]\} > 0 \quad (\mathrm{Im}\,\{\omega\} < 0)$$

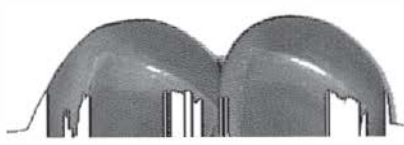
and

$$\mathrm{Re}\,\{\mathrm{i}\omega + \gamma[\omega]\} > 0 \quad (\mathrm{Im}\,\{\omega\} < 0).$$

Therefore, on the right-hand side of (1.6.13) the first term in the bracket is analytic for $\mathrm{Im}\,\{\omega\} < 0$. Correspondingly, the second term is analytic for $\mathrm{Im}\,\{\omega\} > 0$ and henceforth the integral containing this vanishes in (1.6.13), because the integral can be supplemented by a large semicircle on the half-plane $\mathrm{Im}\,\{\omega\} > 0$ on which $\exp(\mathrm{i}\omega t)$ tends to zero for $t > 0$ and the integrand is analytic inside the closed path of integration. Thus (1.6.13) is simplified to

$$\langle u(t_0)\, u(t_0+t)\rangle = \frac{kT}{2\pi m} \int_{-\infty-\mathrm{i}\epsilon}^{\infty-\mathrm{i}\epsilon} \frac{\mathrm{e}^{\mathrm{i}\omega t}}{\mathrm{i}\omega + \gamma[\omega]}\, d\omega. \tag{1.6.14}$$

Here the integration path lies just below the real axis of ω. If there is no branch cut for the function $\gamma[\omega]$, the path can be made a closed contour by supplementing with a large semicircle as shown in Fig. 1.5. In the limit $t \to 0+$, the sum of the residues of $(\mathrm{i}\omega + \gamma[\omega])^{-1}$ is equal to the residue



around the infinity $\omega = \infty$. If the condition

$$\lim_{|\omega|\to\infty} \gamma[\omega] = \text{finite}$$

is assumed, namely, if the resistance remains finite for $\omega \to \infty$, the residue is simply equal to one. Therefore the equipartition law results:

$$\lim_{t\to 0+} \langle u(t_0)\, u(t_0+t)\rangle \equiv \langle u^2\rangle = \frac{kT}{m}.$$

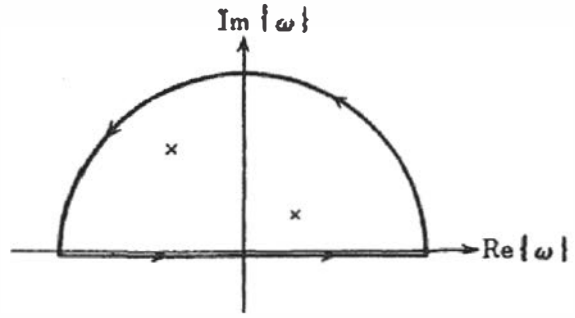


Fig. 1.5. Integration path for (1.6.14). The crosses indicate poles of the integrand

It is interesting to note that the Einstein-Ornstein-Uhlenbeck theory of Brownian motion as formulated by (1.3.1–4) has to be modified because of retardation in viscous resistance due to the hydrodynamic backflow effect. By a hydrodynamic calculation [1.11], it is shown that (1.5.1) is replaced by

$$m^* \frac{du}{dt} + \beta u + \alpha \int_{-\infty}^{t} (t-t')^{-1/2}\, \dot{u}(t')\, dt' = R(t), \tag{1.6.15}$$

where

$$m^* = m + \frac{2}{3}\pi\varrho a^3 = \frac{4\pi}{3}\left(\varrho_0 + \frac{\varrho}{2}\right) a^3,$$

$$\alpha = 6\pi\varrho a^2 (\nu/\pi)^{1/2},$$

$$\beta = 6\pi\nu\varrho a.$$

Here ϱ is the density of the fluid surrounding the particle, ϱ_0 is the average density of the matter composing the particle and ν is the kinetic viscosity equal to η/ϱ. The effective mass m^* contains additional inertia due to the dragging motion of the fluid. Retardation of viscous resistance is caused by the backflow of fluid. Then

$$\gamma[\omega] = \{\beta + (\pi i\omega)^{1/2}\alpha\}/m^*.$$

This function has a branch point at $\omega = 0$ and does not satisfy the condition (1.6.12). Accordingly, the previous argument does not apply. However, the integral (1.6.14) is easily calculated by transforming the path to the contour

$(i\infty - \varepsilon, 0, i\infty + \varepsilon, \varepsilon > 0)$, giving

$$\langle u(t_0)\, u(t_0 + t)\rangle = \frac{kT}{m^*}\phi(t),$$

where

$$\phi(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\exp(-z^2\tau)\,\sigma\, z^2\, dz}{(z^2-1)^2 + \sigma^2 z^2}$$

$$= \frac{1}{\pi}\sigma\tau^{-3/2}\int_{-\infty}^{\infty} \frac{\exp(-\zeta^2)\,\zeta^2\, d\zeta}{(\zeta^2/\tau - 1)^2 + \sigma^2\zeta^2/\tau}$$

with

$$\sigma = \left(\frac{9}{2}\frac{\varrho}{\varrho_0 + \varrho/2}\right)^{1/2}, \quad \tau = \beta t/m^*.$$

In the ideal limit of an extremely heavy particle ($\sigma \to 0$), $\phi(t)$ is reduced to $\exp(-\tau)$ and the retardation effect disappears. For a finite value of σ, the retardation effect shows up in a slow decay of the correlation function

$$\phi(t) \sim \frac{\sigma}{2\sqrt{\pi}}\tau^{-3/2}. \tag{1.6.16}$$

The presence of such a long-time tail was first observed by *Adler* and *Wainwright* in a computer simulation of velocity correlation functions [1.12].

Equation (1.6.14) means

$$\mu(\omega) = \frac{1}{m}\frac{1}{i\omega + \gamma[\omega]} = \frac{1}{kT}\int_0^{\infty}\langle u(t_0)\, u(t_0 + t)\rangle\, e^{-i\omega t}\, dt \tag{1.6.17}$$

because it is nothing but the inverse transformation of the above expression. Equations (1.6.5, 11) can now be written as

$$m\gamma[\omega] = \frac{1}{kT}\int_0^{\infty}\langle R(t_0)\, R(t_0 + t)\rangle\, e^{-i\omega t}\, dt \tag{1.6.18}$$

in analogy to (1.6.17), here yielding two fundamental expressions of the fluctuation-dissipation theorem (F-D theorem).

The first expression gives the complex mobility (complex admittance in general) in terms of the Fourier-Laplace transform of the correlation function of velocity (flow) and is a generalization of the Einstein relation (1.5.11). The second gives the complex resistance (complex impedance in general) in terms of the Fourier-Laplace transform of the correlation function of the random force. Formula (1.6.10), which is equivalent to this, was first obtained by *Nyquist* as the power spectrum of noise voltage caused by thermal fluctuations in a resistance [1.13]. These two expressions imply that the response of a system to an external disturbance is related to thermal

fluctuations spontaneously produced in the system in the absence of external forces. The relation of the dissipative part of the response to fluctuations was first recognized by *Nyquist* (Nyquist theorem) and obtained the name fluctuation-dissipation theorem. However, it is important to recognize that the theorem is true both for the dissipative as well as the dispersive (nondissipative) parts.

To distinguish between the two expressions (1.6.17, 18), we call the first the *F-D theorem of the first kind* and the second the *F-D theorem of the second kind*. As shown in Chap. 4, the first theorem can be derived from the linear response theory in a general way [1.14]. The correlation function on the right-hand side of the equation can be analyzed from microscopic theories by statistical mechanics. On the other hand, the random force appearing in the F-D theorem of the second kind is not simple, because the separation of the force into frictional and random forces is itself a complex problem of statistical mechanics. In this sense, the F-D theorem of the first kind should be regarded as basic and the second as a corollary to the first. In the context of this chapter, the two theorems are related to each other through the Langevin equation (1.6.1).

The mean square average of the displacement of a Brownian particle in a time interval $(0, t)$ is given by

$$\langle x(t)^2 \rangle = \int_0^t dt_1 \int_0^t dt_2 \, \langle u(t_1)\, u(t_2) \rangle .$$

This is transformed into

$$\lim_{t \to \infty} \frac{\langle x(t)^2 \rangle}{2t} = \int_0^\infty \langle u(t_0)\, u(t_0 + t') \rangle \, dt' \tag{1.6.19}$$

as for (1.3.24), since the correlation function $\langle u(t_1)\, u(t_2) \rangle$ is dependent on $t_1 - t_2$ only as a stationary process. Through (1.2.33, 6.17) we obtain

$$D = \mu(0)\, kT .$$

This shows that the Einstein relation (1.5.11) is a special case of the F-D theorem of the first kind.

In (1.6.1, 7) the retarded friction is expressed as an integral from the infinite past to the present. It is possible to modify this to

$$\frac{d}{dt} u(t) = - \int_{t_0}^t \gamma(t - t')\, u(t')\, dt' + \frac{1}{m} \bar{R}(t), \quad (t > t_0) . \tag{1.6.20}$$

For stationary Brownian motion, the initial time t_0 in this equation can be chosen arbitrarily. In that sense the correlation function obtained from this equation does not depend on the choice of t_0.

The random force $\bar{R}(t)$ is not the same as $R(t)$ in (1.6.7) and is

$$\bar{R}(t) = R(t) - \int_{-\infty}^{t_0} \gamma(t-t')\, m u(t')\, dt'. \tag{1.6.21}$$

For (1.6.20) to represent the same Brownian motion as represented by (1.6.7), $\bar{R}(t)$ must fulfill the conditions

$$\langle u(t_0)\, \bar{R}(t)\rangle = 0 \quad \text{for} \quad t > t_0 \quad \text{and} \tag{1.6.22}$$

$$\langle \bar{R}(t_0)\, \bar{R}(t_0+t)\rangle = \langle R(t_1)\, R(t_1+t)\rangle = m k T \gamma(t). \tag{1.6.23}$$

This can be seen as follows. From (1.6.17) follow

$$\int_0^\infty \langle \dot{u}(t_0)\, u(t_0+t)\rangle\, e^{-i\omega t}\, dt = -\int_0^\infty \langle u(t_0)\, \dot{u}(t_0+t)\rangle\, e^{-i\omega t}\, dt = \frac{\langle u^2\rangle\, \gamma[\omega]}{i\omega + \gamma[\omega]}, \tag{1.6.24}$$

$$\int_0^\infty \langle \dot{u}(t_0)\, \dot{u}(t_0+t)\rangle\, e^{-i\omega t}\, dt = \frac{\langle u^2\rangle\, i\omega\, \gamma[\omega]}{i\omega + \gamma[\omega]}, \tag{1.6.25}$$

and from (1.6.20)

$$\langle u(t_0)\, \dot{u}(t_0+t)\rangle = -\int_0^t \gamma(t-t')\langle u(t_0)\, u(t_0+t')\rangle\, dt' + \frac{\langle u(t_0)\, \bar{R}(t_0+t)\rangle}{m}.$$

Condition (1.6.22) is necessary for the Laplace transform of the above equation to be satisfied by (1.6.17, 24). Furthermore,

$$\langle \bar{R}(t_0)\, \bar{R}(t_0+t)\rangle = m^2 \left\langle \dot{u}(t_0) \left[\dot{u}(t_0+t) + \int_0^t \gamma(t-t')\, u(t_0+t')\, dt' \right] \right\rangle,$$

which yields

$$\int_0^\infty \langle \bar{R}(t_0)\bar{R}(t_0+t)\rangle\, e^{-i\omega t}\, dt = m^2 \langle u^2\rangle\, \gamma[\omega]$$

by (1.6.17, 24, 25). Therefore (1.6.23) should hold. We can show by direct calculation [1.15] using (1.6.21) that for an arbitrary $t_1 > t_0$

$$\langle \bar{R}(t_0)\, \bar{R}(t_0+t)\rangle = \langle \bar{R}(t_1)\, \bar{R}(t_1+t)\rangle,$$

although $R(t)$ is by itself nonstationary since it depends on the arbitrarily chosen initial time t_0. Despite such a somewhat unnatural artifice, the form of the Langevin equation (1.6.20) is sometimes more convenient than that of (1.6.7). As shown in Sect. 2.9, there is a way of transforming the equation of motion into this form which gives a basis to formulate a statistical theory of Brownian motion.

Condition (1.6.22) means that the random force $\bar{R}(t)$ is uncorrelated (in a weak sense) with $u(t_0)$. This does not, however, mean causality, because $\bar{R}(t)$ is generally correlated with $u(t)$ in the future as well as in the past. From (1.6.21) we can show that

$$\langle u(t)\bar{R}(t+\tau)\rangle = m\int_0^{t-t_0} \gamma(\tau+t')\langle u(0)\,u(t')\rangle\,dt' \tag{1.6.26}$$

and from (1.6.7)

$$\langle u(t)\,R(t+\tau)\rangle = m\int_0^{\infty} \gamma(\tau+t')\langle u(0)\,u(t')\rangle\,dt'\,. \tag{1.6.27}$$

Both expressions tend to zero with increasing τ to ∞. But they are not equal to zero except when there is no retardation, where $\gamma(t)$ is a delta function and both correlation functions are zero for $\tau > 0$ and equal to $2m\gamma\langle u^2\rangle\exp(\gamma\tau)$ for $\tau < 0$. In general cases of retarded friction, the random force must be correlated with the velocity in the past, which is not surprising.