nger Series in Solid-State Sciences

s: M. Cardona P. Fulde K. von Klitzing H.-J. Queisser

ging Editor: H. K.V. Lotsch

Volumes 1–89 are listed at the end of the book

R. Kubo M. Toda N. Hashitsume

Statistical Physics II

Nonequilibrium Statistical Mechanics

Second Edition
With 28 Figures

Springer

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:

$$\Phi(\xi) = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} \exp\left(-\frac{(x-m)^2}{2\sigma^2} + ix\xi \right) dx$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma} \int_{-\infty}^{\infty} \exp\left[im\xi - \frac{\sigma^2}{2}\xi^2 - \frac{1}{2\sigma^2}(x - m - i\sigma^2\xi)^2 \right] dx .$$

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.

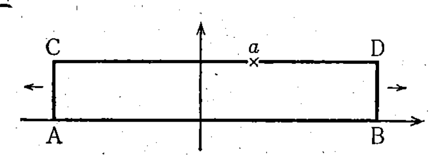

Fig. 1.3. The integration path to prove (1.2.21)

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

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

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

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

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

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

are all possible, then

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

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

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

is

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

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

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

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

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

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

setting

$$m = (m_1, m_2, \ldots, m_n), \quad z - m = y = u + \mathrm{i}v .$$

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

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

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,

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

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

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

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

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

if the expansion is possible, where $\sum{}'$ means the omission of the term with $r_1 = r_2 = \dots = r_n = 0$. 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

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

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 $(t_1, t_2, \dots, t_n)$,

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

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

summation means the following: we divide the set $t_1, t_2, \dots, t_n$ (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

$$\begin{aligned} &\langle z(t_1)\, z(t_2)\, z(t_3)\, z(t_4) \rangle \\ &= \phi(t_1, t_2)\, \phi(t_3, t_4) + \phi(t_1, t_3)\, \phi(t_2, t_4) + \phi(t_1, t_4)\, \phi(t_2, t_3). \end{aligned}$$

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

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

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

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

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

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

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

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