Chapter 5. Fluctuations

This chapter discusses fluctuations of statistical variables, mostly at thermodynamic equilibrium. In particular, I will describe the intimate connection between fluctuations and dissipation (damping) in a dynamic system weakly coupled to a multi-particle environment, which culminates in the Einstein relation between the diffusion coefficient and mobility, the Nyquist formula, and their quantum-mechanical generalization - the fluctuation-dissipation theorem. An alternative approach to the same problem, based on the Smoluchowski and Fokker-Planck equations, is also discussed in brief.

5.1. Characterization of fluctuations

In the beginning of Chapter 2, we have discussed the notion of averaging, \( \langle f \rangle \), of a variable \( f \) over a statistical ensemble – see Eqs. (2.7) and (2.10). Now, the variable’s fluctuation may be defined simply as its deviation from the average:

\[
\tilde{f} \equiv f - \langle f \rangle;
\]  
(5.1)

this deviation is, evidently, also a random variable. The most important property of any fluctuation is that its average (over the same statistical ensemble) equals zero:

\[
\langle \tilde{f} \rangle = \langle f - \langle f \rangle \rangle = \langle f \rangle - \langle \langle f \rangle \rangle = \langle f \rangle - \langle f \rangle = 0.
\]  
(5.2)

As a result, such average cannot characterize fluctuations’ intensity, whose simplest characteristic is the variance (also called “dispersion”):

\[
\langle \tilde{f}^2 \rangle = \langle (f - \langle f \rangle)^2 \rangle.
\]  
(5.3)

The following simple property of the variance is frequently convenient for its calculation:

\[
\langle \tilde{f}^2 \rangle = \langle (f - \langle f \rangle)^2 \rangle = \langle f^2 - 2f \langle f \rangle + \langle f \rangle^2 \rangle = \langle f^2 \rangle - 2\langle f \rangle^2 + \langle f \rangle^2,
\]  
(5.4a)

so that, finally:

\[
\langle \tilde{f}^2 \rangle = \langle f^2 \rangle - \langle f \rangle^2.
\]  
(5.4b)

As the simplest example of its application, consider a variable which can take only two values, \( \pm 1 \), with equal probabilities \( W_j = \frac{1}{2} \). For such a variable,

\[
\langle f \rangle = \sum_j W_j f_j = \frac{1}{2}(+1) + \frac{1}{2}(-1) = 0, \quad \text{but} \quad \langle f^2 \rangle = \sum_j W_j f_j^2 = \frac{1}{2}(+1)^2 + \frac{1}{2}(-1)^2 = 1,
\]  
(5.5)

so that \( \langle \tilde{f}^2 \rangle = \langle f^2 \rangle - \langle f \rangle^2 = 1 \).

The square root of variance,

\[
\tilde{f} \equiv \langle \tilde{f}^2 \rangle^{1/2}.
\]  
(5.6)
is called the \textit{root-mean-square (r.m.s.) fluctuation}. An advantage of this measure is that it has the same dimensionality as the variable itself, so that ratio \( \delta f / \langle f \rangle \) is dimensionless, and may be used to characterize the \textit{relative intensity} of fluctuations. In particular, as has been mentioned in Chapter 1, all results of thermodynamics are valid only if the fluctuations of thermodynamic variables (internal energy \( E \), entropy \( S \), etc.) are relatively small.\(^1\) Let us make the simplest estimate of the relative intensity of fluctuations by considering a system of \( N \) independent, similar particles, and an extensive variable

\[
\mathcal{F} = \sum_{j=1}^{N} f_j. \tag{5.7}
\]

where \( f_j \) depends on the state of just one \( j^{\text{th}} \) particle. The statistical average of \( \mathcal{F} \) is evidently

\[
\langle \mathcal{F} \rangle = \sum_{j=1}^{N} \langle f_j \rangle = N \langle f \rangle, \tag{5.8}
\]

while the variance is

\[
\langle \mathcal{F}^2 \rangle = \left( \sum_{j=1}^{N} \sum_{j'=1}^{N} f_j f_{j'} \right) = \sum_{j=1}^{N} \sum_{j'=1}^{N} \langle f_j f_{j'} \rangle = \sum_{j, j'=1}^{N} \langle f_j f_{j'} \rangle. \tag{5.9}
\]

Now we may use the fact that for two independent variables

\[
\langle f_j f_{j'} \rangle = 0; \tag{5.10}
\]

actually, this equation may be considered as the mathematical definition of the independence. Hence, in the sum (9), only the terms with \( j' = j \) survive, and

\[
\langle \mathcal{F}^2 \rangle = \sum_{j, j'=1}^{N} \langle f_j^2 \rangle \delta_{j, j'} = N \langle f^2 \rangle. \tag{5.11}
\]

Comparing Eqs. (8) and (11), we see that the relative intensity of fluctuations of variable \( \mathcal{F} \),

\[
\frac{\delta \mathcal{F}}{\langle \mathcal{F} \rangle} = \frac{1}{N^{1/2}} \frac{\delta f}{\langle f \rangle}, \tag{5.12}
\]

tends to zero as the system size grows (\( N \to \infty \)). It is this fact that justifies the thermodynamic approach to typical physical systems, with the number \( N \) of particles of the order of the Avogadro number \( N_A \approx 10^{24} \). Nevertheless, in many situations even small fluctuations of thermodynamic variables are important, and in this chapter we will calculate their basic properties, starting from the variance.

It will be pleasant for the reader to notice that for some simple (but important) cases, such calculation has already been done in our course. For example, for any generalized coordinate \( q_j \) and generalized momentum \( p_j \) that give quadratic contributions to system’s Hamiltonian (2.46), we have derived the equipartition theorem (2.48), valid in the classical limit. Since the average values of these

\(^1\) Let me remind the reader that up to this point, the averaging signs \( \langle \ldots \rangle \) were dropped in most formulas, for the sake of notation simplicity. In this chapter I have to restore these signs to avoid confusion. The only exception will be temperature whose average, following (bad :-/) tradition, will be still call \( T \) everywhere besides the last part of Sec. 3 where temperature fluctuations are discussed explicitly.
variables, in the thermodynamic equilibrium, equal zero, Eq. (6) immediately yields their r.m.s. fluctuations:

$$\delta p_j = (mT)^{1/2}, \quad \delta q_j = \left( \frac{T}{m\omega^2} \right)^{1/2}. \quad (5.13)$$

The generalization of these classical relations to the quantum-mechanical case \((T \sim \hbar \omega)\) for a 1D harmonic oscillator is provided by Eqs. (2.78) and (2.81):

$$\delta p_j = \left[ \frac{\hbar m \omega}{2 \coth \frac{\hbar \omega}{2T}} \right]^{1/2}, \quad \delta q_j = \left[ \frac{\hbar}{2m \omega} \coth \frac{\hbar \omega}{2T} \right]^{1/2}. \quad (5.14)$$

However, the intensity of fluctuations in other systems requires special calculations. Moreover, only a few cases allow for general, model-independent results. Let us review some of them.

5.2. Energy and the number of particles

First of all, note that fluctuations of macroscopic variables depend on particular conditions.\(^2\) For example, in a mechanically- and thermally-insulated system, e.g., a member of a microcanonical ensemble, there are no fluctuations of internal energy: \(\delta E = 0\).

However, if a system is in a thermal contact with environment, for example is a member of a canonical ensemble (Fig. 2.6), the Gibbs distribution (2.58)-(2.59) is valid. We already know that application of this distribution to energy itself,

$$\langle E \rangle = \sum_m W_mE_m, \quad W_m = \frac{1}{Z} \exp \left\{ -\frac{E_m}{T} \right\}, \quad Z = \sum_m \exp \left\{ -\frac{E_m}{T} \right\}, \quad (5.15)$$
yields Eq. (2.61b), which may be rewritten in the form

$$\langle E \rangle = \frac{1}{Z} \frac{\partial Z}{\partial (-\beta)}, \quad \text{with } \beta \equiv \frac{1}{T}, \quad (5.16)$$
more convenient for our current purposes. Now let us carry out a similar calculation for variable \(E^2\):

$$\langle E^2 \rangle = \sum_m W_mE_m^2 = \frac{1}{Z} \sum_m E_m^2 \exp \left\{ -\beta E_m \right\}. \quad (5.17)$$

It is straightforward to check, by double differentiation, that this expression may be rewritten as

$$\langle E^2 \rangle = \frac{1}{Z} \frac{\partial^2 Z}{\partial (-\beta)^2} \sum_m \exp \left\{ -\beta E_m \right\} = \frac{1}{Z} \frac{\partial^2 Z}{\partial (-\beta)^2}. \quad (5.18)$$

Now it is straightforward to use Eq. (4) to calculate the energy fluctuation variance:

$$\langle E^2 \rangle - \langle E \rangle^2 = \frac{1}{Z} \frac{\partial^2 Z}{\partial (-\beta)^2} - \frac{1}{Z^2} \left( \frac{\partial Z}{\partial (-\beta)} \right)^2 = \frac{\partial}{\partial (-\beta)} \left( \frac{1}{Z} \frac{\partial Z}{\partial (-\beta)} \right) = \frac{\partial \langle E \rangle}{\partial (-\beta)}. \quad (5.19)$$

\(^2\) Unfortunately, even in some popular textbooks, a few formulas pertaining to fluctuations are either incorrect, or given without specifying the conditions of their applicability, so that reader’s caution is advised.
Since Eq. (15) is valid only if system’s volume $V$ is fixed, it is customary to rewrite this extremely simple and important result as follows:

$$
\langle \bar{E}^2 \rangle = \frac{\partial \langle E \rangle}{\partial (-1/T)} = T^2 \left( \frac{\partial \langle E \rangle}{\partial T} \right)_V = C_v T^2.
$$

(5.20)

This is a remarkably simple, fundamental result. As a sanity check, for a system of $N$ similar, independent particles, $\langle E \rangle$ and hence $C_v$ and are proportional to $N$, so that $\delta E \propto N^{1/2}$ and $\delta E/\langle E \rangle \propto N^{-1/2}$, in agreement with Eq. (12). Let me emphasize that the classically-looking Eq. (20) is based on the general Gibbs distribution, and hence is valid for any system – either classical or quantum.

We will discuss the corollaries of this result in the next section, and now let me carry out a very similar calculation for a system whose number $N$ of particles in a system is not fixed, because they may go to, and come from the environment at will. If the chemical potential $\mu$ of the environment and its temperature $T$ are fixed, we are dealing with the grand canonical ensemble (Fig. 2.13), and may use the grand canonical distribution (2.106)-(2.107):

$$
W_{m,N} = \frac{1}{Z_G} \exp \left\{ \frac{\mu N - E_{m,N}}{T} \right\}, \quad Z_G = \sum_{N,m} \exp \left\{ \frac{\mu N - E_{m,N}}{T} \right\}.
$$

(5.21)

Acting exactly as we did above for energy, we get

$$
\langle N \rangle = \frac{1}{Z_G} \sum_{m,N} N \exp \left\{ \frac{\mu N - E_{m,N}}{T} \right\} = \frac{T}{Z_G} \frac{\partial Z_G}{\partial \mu},
$$

(5.22)

$$
\langle N^2 \rangle = \frac{1}{Z_G} \sum_{m,N} N^2 \exp \left\{ \frac{\mu N - E_{m,N}}{T} \right\} = \frac{T^2}{Z_G} \frac{\partial^2 Z_G}{\partial \mu^2},
$$

(5.23)

so that the particle number variance is

$$
\langle \bar{N}^2 \rangle = \langle N^2 \rangle - \langle N \rangle^2 = \frac{T^2}{Z_G} \frac{\partial Z_G}{\partial \mu} - \frac{T^2}{Z_G^2} \left( \frac{\partial Z_G}{\partial \mu} \right)^2 = T \frac{\partial}{\partial \mu} \left( \frac{T}{Z_G} \frac{\partial Z_G}{\partial \mu} \right) = T \frac{\partial \langle N \rangle}{\partial \mu},
$$

(5.24)

in the full analogy with Eq. (19).

For example, for the ideal classical gas we had Eq. (3.32). As was already emphasized in Sec. 3.2, though that result has been obtained from the canonical ensemble in that the number of particles $N$ is fixed, at $N >> 1$ the fluctuations of $N$ in the grand canonical ensemble should be relatively small, so that the same relation should be valid for average $\langle N \rangle$ in that ensemble. Solving that relation for $\langle N \rangle$, we get

$$
\langle N \rangle = \text{const} \times \exp \left\{ \frac{\mu}{T} \right\},
$$

(5.25)

where “const” means a factor that is constant at the differentiation of $\langle N \rangle$ over $\mu$, required by Eq. (24). Performing the differentiation and then using Eq. (25) again,

$$
\frac{\partial \langle N \rangle}{\partial \mu} = \text{const} \times \frac{1}{T} \exp \left\{ \frac{\mu}{T} \right\} = \frac{\langle N \rangle}{T},
$$

(5.26)
we get from Eq. (24) a surprisingly simple result:

\[
\langle \hat{N}^2 \rangle = \langle N \rangle, \quad \text{i.e.} \quad \delta N = \langle N \rangle^{1/2}
\]

(5.27)

This relation is so simple and important that I will now show how it may be derived in a different way, in order to prove that this result is valid for systems with an arbitrary (say, small) \( N \), and also get more detailed information about the statistics of fluctuations of that number. Let us consider an ideal classical gas of \( N_0 \) particles in a volume \( V_0 \), and calculate the probability \( W_N \) to have exactly \( N \leq N_0 \) of these particles in a part \( V \leq V_0 \) of this volume – see Fig. 1.

For one particle such probability is of course \( W = V/V_0 \leq 1 \), while the probability of one particle being in the remaining part of the volume is \( W' = 1 - W = 1 - V/V_0 \). If all particles were distinguishable, the probability of having \( N \leq N_0 \) specific particles in volume \( V \), and \( (N - N_0) \) specific particles in volume \( (V - V_0) \), would be \( W^N W^{(N_0-N)} \). However, if we do not distinguish the particles, we should multiply the probability by the number of possible particle combinations keeping numbers \( N \) and \( N_0 \) constant, i.e. by the binomial coefficient \( N_0!/(N_0 - N)! \). As the result, the required probability is

\[
W_N = W^N W'^{(N_0-N)} \frac{N_0!}{N!(N_0 - N)!} = \left( \frac{\langle N \rangle}{N_0} \right)^N \left( 1 - \frac{\langle N \rangle}{N_0} \right)^{N_0-N} \frac{N_0!}{N!(N_0 - N)!},
\]

(5.28)

where in the second instance I have used the evident expression \( \langle N \rangle = WN_0 = (V/V_0)N_0 \) for the average number of particles in volume \( V \). Relation (28) is the so-called binomial probability distribution, valid for any \( \langle N \rangle \) and \( N_0 \).

If we are interested in keeping \( \langle N \rangle \) arbitrary, but do not care how large the additional volume \( (V_0 - V) \) is, we can simplify the binomial distribution by assuming that the external part, and hence \( N_0 \), are very large:

\[
N_0 >> N,
\]

(5.29)

where \( N \) means all values of interest, including \( \langle N \rangle \). In this limit we can neglect \( N \) in comparison with \( N_0 \) in the second exponent of Eq. (28), and also approximate the fraction \( N_0!/(N_0 - N)! \), i.e. the product of \( N \) terms, \( (N_0 - N + 1) (N_0 - N + 2) \ldots (N_0 - 1)N_0 \), as just \( N_0^N \). As a result, we get

\[
W_N \approx \left( \frac{\langle N \rangle}{N_0} \right)^N \left( 1 - \frac{\langle N \rangle}{N_0} \right)^{N_0} \frac{N_0^N}{N!} \left( 1 - \frac{\langle N \rangle}{N_0} \right)^N \frac{\langle N \rangle^N}{N!} = \frac{\langle N \rangle^N}{N!} \left( \frac{(1 - W)}{\bar{W}} \right)^N.
\]

(5.30)

\[\text{See, e.g., MA Eq. (2.2).}\]
In the limit (29), $W \to 0$, and factor inside the square brackets tends to $1/e$, the reciprocal of the natural logarithm base.\(^4\) Thus, we finally get an expression independent of $N_0$:

$$W_N = \frac{\langle N \rangle^N}{N!} e^{-\langle N \rangle}. \quad (5.31)$$

This is the much celebrated *Poisson distribution*, which describes a very broad family of random phenomena. Figure 2 shows this distribution for several values of $\langle N \rangle$ - which, in contrast to $N$, are not necessarily integer.

At very small $\langle N \rangle$, function $W_N(N)$ distribution is close to an exponential one, $W_N \approx W^N \propto \langle N \rangle^N$, while in the opposite limit, $\langle N \rangle >> 1$, it rapidly approaches the *Gaussian* (alternatively called “normal”) distribution

$$W_N = \frac{1}{(2\pi)^{1/2} \delta N} \exp \left\{ -\frac{(N - \langle N \rangle)^2}{2(\delta N)^2} \right\}, \quad \text{with } \delta N = \langle N \rangle^{1/2}. \quad (5.32)$$

(Note that the Gaussian distribution is also valid if both $N$ and $N_0$ are large, regardless of relation (29) between them - see Fig. 3.)

\(^4\) Indeed, this is the most popular definition of this major mathematical constant – see, e.g., MA Eq. (1.2a) with $n$ replaced with $-1/W$. 

---

**Note:**

- **Fig. 5.2.** The Poisson distribution for several values of $\langle N \rangle$. In contrast to that average, argument $N$ may take only integer values, so that lines are only guides for the eye.

- **Fig. 5.3.** Hierarchy of three major probability distributions.
The key property of the Poisson (and hence of the Gaussian) distribution is that it has the same variance as given by Eq. (27):

$$\langle \tilde{N}^2 \rangle \equiv \langle (N - \langle N \rangle)^2 \rangle = \langle N \rangle.$$  \hspace{1cm} (5.33)

(This is not true for the general binomial distribution.) For our current purposes, this means that for the ideal classical gas, Eq. (27) is valid for any number of particles.

### 5.3. Volume and temperature

What are the r.m.s. fluctuations of other thermodynamic variables – like $V$, $T$, etc.? Again, the answer depends on conditions. For example, if the volume $V$ occupied by a gas is externally fixed (say, by rigid walls), it evidently does not fluctuate at all: $\delta V = 0$. On the other hand, the volume may fluctuate in the situation when average pressure is fixed – see, e.g., Fig. 1.5. A formal calculation of these fluctuations, using the approach applied in the last section, is hampered by the fact that it is physically impracticable to fix its conjugate variable, $P$, i.e. suppress its fluctuations. For example, the force $\mathcal{F}(t)$ exerted by an ideal classical gas on vessel’s wall (whose measure the pressure is) is the result of individual, independent hits of the wall by particles (Fig. 4), with time scale $\tau_c \sim r_B/(T/m)^{1/2} \sim 10^{-16}$ s, so that its frequency spectrum extends to very high frequencies, virtually impossible to control.

However, we can use the following trick, very typical for the theory of fluctuations. It is almost evident that r.m.s. fluctuations of volume are independent of the shape of the container. Let us consider the particular situation similar to that shown in Fig. 1.5, with the container of a cylindrical shape, with the base area $A$. Then the coordinate of the piston is just $q = V/A$, while the average force exerted by the gas on the cylinder is $\mathcal{F} = PA$ – see Fig. 5. Now if the piston is sufficiently massive, its free oscillation frequency $\omega$ near the equilibrium position is small enough to satisfy the following three conditions.

First, besides balancing the average force $\langle \mathcal{F} \rangle$, and thus sustaining average pressure $\langle P \rangle = \langle \mathcal{F} \rangle/A$ of the gas, the interaction between the heavy piston and light molecules of the gas is weak because of a relatively short duration of the wall hits (Fig. 4). Because of that, the full energy of the system may be presented as a sum of those of the gas and the piston, with a quadratic contribution to piston’s potential energy from small deviations of equilibrium:

$$U_p = \frac{K}{2} \tilde{q}^2, \quad \tilde{q} = q - \langle q \rangle = \frac{\mathcal{V}}{A}, \hspace{1cm} (5.34)$$

As a reminder, in geometry the term “cylinder” does not necessarily mean the “circular cylinder”; the shape of base $A$ may be arbitrary; it just should not change with height.
where $\kappa$ is the effective spring constant arising from gas’ compressibility.

Second, at $\omega \to 0$, that spring constant may be calculated just as for constant variations of volume, with the gas remaining in quasi-equilibrium at all times:

$$\kappa = -\frac{\partial \langle \mathcal{F} \rangle}{\partial q} = A^2 \left( -\frac{\partial \langle P \rangle}{\partial \langle V \rangle} \right). \quad (5.35)$$

This partial derivative\(^6\) should be taken at whatever the given thermal conditions are, e.g., with $S = \text{const}$ for adiabatic conditions (i.e., thermally insulated gas), or with $T = \text{const}$ for isothermic conditions (gas in a good thermal contact with a heat bath), etc. With that constant denoted as $X$, Eqs. (34)-(35) give

$$U_P = \frac{1}{2} \left( -A^2 \frac{\partial \langle P \rangle}{\partial \langle V \rangle} \right) \left( \frac{\langle V \rangle}{A} \right)^2 = \frac{1}{2} \left( -\frac{\partial \langle P \rangle}{\partial \langle V \rangle} \right)_X \langle \tilde{V}^2 \rangle. \quad (5.36)$$

Finally, making $\omega$ sufficiently small (namely, $\hbar \omega \ll T$) by a sufficiently large piston mass, we can apply, to the piston’s fluctuations, the classical equipartition theorem: $\langle U_P \rangle = T/2$, giving

$$\langle \tilde{V}^2 \rangle_X = T \left( -\frac{\partial \langle V \rangle}{\partial \langle P \rangle} \right)_X. \quad (5.37a)$$

Since this result is valid for any $A$ and $\omega$, it should not depend on system’s geometry and piston mass, provided that it is large in comparison with the effective mass of a single system component (say, a gas molecule) – the condition that is naturally fulfilled in most experiments.\(^7\) For the particular case of fluctuations at constant temperature ($X = T$), we may use the second of Eqs. (1.39) to rewrite Eq. (37a) as

---

\(^6\) As already was discussed in Sec. 4.1 in the context of the van der Waals equation, for mechanical stability of a gas (or liquid), derivative $\partial P/\partial V$ has to be negative, so that $\kappa$ is positive.

\(^7\) One may meet statements that a similar formula,

$$\langle \tilde{P}^2 \rangle = T \left( -\frac{\partial \langle P \rangle}{\partial \langle V \rangle} \right)_X,$$

is valid for pressure fluctuations. However, such statement does not take into account a different physical nature of pressure (Fig. 4), with its very broad frequency spectrum. This issue will be discussed later in this chapter.
\[ \langle \tilde{V}^2 \rangle_T = -T \left( \frac{\partial^2 G}{\partial \langle P \rangle^2} \right)_T. \]  

(5.37b)

In the specific case of an ideal classical gas of \( N \) particles, with the equation of state \( \langle V \rangle = NT/\langle P \rangle \), it is easier to use directly Eq. (37a), with \( X = T \), to get

\[ \langle \tilde{V}^2 \rangle_T = -T \left( -\frac{NT}{\langle P \rangle^2} \right) = \frac{(V)^2}{N}, \quad \text{i.e.,} \quad \delta V_T = \frac{1}{N^{1/2}}, \]  

(5.38)

in agreement with the trend given by Eq. (12).

Now let us proceed to fluctuations of temperature, for simplicity focusing on the case \( V = \text{const} \). Let us again assume that the system we are considering is weakly coupled to a heat bath of temperature \( T_0 \), in the sense that the time \( \tau \) of temperature equilibration between the two is much larger than the internal temperature relaxation (thermalization) time. Then we may assume that \( T \) changes in the whole system virtually simultaneously, and consider it a function of time alone:

\[ T = \langle T \rangle + \tilde{T}(t). \]  

(5.39)

Moreover, due to the (relatively) large \( \tau \), we may use the stationary relation between small fluctuations of temperature and the internal energy of the system:

\[ \tilde{T}(t) = \frac{\tilde{E}(t)}{C_v}, \quad \text{so that} \quad \delta T = \frac{\delta E}{C_v}. \]  

(5.40)

With those assumptions, Eq. (20) immediately yields the famous expression for the so-called thermodynamic fluctuations of temperature:

\[ \delta T = \frac{\delta E}{C_v} = \frac{\langle T \rangle}{C_v^{1/2}}. \]  

(5.41)

The most straightforward application of this result is to analysis of so-called bolometers - broadband detectors of electromagnetic radiation in microwave and infrared frequency bands. In such a detector (Fig. 6), the incoming radiation it focused on a small sensor (e.g., either a small piece of a Ge crystal, or a superconductor thin film at temperature \( T \approx T_c \), etc.) that is well isolated thermally from the environment.

As a result, the absorption of even small radiation power \( \mathcal{P} \) leads to a noticeable change \( \Delta T \) of sensor’s average temperature \( \langle T \rangle \) and hence of its electric resistance \( R \), which is probed up by low-noise
external electronics.\textsuperscript{8} If power does not change in time too fast, $\Delta T$ is a certain function of $P$, turning into 0 at $P=0$. Hence, if $\Delta T$ is much lower than the environment temperature $T_0$, we may keep only the main, linear term in it Taylor expansion in $P$:

$$\Delta T \equiv \langle T \rangle - T_0 = \frac{P}{q}, \quad (5.42)$$

where coefficient $q \equiv \partial P/\partial T$ is called the \textit{thermal conductance} of the unavoidable thermal coupling between the sensor and the heat bath – see Fig. 6. The power may be detected if the electric signal from the sensor, which results from change $\Delta T$, is not drowned in spontaneous fluctuations. In practical systems, these fluctuations are is contributed by several sources including electronic amplifiers, sensor, etc. However, in modern systems these “technical” contributions to noise are successfully suppressed, and the dominating noise source are the fundamental fluctuations of sensor temperature, described by Eq. (41). In this case the so-called \textit{noise-equivalent power} (“NEP”), defined as the level of $P$ that produces signal equal to r.m.s. value of noise, may be calculated by equating Eqs. (41) (with $\langle T \rangle \approx T_0$) and (42):

$$\text{NEP} = \frac{T_0 q}{C_V^{1/2}}. \quad (5.43)$$

This expression shows that in order to decrease $\text{NEP}$, i.e. improve the device sensitivity, both the environment temperature $T_0$ and thermal conductance $q$ should be reduced. In modern receivers of radiation, their typical values (in SI units) are of the order of 0.1 K and $10^{-10}$ W/K, respectively.

On the other hand, Eq. (43) implies that in order to increase bolometer sensitivity, i.e. reduce $\text{NEP}$, the $C_V$ of the sensor, and hence its mass, should be \textit{increased}. This conclusion is valid only to a certain extent, because due to technical reasons (parameter drift and the so-called 1/f noise of the sensor and external electronics), incoming power has to be modulated with as high frequency $\omega$ as possible (in most cases, the cyclic frequency $\nu = \omega/2\pi$ of the modulation is between 10 to 1,000 Hz), so that the electrical signal may be picked up from the sensor at that frequency. As a result, $C_V$ may be increased only until the thermal constant of the sensor,

$$\tau = \frac{C_V}{q}, \quad (5.44)$$

becomes close to $1/\omega$, because at $\omega \tau >> 1$ the useful signal drops faster than noise. As a result, the lowest (i.e. the best) value of $\text{NEP}$,

$$\frac{(\text{NEP})_{\min}}{\nu^{1/2}} = \alpha T_0 q^{1/2}, \quad \alpha \sim 1, \quad (5.45)$$

is reached at $\nu \tau \approx 1$. (The exact values of the optimal product $\omega \tau$, and the numerical constant $\alpha \sim 1$ in Eq. (45), depend on the exact law of power modulation in time, and the output signal processing procedure.) With the parameters cited above, this estimate yields $(\text{NEP})_{\min}/\nu^{1/2} \sim 3 \times 10^{-17}$ W/Hz$^{1/2}$ – a very low power indeed.

\textsuperscript{8} Besides low internal electric noise, the sensor should have a sufficiently large \textit{temperature responsivity} $dR/dT$, making the noise contribution by the pickup electronics insignificant – see below.
However, surprisingly enough, the power modulation allows bolometric (and other broadband) receivers to register radiation with power much lower than this NEP! Indeed, picking up the sensor signal at the modulation frequency $\omega$, we can use the following electronics stages to filter out all the noise besides its components within a very narrow band, of width $\Delta \nu \ll \nu$, around the modulation frequency (Fig. 7). This is the idea of a microwave radiometer,\(^9\) currently used in all sensitive broadband receivers.

In order to analyze this opportunity, we need to develop theoretical tools for a quantitative description of the spectral distribution of fluctuations. Another motivation for that description is the need in analysis of variables dominated by fast (high-frequency) components, such as pressure – please have one more look at Fig. 4. Finally, during the analysis, we will run into the fundamental relation between fluctuations and dissipation, which is one of the main results of statistical physics as a whole.

5.4. Fluctuations as functions of time

There are two mathematically-equivalent approaches to time-dependent functions of time, called time-domain and frequency-domain pictures, with their relative convenience depending on the particular problem to be solved.

In the time domain, we cannot characterize a random fluctuation $\tilde{f}(t)$ of a classical variable by its statistical average, because it equals zero – see Eq. (2). Of course, variance (3) does not vanish, but if fluctuations are stationary, it does not depend on time either. Because of that, let us consider the following average:\(^{10}\)

$$\langle \tilde{f}(t)\tilde{f}(t') \rangle. \quad (5.46)$$

Generally, this is a function of two arguments. Moreover, in the systems that are stationary (whose macroscopic parameters and hence the variable expectation values do not change with time), averages like (46) may depend only on the difference,

$$\tau \equiv t' - t, \quad (5.47)$$

\(^9\) It was pioneered in the 1950s by R. Dicke, so that the device is frequently called the Dicke radiometer.

\(^{10}\) Clearly, this is a temporal analog of the spatial correlation function discussed in Sec. 4.2 – see Eq. (4.30).
between the two observation times. In this case, average (46) is called the \textit{correlation function} of variable $f$:

\[
K_f(\tau) \equiv \left\langle \tilde{f}(t)\tilde{f}(t+\tau) \right\rangle. \tag{5.48}
\]

This name\(^{11}\) catches the idea of this notion very well: $K_f(\tau)$ tells us about the average mutual relation between the fluctuations at two times separated by interval $\tau$. Let us list the basic properties of this function.

First of all, $K_f(\tau)$ has to be an even function of the time delay $\tau$. Indeed, we may write

\[
K_f(-\tau) = \left\langle \tilde{f}(t)\tilde{f}(t-\tau) \right\rangle = \left\langle \tilde{f}(t-\tau)\tilde{f}(t) \right\rangle = \left\langle \tilde{f}(t')\tilde{f}(t'+\tau) \right\rangle, \tag{5.49}
\]

with $t' \equiv t - \tau$. For stationary processes, this average cannot depend on the common shift $t'$ of the two observation times, so that averages (48) and (49) have to be equal:

\[
K_f(-\tau) = K_f(\tau). \tag{5.50}
\]

Second, at $\tau \to 0$ the correlation function tends to the variance:

\[
K_f(0) = \left\langle \tilde{f}(t)\tilde{f}(t) \right\rangle = \left\langle \tilde{f}^2 \right\rangle. \tag{5.51}
\]

In the opposite limit, when $\tau$ is much larger than some characteristic \textit{correlation time} $\tau_c$ of the system,\(^{12}\) the correlation function tends to zero, because fluctuations separated by such large time interval are virtually independent (uncorrelated). As a result, the correlation function typically looks like one of the plots sketched in Fig. 8. Note that on a time scale much longer than $\tau_c$, any physically-realistic correlation function may be well approximated with a delta-function of $\tau$.\(^{13}\)

![Fig. 5.8. Correlation function of fluctuations: two typical examples.](image)

In the reciprocal, frequency domain, process $\tilde{f}(t)$ is presented as a Fourier integral,

\[
\tilde{f}(t) = \int_{-\infty}^{+\infty} f_\omega e^{-i\omega t} d\omega, \tag{5.52}
\]

with the reciprocal transform being

\(^{11}\) Another term, the \textit{autocorrelation function}, is sometimes used for average (48) to distinguish it from the \textit{mutual correlation function}, $\langle f(t)g(t+\tau) \rangle$, of two stationary processes.

\(^{12}\) Correlation time $\tau_c$ is the direct temporal analog of the correlation radius $r_c$ which was discussed in Sec. 4.2.

\(^{13}\) For example, for a process which is a sum of independent very short pulses, e.g., the gas pressure force exerted on the container wall (Fig. 4), such approximation is legitimate on time scales longer than the single pulse duration, e.g., the time of particle’s impact on the wall.
\[ f_\omega = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{f}(t)e^{i\omega t} dt . \]  

(5.53)

If the initial function \( \tilde{f}(t) \) is random (as it is in the case of fluctuations), with zero average, its Fourier transform \( f_\omega \) is a random function (now of frequency) as well, also with a vanishing statistical average:

\[ \langle f_\omega \rangle = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \langle \tilde{f}(t)e^{i\omega t} \rangle dt \frac{1}{2\pi} \int_{-\infty}^{+\infty} \langle \tilde{f}(t) \rangle e^{i\omega t} dt = 0 . \]  

(5.54)

The simplest nonvanishing average may be formed similarly to Eq. (46), but with due respect to the complex-variable character of the Fourier images:

\[ \langle f_\omega f_\omega^* \rangle = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \langle \tilde{f}(t)\tilde{f}(t') \rangle e^{i(\omega' t' - \omega t)} dt dt' . \]  

(5.55)

It turns out that for a stationary process, averages (46) and (55) are directly related. Indeed, since the integration over \( t' \) in Eq. (55) is in infinite limits, we may replace it with integration over \( \tau \equiv t' - t \) (at fixed \( t \)), also in infinite limits. Replacing \( t' \) by \( t + \tau \) in expressions under the integral, we see that the average is just the correlation function \( K_f(\tau) \), while the time exponent is equal to \( \exp\{i(\omega' - \omega)t\}\exp\{i\omega'\tau\} \). As a result, changing the order of integration, we get

\[ \langle f_\omega f_\omega^* \rangle = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} d\tau K_f(\tau) e^{i(\omega - \omega')\tau} e^{i\omega t} = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} d\tau K_f(\tau) e^{i\omega'\tau} d\tau \int_{-\infty}^{+\infty} d\tau e^{i(\omega - \omega')\tau} dt . \]  

(5.56)

But the last integral is just \( 2\pi \delta(\omega - \omega') \), so that we finally get

\[ \langle f_\omega f_\omega^* \rangle = S_f(\omega) \delta(\omega - \omega') , \]  

(5.57)

where the real function of frequency,

\[ S_f(\omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} K_f(\tau)e^{i\omega \tau} d\tau = \frac{1}{\pi} \int_{0}^{+\infty} K_f(\tau) \cos \omega \tau d\tau , \]  

(5.58)

is called the \textit{spectral density of fluctuations at frequency} \( \omega \). According to Eq. (58), the spectral density is a Fourier image of the correlation function, and hence the reciprocal Fourier transform is:

\[ K_f(\tau) = \int_{-\infty}^{+\infty} S_f(\omega)e^{-i\omega \tau} d\omega = 2\pi \int_{0}^{+\infty} S_f(\omega) \cos \omega \tau d\omega . \]  

(5.59)

In particular, for the variance, Eq. (59) yields

---

14 See, e.g., MA Eq. (14.4a).

15 The second form of Eq. (59) uses the fact that, according to Eq. (58), \( S_f(\omega) \) is an even function of frequency - just as \( K_f(\tau) \) is an even function of time.

16 Although Eqs. (58) and (59) look not much more than straightforward corollaries of the Fourier transform, they bear a special name of the \textit{Wiener-Khinchin theorem} – after mathematicians N. Wiener and A. Khinchin who have proved that these relations are valid even for functions \( f(t) \) which are not square-integrable, so that from the point of view of rigorous mathematics, their Fourier transforms are not well defined.
This relation shows that term “spectral density” describes the physical sense of function $S_f(\omega)$ very well. Indeed, if a random signal $f(t)$ had been passed through a frequency filter with a small bandwidth $\Delta \nu \ll \nu$ of positive cyclic frequencies, the integral in Eq. (60) had to be limited to interval $\Delta \omega = 2\pi \Delta \nu$, i.e. that the variance of the output signal would become:

$$\left\langle \tilde{f}^2 \right\rangle_{\Delta \nu} = 2S_f(\omega)\Delta \omega = 4\pi S_f(\omega)\Delta \nu.$$  \hspace{1cm} (5.61)

To complete this introductory section, let me note an important particular case. If the spectral density of some process, is nearly constant within the frequency range of interest, $S_f(\omega) = \text{const} = S_f(0),$\(^{18}\) Eq. (59) shows that its correlation function may be well approximated by a delta-function:

$$K_f(\tau) = S_f(0)\int_{-\infty}^{\infty} e^{-i\omega \tau} d\omega = 2\pi S_f(0)\delta(\tau).$$ \hspace{1cm} (5.62)

From this relation stems another popular name of the white noise, the \textit{delta-correlated process}. We have already seen that this is a very reasonable approximation, for example, for the gas pressure force fluctuations (Fig. 4). Of course, for spectral density of a realistic, limited physical variable the approximation of constant spectral density cannot be true for \textit{all} frequencies (otherwise, for example, integral (60) would diverge, giving an unphysical, infinite value of variance), and is valid only at frequencies much lower than $1/\tau_c$.

### 5.5. Fluctuations and dissipation

Now we are mathematically equipped to address one of the most important topics of statistical physics, the relation between fluctuations and dissipation. This relation is especially simple for the following hierarchical situation: a relatively “heavy”, slowly moving system interacting with an environment consisting of rapidly moving, “light” components. A popular theoretical term for such a system is the \textit{Brownian particle}, named after botanist R. Brown who first noticed in 1827 the random motion of pollen grains, caused by their random hits by fluid molecules, under a microscope. However, the family of such systems is much broader than that of mechanical particles.\(^{19}\)

One more important assumption of this theory is that the system’s motion does not violate the thermal equilibrium of the environment - well fulfilled in many cases. (Think, for example, about a usual mechanical pendulum whose motion does not overheat the air around it.) In this case, the statistical averaging over the thermally-equilibrium environment may be performed for any (slow)

\(^{17}\) A popular alternative definition of the spectral density is $S_f(\nu) = 4\pi S_f(\omega)$, making average (61) equal to $S_f(\nu)\Delta \nu$.

\(^{18}\) Such process is frequently called \textit{white noise}, because it consists of all frequency components with equal amplitudes, reminding the white light, which consists of many monochromatic components.

\(^{19}\) Just for one example, such description may be valid for the complex amplitude of an electromagnetic field mode weakly interacting with matter. To emphasize this generality, I will use letter $q$ rather than $x$ for “particle’s” coordinate.
motion of the system of interest, considering the motion fixed. 20 I will denote such a “primary” averaging by angular brackets $\langle \ldots \rangle$. At a later stage we may carry out another, “secondary” averaging, over an ensemble of many similar systems of interest, coupled to similar environments. If we do, it will be denoted by double angle brackets $\langle\langle \ldots \rangle\rangle$.

Let me start from a simple classical system, a 1D harmonic oscillator whose equation of evolution may be presented as

$$m \ddot{q} + \kappa q = F_{\text{det}}(t) + F_{\text{env}}(t) = \dot{F}_{\text{det}}(t) + \dot{F}_{\text{env}}(t) = \langle \dot{F} \rangle + \dot{\vec{F}}(t),$$

(5.63)

where $q$ is the (generalized) coordinate of the oscillator, $F_{\text{det}}(t)$ is the deterministic (generalized) external force, while both components of the random force $\vec{F}(t)$ present the impact of the environment on oscillator’s motion. Again, from the point of view of the fast-moving environmental components, the oscillator’s motion is slow. The average of the force exerted by environment on such a slowly moving object may have a part depending on not only $q$, but on the velocity $\dot{q}$ as well. For most systems, the Taylor expansion of the force in small velocity would have a finite leading, linear term, so that we may take

$$\langle \dot{F} \rangle = -\eta \dot{q},$$

(5.64)

so that Eq. (63) may be rewritten as

$$m \ddot{q} + \eta \dot{q} + \kappa q = F_{\text{det}}(t) + \dot{\vec{F}}(t).$$

(5.65)

This way of describing the effects of environment on an otherwise Hamiltonian system is called the Langevin equation. 21 Due to the linearity of the differential equation (65), its general solution may be presented as a sum of two parts: the deterministic motion of the linear oscillator due to the external force $\dot{F}_{\text{det}}(t)$, and random fluctuations due to the random force exerted by the environment. The former effects are well known from classical dynamics, 22 so let us focus on the latter part by taking $\dot{F}_{\text{det}}(t) = 0$. The remaining term in the right-hand part describes the fluctuating part of the environmental force; in contrast to the average component (64), its intensity (read: its spectral density at relevant frequencies $\omega \sim \omega_0 \equiv (\kappa/m)^{1/2}$) does not vanish at $q(t) = 0$, and hence may be evaluated ignoring system’s motion.

Plugging into Eq. (65) the presentation of both variables in the form similar to Eq. (52), for their Fourier images we get the following relation:

$$-m \omega^2 q_\omega - i \omega \eta q_{\omega} + \kappa q_{\omega} = \dot{F}_\omega.$$  

(5.66)

which immediately gives us $q_{\omega}$.

---

20 For a usual (ergodic) environment, the primary averaging may be interpreted as that over relatively short time intervals, $\tau_c \ll \Delta t \ll \tau$, where $\tau_c$ is the correlation time of the environment, while $\tau$ is the characteristic time scale of motion of our “heavy” system of interest.

21 After P. Langevin whose 1908 work was the first systematic development of A. Einstein’s ideas on Brownian motion (see below) using this formalism. A detailed discussion of this approach, with numerical examples of its application, may be found, e.g., in the monograph by W. Coffey, Yu. Kalmykov, and J. Waldron, The Langevin Equation, World Scientific, 1996.

22 See, e.g., CM Sec. 4.1. In this and the next sections I assume that variable $f(t)$ is classical, with the discussion of the quantum case postponed until Sec. 6.
\[ q_\omega = \frac{\mathcal{F}_\omega}{(\kappa - m\omega^2) - i\eta \omega}. \] (5.67)

Now multiplying Eq. (67) by its complex conjugate, averaging both parts of the resulting equation, and using for each of them Eq. (57), we get the following relation between spectral densities of the oscillations and force:

\[ S_q(\omega) = \frac{1}{(\kappa - m\omega^2)^2 + (\eta \omega)^2} S_\mathcal{F}(\omega). \] (5.68)

As the reader should know well from classical dynamics, at small damping (\( \eta << m\omega_0 \)) the first factor in the right-hand part of Eq. (68) describes the resonance, i.e. has a sharp peak near oscillator’s eigenfrequency \( \omega_0 \), and may be presented in that vicinity as

\[ \frac{1}{(\kappa - m\omega^2)^2 + (\eta \omega)^2} \approx \frac{1}{4m\kappa(\xi^2 + \delta^2)}, \quad \text{at } |\xi| << \omega_0 \quad \text{with } \xi \equiv \omega - \omega_0, \quad \delta \equiv \eta / 2m. \] (5.69)

In contrast, spectral density \( S_\mathcal{F}(\omega) \) of fluctuations of a typical environment is changing slowly near that frequency, so that for the purpose of integration over frequencies near \( \omega_0 \) we may replace \( S_\mathcal{F}(\omega) \) with \( S_\mathcal{F}(\omega_0) \). As a result, the variance of the environment-imposed random oscillations may be calculated as

\[ \left\langle \left\langle \tilde{q}^2 \right\rangle \right\rangle = 2 \int S_q(\omega) d\omega \approx 2 \int_{\omega = \omega_0} S_q(\omega) d\omega \approx 2S_\mathcal{F}(\omega_0) \frac{1}{4m\kappa} \int_{-\infty}^{+\infty} \frac{d\xi}{\xi^2 + \delta^2}. \] (5.70)

The last expression includes a well-known table integral, equal to \( \pi/\delta = 2\pi m/\eta \), so that finally

\[ \left\langle \left\langle \tilde{q}^2 \right\rangle \right\rangle = 2S_\mathcal{F}(\omega_0) \frac{1}{4m\kappa} \frac{2\pi m}{\eta} = \frac{\pi}{\kappa \eta} S_\mathcal{F}(\omega_0). \] (5.71)

But on the other hand, the weak interaction with environment should keep the oscillator in thermodynamic equilibrium at the same temperature \( T \). Since our analysis has been based on the classical Langevin equation (65), we may only use it in the classical limit \( \hbar \omega_0 << T \), in which we may use the equipartition theorem (2.48). In our current notation, it yields

\[ \frac{\kappa}{2} \left\langle \left\langle \tilde{q}^2 \right\rangle \right\rangle = \frac{T}{2}. \] (5.72)

Comparing Eqs. (71) and (72), we see that the spectral density of the random force exerted by environment is fundamentally related to the damping it provides:

\[ S_\mathcal{F}(\omega_0) = \frac{\eta}{\pi} T. \] (5.73a)

Now we may argue (rather convincingly :-) that since this relation does not depend on oscillator’s parameters \( m \) and \( \kappa \), and hence its eigenfrequency \( \omega_0 = (\kappa m)^{1/2} \), it should be valid at any (but

---

23 At this stage we restrict our analysis to random, stationary processes \( q(t) \), so that Eq. (57) is valid for this variable as well, if the averaging is understood in the \( \langle \langle \ldots \rangle \rangle \) sense.

24 See, e.g. MA Eq. (6.5a).
sufficiently low, $\omega \tau_c << 1$) frequency. Using Eq. (58) with $\omega \rightarrow 0$, it may be rewritten as a formula for the effective low-frequency drag (friction) coefficient:

$$\eta = \frac{1}{T} \int_0^\infty K_\phi(\tau) d\tau \equiv \frac{1}{T} \int_0^\infty \langle \tilde{\mathcal{F}}(0) \tilde{\mathcal{F}}(\tau) \rangle d\tau.$$  \hspace{1cm} (5.73b)

Relation (73) reveals an intimate, fundamental connection between fluctuations and dissipation provided by a thermally-equilibrium environment. Verbally, “there is no dissipation without fluctuations” - and vice versa. 25 Historically, this fact was first recognized in 1905 by A. Einstein, 26 in the following form. Let us apply our result (73) to the particular case of a free 1D Brownian particle, by taking $\kappa = 0$. In this case both equations (71) and (72) give infinities. In order to understand the reason for that divergence, let us go back to the Langevin equation (65) with not only $\kappa = 0$, but also, just for the sake of simplicity, $m \rightarrow 0$ as well. (The latter approximation, frequently called the overdamping limit, is quite appropriate for the motion of a small particle in a viscous fluid, when $m << \eta \Delta t$ even for smallest time intervals $\Delta t$ between the successive observations of particle’s positions.) In this approximation, Eq. (65) is reduced to a simple equation,

$$\eta \dot{q} = \mathcal{F}_{\text{det}}(t) + \langle \tilde{\mathcal{F}}(t) \rangle,$$ \hspace{1cm} (5.74)

with a ready solution for particle displacement during a finite time interval $t$:

$$\Delta q(t) \equiv q(t) - q(0) = \langle \langle \Delta q(t) \rangle \rangle + \tilde{\mathcal{F}}(t), \quad \langle \langle \Delta q(t) \rangle \rangle = \frac{1}{\eta} \int_0^t \mathcal{F}_{\text{det}}(t') dt', \quad \Delta \tilde{\mathcal{F}}(t) = \frac{1}{\eta} \int_0^t \tilde{\mathcal{F}}(t') dt'.$$ \hspace{1cm} (5.75)

Evidently, in the statistical average of the displacement, the fluctuation effects vanish, but this does not mean that the particle does not deviate from the deterministic trajectory $\langle \langle q(t) \rangle \rangle$ – just that is has equal probabilities to be shifted either of two possible directions from that trajectory. To see that, let us calculate the variance of the displacement:

$$\langle \langle \Delta \tilde{\mathcal{F}}^2(t) \rangle \rangle = \frac{1}{\eta^2} \int_0^t dt' \int_0^{t'} dt'' \langle \tilde{\mathcal{F}}(t') \tilde{\mathcal{F}}(t'') \rangle = \frac{1}{\eta^2} \int_0^t dt' \int_0^{t'} dt'' K_\phi(t' - t'').$$ \hspace{1cm} (5.76)

As we already know, at times $\tau \gg \tau_c$ (this correlation time, for typical molecular impacts, is of the order of a picosecond), correlation function may be well approximated by the delta-function – see Eq. (62). In this approximation, with $S_\phi(0)$ expressed by Eq. (73), and Eq. (80) yields

$$\langle \langle \Delta \tilde{\mathcal{F}}^2(t) \rangle \rangle = \frac{2\pi}{\eta^2} S_\phi(0) \int_0^t dt' \int_0^{t'} dt'' \delta(t - t') = \frac{2\pi}{\eta^2} \frac{\eta T}{\tau} \int_0^t dt' \int_0^{t'} dt'' \delta(t - t') = 2Dt,$$ \hspace{1cm} (5.77)

with

\begin{itemize}
  \item 25 This means that the phenomenological description of dissipation by bare friction in classical mechanics (see, e.g., CM Sec. 4.1) is only valid approximately, when the energy scale of the process is much larger than $T$.
  \item 26 It was published in one of the three papers of Einstein’s celebrated 1905 “triad”. As a reminder, another paper started the (special) relativity, and one more was the quantum description of photoelectric effect, essentially the prediction of light quanta – photons, which essentially started quantum mechanics. (Not too bad for one year!)
\end{itemize}
The final form of Eq. (77) describes the well-known law of diffusion ("random walk") of a 1D system, with the r.m.s. deviation from the point of origin growing as $(2Dt)^{1/2}$. Coefficient $D$ in this relation is called the coefficient of diffusion, and Eq. (78) describes the extremely simple Einstein relation between that coefficient and particle's damping. Often this relation is rewritten in SI units of temperature as $D = \mu \kappa T$, where $\mu \equiv 1/\eta$ is the mobility of the particle. The physical sense of $\mu$ becomes clear from rewriting the expression for the deterministic viscous motion, $\langle \langle q(t) \rangle \rangle$ (particle's "drift") in the form:

$$v_{\text{drift}} \equiv \frac{d\langle \langle q(t) \rangle \rangle}{dt} = \frac{1}{\eta} J_{\text{det}}(t) = \mu J_{\text{det}}(t), \quad (5.79)$$

so that mobility is just velocity given to the particle by unit force.\(^{27}\)

Another famous example of application of Eq. (73) is to the thermal (or "Johnson", or "Johnson-Nyquist", or just "Nyquist") noise in resistive electron devices. Let us consider a two-terminal "probe" circuit, playing the role of the harmonic oscillator in our analysis above, connected to a resistor $R$ (Fig. 9), playing the role of noisy environment. (The noise is generated by the thermal motion of numerous electrons, randomly moving inside the resistor.) For this system, one convenient choice of conjugate variables (the generalized coordinate and generalized force) is, respectively, the electric charge $Q \equiv \int I(t)dt$ that has passed through the "probe" circuit by time $t$, and voltage $V$ across its terminals, with the polarity shown in Fig. 9. (Indeed, product $VdQ$ is indeed the elementary work $dW$ done by the environment on the probe circuit.)

Making the corresponding replacements, $q \rightarrow Q$ and $F \rightarrow V$ in Eq. (64), we see that it becomes

$$\eta \dot{Q} \equiv -\eta I = \langle V \rangle. \quad (5.80)$$

Comparing this relation with Ohm's law, $R(-I) = \eta$,\(^{28}\) we see that in this case, coefficient $\eta$ has the physical sense of the usual Ohmic resistance $R$,\(^{29}\) so that Eq. (73) becomes

\(^{27}\) In solid-state physics and electronics, mobility is more frequently defined as $|v_{\text{drift}}/E| = e|v_{\text{drift}}/F_{\text{det}}|$ (where $E$ is the applied electric field), and is traditionally measured in cm$^2$/V·s. In these units, the electron mobility in silicon wafers used for integrated circuit fabrication (i.e. the solid most important for engineering practice) at room temperature is close to $10^3$.

\(^{28}\) The minus sign is due to the fact that in our notation, current through the resistor equals $(-I)$ – see Fig. 9.

\(^{29}\) Due to this fact, Eq. (64) is often called the Ohmic model of the environment response, even if the physical nature of variables $q$ and $F$ is completely different from the electric charge and voltage.
Using Eq. (61), and transferring to the SI units of temperature \((T \rightarrow k_B T_K)\), we can bring this famous Nyquist formula\(^{30}\) to its most popular form

\[
\langle \bar{\nu}^2 \rangle_{\Delta \nu} = 4k_B T_K R \Delta \nu .
\] (5.81b)

Note that according to Eq. (65), this result is only valid at a negligible speed of change of the generalized coordinate \(q\) (in this case, negligible current \(I\)), i.e. Eq. (81) expresses the voltage fluctuations as would be measured by an ideal voltmeter, with an input resistance much higher than \(R\).

On the other hand, applying a different choice of generalized coordinate and force, \(q \rightarrow \Phi, \bar{\nu} \rightarrow I\) (where \(\Phi \equiv \int \nu(t) dt\) is the generalized magnetic flux, so that \(d \nu = Id\Phi\), we get \(\eta \rightarrow 1/R\), and Eq. (73) yields the thermal fluctuations of the current through the resistor (as measured by an ideal ammeter, i.e. at \(\nu \rightarrow 0\)):

\[
S_i(\omega) = \frac{1}{\pi R} T, \quad \text{i.e.} \quad \langle I^2 \rangle_{\Delta \nu} = \frac{4k_B T_K}{R} \Delta \nu .
\] (5.81c)

Note that Eqs. (81) as valid for noise in thermal equilibrium only. In electric circuits, which may be readily driven out of equilibrium by applied voltage \(\langle \nu \rangle\), other types of noise are frequently important, notably the shot noise, which arises in short conductors, e.g., tunnel junctions, at applied voltages \(\langle \nu \rangle >> T/q\), due to the discreteness of charge carriers.\(^{31}\) A straightforward analysis using a simple model, described in the assignment of Exercise Problem 9, shows that this noise may be characterized by current fluctuations with low-frequency spectral density

\[
S_i(\omega) = \frac{|q\nu|}{2\pi}, \quad \text{i.e.} \quad \langle I^2 \rangle_{\Delta \nu} = 2|q\nu| \Delta \nu ,
\] (5.82)

where \(q\) is the electric charge of a single current carrier. This is the Schottky formula, valid for any relation between \(I\) and \(\nu\). Comparison of Eqs. (81c) and (82) for a device that obeys the Ohm law shows that the shot noise has the same intensity as the thermal noise with effective temperature

\[
T_{\text{ef}} = \frac{|q\nu|}{2} >> T .
\] (5.83)

This relation may be interpreted as a result of charge carrier overheating by the applied electric field, and explains why the Schottky formula (82) is only valid in conductors much shorter than the energy

---

\(30\) Named after H. Nyquist who derived this formula in 1928 (independently of the prior work by A. Einstein, M. Smoluchowski, and P. Langevin) to describe the noise which had been just discovered experimentally by his Bell Labs’ colleague J. B. Johnson. The derivation of Eq. (73) and hence Eq. (81) in these notes is essentially a twist of the derivation used by Nyquist.

\(31\) Another practically important type of fluctuations in electronic devices is the low-frequency \(1/f\) noise which was already mentioned in Sec. 3 above. I will briefly discuss it in Sec. 8.
relaxation length \( l_e \) of the charge carriers.\(^{32}\) Another mechanism of the shot noise suppression, that becomes noticeable if system’s transparency is high, is the Fermi-Dirac statistics of electrons.\(^{33}\)

Returning to the bolometric Dicke radiometer (see Figs. 6-7 and their discussion), we may now use the Langevin equation formalism to finalize its analysis. For this system, the Langevin equation is just the usual equation of heat balance:

\[
C_v \frac{dT}{dt} + \mathcal{g}(T - T_0) = \mathcal{P}_{\text{det}}(t) + \tilde{\mathcal{P}}(t),
\]

where \( \mathcal{P}_{\text{det}} = \langle \mathcal{P} \rangle \) describes the (deterministic) power of absorbed radiation, and \( \tilde{\mathcal{P}} \) presents the effective source of temperature fluctuations. Now we can use Eq. (84) to carry out a calculation of the spectral density \( S_T(\omega) \) of temperature fluctuations absolutely similar to how this was done with Eq. (65), assuming that the frequency spectrum of the fluctuation source is much broader than the intrinsic bandwidth \( 1/\tau = \mathcal{g}/C_v \) of the bolometer, so that its spectral density at frequencies \( \omega \tau \sim 1 \) may be well approximated by its low-frequency value \( S_p(0) \):

\[
S_T(\omega) = \left| \frac{1}{-i\omega C_v + \mathcal{g}} \right|^2 S_p(0).
\]

Then, requiring the variance of temperature fluctuations,

\[
(\delta T)^2 \equiv \langle \tilde{T}^2 \rangle = 2 \int_0^\infty S_T(\omega) d\omega = 2S_p(0) \int_0^\infty \frac{1}{-i\omega C_v + \mathcal{g}} d\omega = 2S_p(0) \frac{1}{C_v} \int_0^\infty \frac{d\omega}{\omega^2 + (\mathcal{g}/C_v)^2} = \pi S_p(0) \frac{\mathcal{g}}{C_v} = \frac{\mathcal{g}}{\pi} T_0^2.
\]

The r.m.s. value of the “power noise” \( \tilde{\mathcal{P}} \) within bandwidth \( \Delta \nu \ll 1/\tau \) (Fig. 7) becomes equal to the deterministic signal power \( \mathcal{P}_{\text{det}} \) (or more exactly, the main harmonic of its modulation law) at

\[
\mathcal{P}_{\text{min}} = \left( \left\langle \tilde{\mathcal{P}}^2 \right\rangle_{\Delta \nu} \right)^{1/2} = (2S_p(0)\Delta \omega)^{1/2} = 2(\mathcal{g}\Delta \nu)^{1/2} T_0.
\]

This result shows that our earlier prediction (45) may be improved by a substantial factor of the order of \((\Delta \nu/\nu)^{1/2}\), where the reduction of the output bandwidth is limited only by the signal accumulation time \( \Delta t \sim 1/\Delta \nu \), while the increase of \( \nu \) is limited by the speed of (typically, mechanical) devices performing the power modulation. In practical systems this factor may improve the sensitivity by a couple orders of magnitude, enabling observation of extremely weak radiation. Maybe the most spectacular example are the recent measurements of the CMB radiation (discussed in Sec. 2.6), which corresponds to blackbody temperature \( T_K \approx 2.725 \text{ K} \), with accuracy \( \delta T_K \sim 10^{-6} \text{ K} \), using microwave

---

\(^{32}\) See, e.g., Y. Naveh et al., Phys. Rev. B **58**, 15371 (1998). In practically used metals, \( l_e \) is of the order of 30 nm even at liquid helium temperatures (and even shorter at ambient conditions), so that the usual “macroscopic” resistors do not exhibit the shot noise.

receivers with physical temperature of all their components much higher than $\delta T$. The observed weak ($\sim 10^{-5}$ K) anisotropy of the CMB radiation is a major experimental basis of all modern cosmology.

Let me also note that Eq. (73) may be readily generalized to the case when environment’s response is different from the Ohmic model (64). This generalization is virtually evident from Eq. (66). Indeed, the second term in its left-hand part is just the Fourier component of the average response of the environment:

$$\left\langle \hat{f}_\omega \right\rangle = i \omega \eta \hat{q}_\omega .$$  \hspace{1cm} (5.89)

Let the environment’s response be still linear, but have an arbitrary dispersion,

$$\left\langle \hat{f}_\omega \right\rangle = \chi(\omega)q_\omega ,$$  \hspace{1cm} (5.90)

where the function $\chi(\omega)$, called the generalized susceptibility of the environment, may be complex, i.e. have both the imaginary and real parts:

$$\chi(\omega) = \chi'(\omega) + i \chi''(\omega) .$$  \hspace{1cm} (5.91)

Then Eq. (73) remains valid$^{34}$ with the replacement $\eta \rightarrow \chi''(\omega)/\omega$:

$$S_f(\omega) = \frac{\chi''(\omega)}{\pi \omega} T .$$  \hspace{1cm} (5.92)

This fundamental relation$^{35}$ is used not only to calculate the fluctuation intensity from the known generalized responsibility (i.e. the deterministic response of a complex system to a small perturbation), but sometimes in the opposite direction – to calculate the linear response from the known fluctuations. (The latter use is especially attractive at numerical simulations, such as molecular dynamics approaches, because it allows to avoid filtering a weak response from the noisy background.)

Now let us discuss what generalization of Eq. (92) is necessary to make that fundamental result suitable for arbitrary temperatures, $T \sim \hbar \omega$. The calculations we had performed started from the apparently classical equation of motion, Eq. (63). However, quantum mechanics shows$^{36}$ that a similar equation is valid for the corresponding Heisenberg-picture operators, so that repeating all arguments leading to the Langevin equation (65), we may write its quantum-mechanical version

$$\hat{m} \ddot{q} + \eta \dot{q} + \kappa \hat{q} = \hat{q}_{\text{det}} + \hat{\mathcal{F}} .$$  \hspace{1cm} (5.93)

---

$^{34}$ Reviewing the calculations leading to Eq. (73), we may see that if the possible real part $\chi'(\omega)$ of the susceptibility just adds up to $(k - m \omega^2)$ in the denominator of Eq. (67), resulting in a change of oscillator’s eigenfrequency. This renormalization is insignificant if the oscillator-to-environment coupling is weak, i.e. susceptibility $\chi(\omega)$ small, as had been assumed at the derivation of Eq. (69) and hence Eq. (73).

$^{35}$ It is sometimes called the Green-Kubo (or just “Kubo”) formula. This is hardly fair, because, as the reader could see, Eq. (92) is just an elementary generalization of the Nyquist formula (81). Moreover, the corresponding works of M. Green and R. Kubo were published, respectively, in 1954 and 1957, i.e. after the 1950 paper by H. Callen and T. Welton, where a more general result (see below) had been derived. More adequately, the Green / Kubo names are associated with a related relation between the response function and the operator commutator – see, e.g., QM Eq. (7.109).

$^{36}$ See, e.g., QM Sec. 4.6.
This is the so-called the Heisenberg-Langevin (or “quantum Langevin”) equation – in this particular case, for a harmonic oscillator.

The further operations, however, require certain caution, because the right-hand part of the equation is now an operator, and has some nontrivial properties. For example, the “values” of the Heisenberg operator, representing the same variable \( f(t) \) at different times, do not necessarily commute:

\[
\left[ \hat{f}(t), \hat{f}(t') \right] \neq 0, \quad \text{if } t' \neq t. \tag{5.94}
\]

As a result, the function defined by Eq. (46) may not be an even function of time delay \( \tau = t' - t \) even for a stationary process, making it inadequate for representation of the real correlation function - which has to obey Eq. (51). This technical difficulty may be circumvented by the introduction of the following symmetrized correlation function

\[
K_f(\tau) \equiv \frac{1}{2} \left\langle \hat{f}(t)\hat{f}(t+\tau) + \hat{f}(t+\tau)\hat{f}(t) \right\rangle = \frac{1}{2} \left\langle \left\{ \hat{f}(t), \hat{f}(t+\tau) \right\} \right\rangle, \tag{5.95}
\]

(where \( \left\{ \ldots, \ldots \right\} \) denotes the anticommutator of the two operators), and, similarly, the symmetrical spectral density \( S_f(\omega) \), defined by relation

\[
S_f(\omega)\delta(\omega - \omega') \equiv \frac{1}{2} \left\langle \hat{f}_\omega\hat{f}^*_\omega + \hat{f}^*_\omega\hat{f}_\omega \right\rangle \equiv \frac{1}{2} \left\langle \left\{ \hat{f}_\omega, \hat{f}^*_\omega \right\} \right\rangle, \tag{5.96}
\]

with \( K_f(\tau) \) and \( S_f(\omega) \) still related by the Fourier transform (59).\(^{37}\)

Now we may repeat all the analysis that was carried out for the classical case, and get Eq. (71) again, but this expression has to be compared not with the equipartition theorem (72), but with its quantum-mechanical generalization (2.78), which, in our current notation, reads

\[
\left\langle \hat{q}^2 \right\rangle = \frac{\hbar \omega_0}{2\kappa} \coth \frac{\hbar \omega_0}{2T}. \tag{5.97}
\]

As a result, we get the following quantum-mechanical generalization of Eq. (92):

\[
S_f(\omega) = \frac{\hbar \chi''(\omega)}{2\pi} \coth \frac{\hbar \omega}{2T}. \tag{5.98}
\]

This is the much-celebrated fluctuation-dissipation theorem, frequently referred to just as FDT.\(^{38}\)

As natural as it seems, this generalization poses a very interesting conceptual dilemma. Let, for the sake of clarity, temperature be relatively low, \( T << \hbar \omega \); then Eq. (98) gives a temperature-independent result

\[
S_f(\omega) = \frac{\hbar \chi''(\omega)}{2\pi}, \tag{5.99}
\]

\(^{37}\) Please note that here (and to the end of this section) brackets \( \langle \ldots \rangle \) mean quantum-statistical averaging (2.12). As was discussed in Sec. 2.1, for a classical-mixture state of the environment, this does not create any difference in either mathematical treatment of the averages or their physical interpretation.

\(^{38}\) It was first derived in 1951 by H. Callen and T. Welton (in a somewhat different way). One more derivation of the FDT, which gives the Kubo formula as a by-product, may be found in QM Sec. 7.4.
which is frequently called the quantum noise. According to the quantum Langevin equation (93), nothing but these fluctuations of the force exerted by the environment, with spectral density proportional to the imaginary part of susceptibility (i.e. damping), are the source of the ground-state “fluctuations” of the coordinate and momentum of a quantum harmonic oscillator, with r.m.s. values

\[
\delta q \equiv \left\langle \left\langle \tilde{q}^2 \right\rangle \right\rangle^{1/2} = \left( \frac{\hbar}{2m\omega_0} \right)^{1/2}, \quad \delta p \equiv \left\langle \left\langle \tilde{p}^2 \right\rangle \right\rangle^{1/2} = m\omega_0\delta q = \left( \frac{\hbar\omega_0}{2} \right)^{1/2}, \quad \text{i.e.} \quad \delta q \cdot \delta p = \frac{\hbar}{2}, \quad (5.100)
\]

and average energy \(\hbar\omega_0/2\). On the other hand, the basic quantum mechanics tells us that exactly these formulas describe the ground state of a dissipation-free oscillator, not coupled to any environment, and are a direct corollary of the Heisenberg uncertainty relation

\[
\delta q \cdot \delta p \geq \frac{\hbar}{2}. \quad (5.101)
\]

(The Gaussian wave packets, pertinent to a harmonic oscillator’ ground state, turn the sign in Eq. (101) into pure equality.) So, what is the genuine source of Eqs. (100)?

The resolution of this paradox is that either interpretation of Eqs. (100) is legitimate, with their relative convenience depending on the particular application. (One can say that since the right-hand part of the quantum Langevin equation (93) is a quantum-mechanical operator, rather than a classical force, it “carries the uncertainty relation within itself”.) However, this opportunistic resolution leaves the following question open: is the quantum noise (99) of the environment observable directly, without any probe oscillator subjected to it? An experimental resolution of this dilemma is not quite simple, because usual scientific instruments have their own zero-point fluctuations, which may be readily confused with those of the system under study. Fortunately, this difficulty may be overcome, for example, using unique frequency-mixing (“down-conversion”) properties of Josephson junctions.\(^{39}\) Special low-temperature experiments using such down-conversion\(^ {40}\) have confirmed that noise (99) is real and measurable. This has been one of the most convincing direct demonstrations of the reality of the zero-point energy \(\hbar\omega_0/2\).\(^ {41}\)

Finally, let me mention briefly an alternative derivation\(^ {42}\) of the fluctuation-theorem from the general quantum mechanics of open systems. This derivation is substantially longer, but gives an interesting sub-product,

\[
\left\langle \left[ \tilde{\mathcal{F}}(t), \tilde{\mathcal{F}}(t+\tau) \right] \right\rangle = i\hbar \mathcal{G}(\tau), \quad (5.102)
\]

where \(\mathcal{G}(\tau)\) is the temporal Green’s function of the environment (as “seen” by the system subjected to the generalized force \(\mathcal{F}\)), defined by equation

\[
\langle \mathcal{F}(t) \rangle = \int_0^\infty \mathcal{G}(\tau)q(t-\tau)d\tau = \int_{-\infty}^t \tilde{g}(t-t')q(t')dt'. \quad (5.103)
\]

---


\(^{41}\) Another one is the Casimir effect - see, e.g., QM Sec. 9.1.

\(^{42}\) See, e.g., QM Sec. 7.4.
Plugging the Fourier transforms of all three functions participating in Eq. (103) into that relation, it is straightforward to check\(^{43}\) that the Green’s function is just the Fourier image of the complex susceptibility \(\chi(\omega)\), defined by Eq. (90):

\[
\int_0^\infty \varphi(\tau) e^{i\omega \tau} d\tau = \chi(\omega) \quad (5.104)
\]

here 0 is used as a lower limit instead of (-\(\infty\)) just to emphasize that due to the causality principle, the Green’s function has to be equal zero for \(\tau < 0\).

In order to reveal the real beauty of Eq. (102), we may use the Wiener-Khinchin theorem (59) to rewrite the fluctuation-dissipation theorem (98) in a form similar to Eq. (102):

\[
\left\langle \left[ \hat{\varphi}(t), \hat{\varphi}(t+\tau) \right] \right\rangle = 2K_\varphi(\tau) \quad (5.105)
\]

where the correlation function \(K_\varphi(\tau)\) is most simply described by its Fourier transform, equal to \(\pi S_\varphi(\omega)\):

\[
\int_0^\infty K_\varphi(\tau) \cos \omega \tau d\tau = \frac{\hbar \chi''(\omega)}{2} \coth \frac{\hbar \omega}{2T} \quad (5.106)
\]

The comparison of Eqs. (102) and (104), on one hand, and Eqs (105)-(106), on the other hand, shows that both the commutation and anticommutation properties of the Heisenberg-Langevin force operator at different moments of time are determined by the same generalized susceptibility \(\chi(\omega)\), but the average anticommutator also depends on temperature, while the average commutator does not.\(^{44}\)

5.6. The Kramers problem and the Smoluchowski equation

Returning to the classical case, it is evident that the Langevin equation (65) provides the means not only for the analysis of stationary fluctuations, but also for the description of an arbitrary time evolution of (classical) dynamic systems coupled to their environment - which, again, provides both dissipation and fluctuations. However, this approach suffers from two major handicaps.

First, this equation does enable us to find the statistical average of variable \(q\), and the variance of its fluctuations (i.e., in the common mathematical terminology, the first and second moments of the probability distribution) as functions of time, but not the distribution \(w(q, t)\) as such. This may not look like a big problem, because in most cases (in particular, in linear systems such as the harmonic oscillator) the distribution is Gaussian – see, e.g., Eq. (2.77).

The second, more painful, drawback of the Langevin approach is that it is instrumental only for the already mentioned “linear” systems - i.e., the systems whose dynamics is described by linear differential equations, such as Eq. (65). However, as we know from classical dynamics, many important problems (for example, the Kepler problem of planetary motion\(^{45}\)) are reduced to 1D motion in substantially anharmonic potentials \(U_{\text{ef}}(q)\), leading to nonlinear equations of motion. If the energy of interaction between the system and its random environment is bilinear – i.e. is a product of variables

---

\(^{43}\) See, e.g., CM Sec. 4.1, part (ii).

\(^{44}\) Only explicitly so, because the complex susceptibility of the environment may depend on temperature as well.

\(^{45}\) See, e.g., CM Sec. 3.4-3.6.
belonging to these sub-systems (as it is very frequently the case), we may repeat all arguments of the last section to derive the following generalized version of the Langevin equation

\[ m \ddot{q} + \eta \dot{q} + \frac{\partial U(q,t)}{\partial q} = \overline{f}(t), \quad (5.107) \]

valid for an arbitrary, possibly time-dependent potential \( U(q,t) \).\(^{46}\) Unfortunately, the solution of this equation may be very hard. Indeed, the Fourier analysis carried out in the last section was essentially based on the linear superposition principle that is invalid for nonlinear equations.

If the fluctuation intensity is low, \( |\delta q| \ll \langle q \rangle \), where \( \langle q \rangle(t) \) is the deterministic solution of Eq. (107) in the absence of fluctuations, this equation may be linearized\(^{47}\) with respect to small fluctuations \( \tilde{q} = q - \langle q \rangle \) to get a linear equation,

\[ m \ddot{\tilde{q}} + \eta \dot{\tilde{q}} + \kappa(t)\tilde{q} = \overline{f}(t), \quad \text{with } \kappa(t) = \frac{\partial^2}{\partial q^2} U(\langle q \rangle(t),t). \quad (5.108) \]

This equation differs from Eq. (65) only by the time dependence of the effective spring constant \( \kappa(t) \), and may be solved by the Fourier expansion of both fluctuations and function \( \kappa(t) \). Such calculations are somewhat more cumbersome than have been performed above, but may be doable (especially if the unperturbed motion \( \langle q \rangle(t) \) is periodic), and sometimes give useful analytical results.\(^{48}\)

However, some important problems cannot be solved by the linearization. Perhaps, the most apparent example is the so-called Kramers problem\(^{49}\) of finding the lifetime of a metastable state of a 1D classical system in a potential well separated from the continuum motion region with a potential barrier (Fig. 10).

In the absence of fluctuations, the system, placed close to well’s bottom \( q = q_1 \), would stay there forever. Fluctuations result not only in a finite spread of the probability density \( w(q,t) \) around that point, but also in the gradual decrease of the total probability

\[ W(t) = \int_{\text{well's bottom}} w(q,t) dq \quad (5.109) \]

\(^{46}\) The generalization of Eq. (107) to higher spatial dimensionality is also straightforward, with the scalar variable \( q \) replaced by vector \( \mathbf{q} \), and the scalar derivative \( dU/dq \) replaced with vector \( \nabla U \).

\(^{47}\) See, e.g., CM Secs. 3.2, 4.2, and beyond.

\(^{48}\) See, e.g., Chapters 5 and 6 in W. Coffey et al., *The Langevin Equation*, World Scientific, 1996.

\(^{49}\) After H. Kramers who, besides solving this important problem in 1940, has made significant contributions to many other areas of physics, including the famous Kramers-Kronig dispersion relations - see, e.g., EM Sec. 7.4.
to find the system in the well, because of the growing probability of escape from the well, over the potential barrier, due to thermal activation. If the barrier height,

$$U_0 = U(q_2) - U(q_1), \quad (5.110)$$

is much larger than temperature $T$, the Boltzmann distribution $w \propto \exp\{-U(q)/T\}$ should be approximately valid in most of the well, so that the probability for the system to overcome the barrier should scale as $\exp\{-U_0/T\}$. From these handwaving arguments, one may reasonably expect that if probability $W(t)$ that the system is still in the well by time $t$ should obey the usual “decay law”

$$\dot{W} = -\frac{W}{\tau}, \quad (5.111)$$

then lifetime $\tau$ has to obey the general Arrhenius law, $\tau = \tau_A \exp\{U_0/T\}$. However, that relation needs to be proved, and the pre-exponential coefficient $\tau_A$ (frequently called the attempt time) needs to be calculated. This cannot be done by the linearization of Eq. (107), because the linearization is equivalent to a quadratic approximation of the potential $U(q)$, which evidently cannot describe the potential well and the potential barrier simultaneously – see Fig. 10.

This and other essentially nonlinear problems may be addressed using an alternative approach to fluctuation analysis, dealing directly with the time evolution of the probability density $w(q,t)$. Due to the shortage of time, I will review this approach a bit superficially, using mostly handwaving arguments, and refer the interested reader to special literature$^{51}$ for strict mathematical proofs. Let us start from the effect of diffusion of a free 1D particle in the high damping limit, described by the Langevin equation (74), and assume that at all times the probability distribution stays Gaussian:

$$w(q,t) = \frac{1}{(2\pi)^{1/2} \delta q(t)} \exp\left\{-\frac{(q-q_0)^2}{2\delta^2(t)}\right\}, \quad (5.112)$$

where $q_0$ is the initial position of the particle, and $\delta q(t)$ is the time-dependent distribution width, which grows in time in accordance with Eq. (77):

$$\delta q(t) = (2Dt)^{1/2}. \quad (5.113)$$

It is straightforward to check, by substitution, that this solution satisfies the following simple partial differential equation,$^{52}$

$$\frac{\partial w}{\partial t} = D \frac{\partial^2 w}{\partial q^2}, \quad (5.114)$$

with the delta-functional initial condition

$$w(q,0) = \delta(q-q_0). \quad (5.115)$$

---

$^{50}$ If $U_0$ is comparable with $T$, system’s behavior also depends substantially on the initial probability distribution, i.e., do not follow the universal law (111).


$^{52}$ By the way, the goal of the traditional coefficient 2 in Eq. (77) is exactly to have the fundamental Eq. (114) free of numerical coefficients.
The simple and important equation of diffusion (114) may be naturally generalized to the 3D motion:\(^{53}\)

\[ \frac{\partial w}{\partial t} = D \nabla^2 w. \]  

(5.116)

Now let us compare this equation with the probability conservation law,\(^ {54}\)

\[ \frac{\partial w}{\partial t} + \nabla \cdot \mathbf{j}_w = 0, \]

(5.117a)

where vector \( \mathbf{j}_w \) has the physical sense of the probability current density. (The validity of this relation is evident from its integral form,

\[ \frac{d}{dt} \int_V w d^3r + \oint_S \mathbf{j}_w \cdot d^2r = 0, \]

(5.117b)

that results from integration of Eq. (117a) over an arbitrary time-independent volume \( V \) limited by surface \( S \), and applying the divergence theorem\(^ {55}\) to the second term.) The continuity relation (117a) coincides with Eq. (116), with \( D \) given by Eq. (78), only if we take

\[ \mathbf{j}_w = -D \nabla w = -\frac{T}{\eta} \nabla w. \]

(5.118)

The first form of this relation allows a simple interpretation: the probability flow is proportional to the spatial gradient of probability density (i.e., in application to many \( N \) similar and independent particles, just to the gradient of their concentration \( n = Nw \)), with the sign corresponding to the flow from the higher to lower concentration. This flow is the very essence of the effect of diffusion.

The fundamental Eq. (117) has to be satisfied also for a force-driven particle at negligible diffusion \( (D \rightarrow 0) \); in this case

\[ \mathbf{j}_w = w \mathbf{v}, \]

(5.119)

where \( \mathbf{v} \) is the deterministic velocity of the particle. In the high-damping limit we are considering right now, \( \mathbf{v} \) is just the drift velocity:

\[ \mathbf{v} = \frac{1}{\eta} \mathbf{F}_{\text{det}} = -\frac{1}{\eta} \nabla U(\mathbf{r}), \]

(5.120)

where \( \mathbf{F}_{\text{det}} \) is the deterministic force described by potential energy \( U(\mathbf{r}) \). Now, as we have descriptions of \( \mathbf{j}_w \) due to both drift and diffusion separately, we may rationally assume that in the general case when both effects are present, the corresponding components of the probability current just add up, so that

\[ \mathbf{j}_w = \frac{1}{\eta} \left[ w(-\nabla U) - T \nabla w \right], \]

(5.121)

\(^{53}\) As will be discussed in Chapter 6, the equation of diffusion also describes several other physical phenomena – in particular, heat propagation in a uniform, isotropic solid, and in this context is called the heat conduction equation or (rather inappropriately) just the “heat equation”.

\(^{54}\) Both forms of Eq. (117) are similar to the mass conservation law in classical dynamics (see, e.g., CM Sec. 8.2), and the electric charge conservation law in electrodynamics (see, e.g., EM Sec. 4.1).

\(^{55}\) See, e.g., MA Eq. (12.2),
and Eq. (117a) takes the form

\[
\eta \frac{\partial w}{\partial t} = \nabla (w \nabla U) + T \nabla^2 w.
\]  

(5.122)

This is the Smoluchowski equation,\textsuperscript{56} which is closely related to the Boltzmann equation in multi-particle kinetics - to be discussed in the next chapter.

As a sanity check, let us see what does the Smoluchowski equation give in the stationary limit, \( \partial w/\partial t \to 0 \) (which evidently may be achieved only if the deterministic potential \( U \) is time-independent.) Then Eq. (117a) yields \( j_w = \text{const} \), where the constant describes the motion of the system as the whole. If such motion is absent, \( j_w = 0 \), then according to Eq. (121),

\[
w \nabla U + T \nabla w = 0, \quad \text{i.e.} \quad \frac{\nabla w}{w} = -\frac{\nabla U}{T}.
\]

(5.123)

Since the left-hand part of the last form of the last relation is just \( \nabla (\ln w) \), Eq. (123) may be immediately integrated, giving

\[
\ln w = \frac{-U}{T} + \ln C, \quad \text{i.e.} \quad w(r) = C \exp \left\{ -\frac{U(r)}{T} \right\}, \quad (5.124)
\]

Multiplied by the number \( N \) of similar, independent systems, with spatial density \( n(r) = N w(r) \), this is just the Boltzmann distribution (3.26).

Now, as a less trivial example of the Smoluchowski equation’s applications, let us use it to solve the 1D Kramers problem (Fig. 10) in the corresponding high-damping limit, \( m \ll \eta \tau_L \). It is straightforward to check that the 1D version of Eq. (121),

\[
I_w = \frac{1}{\eta} \left[ w \left( -\frac{\partial U}{\partial q} \right) - T \frac{\partial w}{\partial q} \right],
\]

(5.125a)

is equivalent to

\[
I_w = -\frac{T}{\eta} \exp \left\{ -\frac{U(q)}{T} \right\} \frac{\partial}{\partial q} \left( w \exp \left\{ \frac{U(q)}{T} \right\} \right),
\]

(5.125b)

(where \( I_w \) is the probability current at a certain location \( q \), rather than its density), so that we can write

\[
I_w \exp \left\{ \frac{U(q)}{T} \right\} = -\frac{T}{\eta} \frac{\partial}{\partial q} \left( w \exp \left\{ \frac{U(q)}{T} \right\} \right).
\]

(5.126)

As was discussed above, the notion of metastable state’s lifetime is well defined only for sufficiently low temperatures

\[
T \ll U_0.
\]

(5.127)

\textsuperscript{56} Named after M. Smoluchowski who developed this formalism in 1906, apparently independently from the slightly earlier Einstein’s work, and in much more detail. This equation has important applications in many fields of science, including such surprising topics as statistics of spikes in neural networks.
when the lifetime is relatively long, \( \tau >> \tau_d \), where \( \tau_d \) has to be of the order of the time of the system relaxation inside the well. Since the first term of the continuity equation (117b) is of the order of \( W/\tau \), in this limit the term, and hence the gradient of \( I_w \), are negligibly small, so the probability current does not depend on \( q \) in the potential barrier region. Let us integrate both sides of Eq. (126) over that region, using that fact:

\[
I_w \int_{q'}^{q''} \exp \left\{ \frac{U(q)}{T} \right\} dq = -\frac{T}{\eta} \left\{ w \exp \left\{ \frac{U(q)}{T} \right\} \right\} q'', \tag{5.128}
\]

where the integration limits \( q' \) and \( q'' \) (Fig. 10) are selected so that so that

\[
T << U(q') - U(q_1), U(q_2) - U(q'') << U_0. \tag{5.129}
\]

(Evidently, such selection is only possible if condition (127) is satisfied.) In this limit the contribution to the right-hand part from point \( q'' \) is negligible because the probability density behind the barrier is exponentially small. On the other hand, the probability at point \( q' \) is close to its stationary, Boltzmann value (124), so that

\[
w(q') \exp \left\{ \frac{U(q')}{T} \right\} = w(q_1) \exp \left\{ \frac{U(q_1)}{T} \right\}, \tag{5.130}
\]

and Eq. (128) yields

\[
I_w = \frac{T}{\eta} \left[ w(q_1) / \int_{q'}^{q''} \exp \left\{ \frac{U(q) - U(q_1)}{T} \right\} dq \right]. \tag{5.131}
\]

We are almost done. The probability density \( w(q_1) \) at the well’s bottom may be expressed in terms of the total probability \( W \) of the particle being in the well by using the normalization condition

\[
W = \int_{\text{well's bottom}} w(q_1) \exp \left\{ \frac{U(q_1) - U(q)}{T} \right\} dq; \tag{5.132}
\]

the integration here may be limited by the region where the difference \( U(q) - U(q_1) \) is larger then \( T \) but still much smaller than \( U_0 \) - cf. Eq. (129). According to the Taylor expansion, the shape of any smooth potential well near its bottom may be well approximated by a quadratic parabola:

\[
U(q \approx q_1) - U(q_1) \approx \frac{\kappa_1}{2} (q - q_1)^2, \quad \text{where} \quad \kappa_1 \equiv \frac{d^2U}{dq^2} \bigg|_{q=q_1} > 0. \tag{5.133}
\]

With this approximation, Eq. (132) is reduced to the standard Gaussian integral:

\[
W = w(q_1) \int_{\text{well's bottom}} \exp \left\{ -\frac{\kappa_1 (q - q_1)^2}{2T} \right\} dq \approx w(q_1) \int_{-\infty}^{\infty} \exp \left\{ \frac{\kappa_1 \tilde{q}^2}{2T} \right\} d\tilde{q} = w(q_1) \left( \frac{2\pi T}{\kappa_1} \right)^{1/2}. \tag{5.134}
\]

To complete the calculation, we may use the similar approximation,

---

57 If necessary, see MA Eq. (6.9b) again.
\[ U(q \approx q_2) - U(q_1) \approx \left[ U(q_2) - \frac{\kappa_2}{2} (q - q_2)^2 \right] - U(q_1) = U_0 - \frac{\kappa_2}{2} (q - q_2)^2, \]  
\begin{equation}
\tag{5.135}
\end{equation}

where \( \kappa_2 \equiv -\frac{d^2 U}{dq^2} \bigg|_{q=q_2} > 0, \)

to work out the remaining integral in Eq. (131), because in the limit (129) this integral is dominated by the contribution from a region very close to the barrier top, where approximation (135) is asymptotically exact. As a result, we get

\[ \int_{q'}^{q''} \exp\left\{ \frac{U(q) - U(q_1)}{T} \right\} dq \approx \exp\left\{ \frac{U_0}{T} \right\} \left( \frac{2\pi T}{\kappa_2} \right)^{1/2}. \]  
\begin{equation}
\tag{5.136}
\end{equation}

Plugging Eqs. (136), and \( w(q_1) \) expressed from Eq. (134), into Eq. (131), we finally get

\[ I_w = W \left( \frac{\kappa_1 \kappa_2}{2\pi \eta} \right)^{1/2} \exp\left\{ -\frac{U_0}{T} \right\}. \]  
\begin{equation}
\tag{5.137}
\end{equation}

This expression should be compared with the 1D version of Eq. (117b) for the segment \([-\infty, q']\]. Since this interval covers the region near \( q_1 \) where most of the probability density resides, and \( I_q(\infty) = 0 \), the result is merely

\[ \frac{dW}{dt} + I_w(q') = 0. \]  
\begin{equation}
\tag{5.138}
\end{equation}

In our approximation, \( I_w(q') \) does not depend on the exact position of point \( q' \), and is given by Eq. (137), so that plugging it into Eq. (138), we recover the exponential decay law (111), with lifetime

\[ \tau = \frac{2\pi \eta}{(\kappa_1 \kappa_2)^{1/2}} \exp\left\{ \frac{U_0}{T} \right\} = 2\pi \left( \tau_1 \tau_2 \right)^{1/2} \exp\left\{ \frac{U_0}{T} \right\}, \]  
\begin{equation}
\tag{5.139}
\end{equation}

where \( \tau_{1,2} \equiv \frac{\eta}{\kappa_{1,2}} \).

Thus the metastable state lifetime is indeed described by the Arrhenius law, with the attempt time scaling as the geometric mean of system’s “relaxation times” near the potential well bottom (\( \tau_1 \)) and the potential barrier top (\( \tau_2 \)).\textsuperscript{58} Let me leave for reader’s exercise to prove that if the potential profile near well’s bottom and/or top is sharp, the pre-exponential factor in Eq. (139) should be modified, but the Arrhenius exponent is not affected.

### 5.7. The Fokker-Planck equation

Expression (139) is just a particular, high-damping limit of a more general result obtained by Kramers. In order to recover all of it, we need to generalize the Smoluchowski equation to arbitrary values of damping \( \eta \). In this case, the probability density \( w \) is a function of not only the particle’s position \( q \) (and time \( t \)), but also its momentum \( p \) – see Eq. (2.11). Thus the continuity equation (117a) needs to be generalized to 6D phase space. Such generalization is natural:

\textsuperscript{58} Actually, \( \tau_2 \) describes the characteristic time of the exponential growth of small deviations from the unstable fixed point \( q_2 \) at the barrier top, rather than their decay, as near point \( q_1 \).
\[
\frac{\partial W}{\partial t} + \nabla_q \cdot \mathbf{j}_q + \nabla_p \cdot \mathbf{j}_p = 0, \tag{5.140}
\]

where \(\mathbf{j}_q\) (which was called \(\mathbf{j}_w\) in the last section) is the probability current density in the coordinate space, while \(\mathbf{j}_p\) is the current density in the momentum space, and \(\nabla_p\) is the gradient operator in that space,

\[
\nabla_p \equiv \sum_{j=1}^{3} n_j \frac{\partial}{\partial p_j}, \tag{5.141}
\]

while \(\nabla_q\) is the usual gradient operator in the coordinate space, that was denoted as \(\nabla\) in the previous section - with index \(q\) added here just for additional clarity. At negligible fluctuations \(T \to 0\), \(\mathbf{j}_p\) in the momentum space may be evaluated using the natural analogy with \(\mathbf{j}_q\) – see Eq. (119). In our new notation, that relation takes the following form,

\[
\mathbf{j}_q = w\mathbf{v} = w\dot{\mathbf{q}} = \frac{\mathbf{p}}{m}, \tag{5.142}
\]

so it is naturally to take

\[
\mathbf{j}_p = w\dot{\mathbf{p}} = w\langle \mathbf{F} \rangle = w(-\nabla_q U - \eta \mathbf{v}) = w(-\nabla_q U - \eta \frac{\mathbf{p}}{m}). \tag{5.143}
\]

As a sanity check, it is straightforward to verify that the diffusion-free equation resulting from the combination of Eqs. (140), (142) and (143),

\[
\frac{\partial w}{\partial t}\bigg|_{\text{drift}} = -\nabla_q \cdot \left( w\frac{\mathbf{p}}{m} \right) + \nabla_p \cdot \left[ w(-\nabla_q U + \eta \frac{\mathbf{p}}{m}) \right], \tag{5.144}
\]

allows the following particular solution

\[
w(\mathbf{q}, \mathbf{p}, t) = \delta(\mathbf{q} - \langle \mathbf{q} \rangle(t))\delta(\mathbf{p} - \langle \mathbf{p} \rangle(t)), \tag{5.145}
\]

where the statistical-average coordinate and momentum satisfy the deterministic equations of motion,

\[
\langle \dot{\mathbf{q}} \rangle = \frac{\langle \mathbf{p} \rangle}{m}, \quad \langle \dot{\mathbf{p}} \rangle = -\nabla_q U - \eta \frac{\langle \mathbf{p} \rangle}{m}, \tag{5.146}
\]

describing particle’s drift, with the appropriate deterministic initial conditions.

In order to understand how the diffusion may be accounted for, let us consider a statistical ensemble of free \((\nabla_q U = 0, \eta \to 0)\) particles that are uniformly distributed in direct space (so that \(\nabla_q w = 0\), but possibly localized in the momentum space. For this case, the right-hand part of Eq. (144) vanishes, i.e. the time evolution of the probability density \(w\) may be only due to diffusion. In the corresponding limit \(\langle \mathbf{F} \rangle \to 0\), the Langevin equation (107) for each Cartesian coordinate is reduced to

\[
m\ddot{q}_j = \tilde{\mathbf{F}}_j(t), \quad \text{i.e.} \quad \dot{p}_j = \tilde{\mathbf{F}}_j(t). \tag{5.147}
\]

This equation is similar to the high-damping 1D equation (74) (with \(\mathbf{F}_{\text{det}} = 0\), with replacement \(q \to \mathbf{p}/\eta\), and hence the corresponding contribution to \(\partial w/\partial t\) may be described by the second term of Eq. (122) with that replacement:
Now the reasonable assumption that in the arbitrary case the drift and diffusion contributions to $\frac{\partial w}{\partial t}$ just add up, immediately leads us to the full Fokker-Planck equation:

$$\frac{\partial w}{\partial t} = -\nabla_q \cdot \left( w \frac{p}{m} \right) + \nabla_p \cdot \left[ w \left( \nabla_q U + \eta \frac{p}{m} \right) \right] + \eta T \nabla^2_p w. \quad (5.149)$$

As a sanity check, let us use this equation to find the stationary probability distribution of momentum of free particles, at arbitrary damping $\eta$, in the momentum space, assuming their uniform distribution in the direct space, $\nabla_q = 0$. In the stationary case $\frac{\partial w}{\partial t} = 0$, so that Eq. (149) is reduced to

$$\nabla_p \cdot \left[ w \left( \eta \frac{p}{m} \right) \right] + \eta T \nabla^2_p w = 0. \quad (5.150)$$

The damping coefficient $\eta$ cancels, and the first integration over momentum yields

$$\frac{p}{m} w + T \nabla_p w = j, \quad (5.152)$$

where $j$ is a vector constant describing a possible motion of the system as the whole. In the absence of such motion, $j = 0$, the second integration over momentum gives

$$w = \text{const} \times \exp \left\{ -\frac{p^2}{2mT} \right\}, \quad (5.153)$$

i.e. the Maxwell distribution (3.5). However, result (153) is more general than that obtained in Sec. 3.1, because it shows that the distribution stays the same even at nonvanishing damping.

It is also easy to show that if the damping is large (in the sense assumed in the last section), the solution of the Fokker-Plank equation tends to the following product

$$w(q,p,t) \to \text{const} \times \exp \left\{ -\frac{p^2}{2mT} \right\} \times w(q,t), \quad (5.154)$$

where the direct-space distribution $w(q,t)$ obeys the Smoluchowski equation (122). However, in the general case, solutions of Eq. (149) may be rather complex, so I would mention (rather than derive) only one of them, that of the Kramers problem (Fig. 10). Acting virtually exactly as in Sec. 6, one can show at arbitrary damping (but still in the limit (127), $T << U_0$, with the additional restriction $\tau >> m/\gamma$), the metastable state’s lifetime is again given by the Arrhenius formula (139), with the same exponent $\exp\{U_0/T\}$, but with the reciprocal time constants $1/\tau_{1,2}$ replaced with

$$\omega_{1,2}^\prime \equiv \left[ \omega_{1,2}^2 + \left( \frac{\eta}{2m} \right)^2 \right]^{1/2} - \frac{\eta}{2m} \to \begin{cases} \omega_{1,2}, & \text{for } \eta << m\omega_{1,2}, \\ 1/\tau_{1,2}, & \text{for } m\omega_{1,2} << \eta, \end{cases} \quad (5.155)$$

59 It was derived in 1913 in A. Fokker’s PhD thesis work; M. Planck was his thesis adviser.

60 The reader should remember that these solutions embody, as the particular case $T = 0$, all classical dynamics of a particle.
where \( \omega_{1,2} \equiv (\kappa_{1,2}/m)^{1/2} \), while \( \kappa_{1,2} \) are the effective spring constants defined by Eqs. (133) and (135). Thus, in the most important particular limit of low damping, Eq. (139) is replaced with the famous formula

\[
\tau = \frac{2\pi}{(\omega_1\omega_2)^{1/2}} \exp\left[\frac{U_0}{T}\right].
\]

(5.156)

This Kramers’ result for the classical thermal activation of the virtually-Hamiltonian system over the potential barrier may be compared with that for its quantum-mechanical tunneling through the barrier.\(^61\) Even the simplest, WKB approximation for the latter time,

\[
\tau_Q = \tau_A \exp\left\{-2 \int_{\kappa^2(q)>0} \kappa(q) dq\right\}, \quad \text{with} \quad \frac{\hbar^2 \kappa^2(q)}{2m} \equiv U(q) - E,
\]

(5.157)

shows that generally those two lifetimes have different dependences on the barrier shape. For example, for a nearly-rectangular potential barrier, the exponent that determines the classical lifetime (156) depends (linearly) only on the barrier height \( U_0 \), while that defining the quantum lifetime is proportional to the barrier width, while scaling as a square root of \( U_0 \). However, in the important case of “soft” potential profiles, which are typical for the case of barely emerging (or nearly disappearing) quantum wells (Fig. 11) the classical and quantum results may be simply related.

Indeed, such potential profile \( U(q) \) may be well approximated by 4 leading terms of its Taylor expansion, with the highest term proportional to \((q - q_0)^3\), near some point \( q_0 \) in the vicinity of the well. In this approximation, the second derivative \( d^2U/dq^2 \) vanishes at the point \( q_0 = (q_1 + q_2)/2 \), exactly between the well’s bottom and the barrier’s top (in Fig. 11, \( q_1 \) and \( q_2 \)). Selecting the origin at this point, we may reduce the approximation to just two terms:\(^62\)

\[
U(q) = aq - \frac{b}{3} q^3,
\]

(5.158)

with \( ab > 0 \). Using a straightforward calculus, we can find all important parameters of this cubic-parabola: the positions of its minimum and maximum:

\[
q_2 = -q_1 = \left(\frac{a}{b}\right)^{1/2},
\]

(5.159)

the barrier height over the well’s bottom:

\(^61\) See, e.g., QM Secs. 2.3-2.4.

\(^62\) As a reminder, an absolutely similar approximation is used in Exercise Problem 4.3 for the \( P(V) \) function, in order to analyze properties of the van der Waals model near the critical temperature.
\[ U_0 \equiv U(q_2) - U(q_1) = \frac{4}{3} \left( \frac{a^3}{b} \right)^{1/2}, \quad (5.160) \]

and the effective spring constants:

\[ \kappa_1 = \kappa_2 \equiv \left. \frac{d^2U}{dq^2} \right|_{q_{1,2}} = 2(ab)^{1/2}. \quad (5.161) \]

The last expression shows that for this potential profile, frequencies \( \omega_{1,2} \) participating in Eq. (161) are equal to each other, so that this result may be rewritten as

\[ \tau = \frac{2\pi}{\omega_0} \exp\left\{ \frac{U_0}{T} \right\}, \quad \text{with} \quad \omega_0^2 = \frac{2(ab)^{1/2}}{m}. \quad (5.162) \]

On the other hand, for the same profile, the WKB approximation (157) (which is accurate when the height of the metastable state energy over the well’s bottom, \( E - U(q_1) \approx \hbar \omega_0/2 \), is much less than the barrier height \( U_0 \)) yields\(^{63}\)

\[ \tau_0 = \frac{2\pi}{\omega_0} \left( \frac{\hbar \omega_0}{864\pi U_0} \right)^{1/2} \exp\left\{ \frac{36}{5} \frac{U_0}{\hbar \omega_0} \right\}. \quad (5.163) \]

Comparison of the dominating, exponential factors in these two results shows that the thermal activation yields lower lifetime (i.e., dominates the metastable state decay) if temperature is above the crossover value

\[ T_c = \frac{36}{5} \hbar \omega_0 = 7.2 \hbar \omega_0. \quad (5.164) \]

This expression for the cubic-parabolic barrier may be compared with the similar crossover for a quadratic-parabolic barrier,\(^ {64}\) for which \( T_c = 2\pi \hbar \omega_0 \approx 6.28 \hbar \omega_0 \). We see that the numerical factors for these two different soft potential profiles are very substantial, but rather close.

5.8. Back to the correlation function

Unfortunately I will not have time to review solutions of other problems using the Smoluchowski and Fokker-Planck equations, but have to mention one conceptual issue. Since it is intuitively clear that these equations provide the complete statistical information about the system under analysis, one may wonder whether they may be used to find the temporal characteristics of the system, which were discussed in Secs. 4-5 using the Langevin formalism. For any statistical average of a function taken at the same time instant, the answer is evidently yes – cf. Eq. (2.11):

\[ \langle f(q(t),p(t)) \rangle = \int f(q,p)w(q,p,t)d^3qd^3p, \quad (5.165) \]

\(^{63}\) The main, exponential factor in this result may be obtained simply by ignoring the difference between \( E \) and \( U(q_1) \), but the correct calculation of the pre-exponent requires to take this difference, \( \hbar \omega_0/2 \), into account – see K. Likharev, *Physica B* **108**, 1079 (1981).

\(^ {64}\) See, e.g., QM Sec. 2.4.
but what if the function depends on variables taken at different times, for example the components of the correlation function $K_f(\tau)$ defined by Eq. (49)?

To answer this question, let us start from the discrete variable case when Eq. (165) takes form (2.7), which, for our current purposes, may be rewritten as

$$\langle f(t) \rangle = \sum_m f_m W_m(t).$$  \hspace{1cm} (5.166)

In plain English, this is a sum of all possible values of the function, each multiplied by its probability as a function of time. But this means that average $\langle f(t) f(t') \rangle$ may be calculated as the sum of all possible products $f_m f_{m'}$, multiplied by the joint probability for measurement outcome $m$ at moment $t$, and outcome $m'$ at moment $t'$. The joint probability may be presented as a product of $W_m(t)$ by the conditional probability $W(m', t' | m, t)$. Since the correlation function is well defined only for stationary systems, in the last expression we can take $t = 0$, i.e. find the conditional probability as the result, $W_{m'}(\tau)$, of solution of the equation describing system’s probability evolution, at time $\tau = t' - t$ (rather than $t'$), with the special initial condition

$$W_{m'}(0) = \delta_{m', m},$$  \hspace{1cm} (5.167)

On the other hand, since the average $\langle f(t) f(t + \tau) \rangle$ of a stationary process should not depend on $t$, instead of $W_m(t)$ we may take the stationary probability distribution $W_m(\infty)$, independent of the initial conditions, and may be found as the same special solution, but at time $\tau \to \infty$. As a result, we may write

$$\langle f(t) f(t + \tau) \rangle = \sum_{m,m'} f_m W_m(\infty) f_{m'} W_{m'}(\tau).$$  \hspace{1cm} (5.168)

This expression looks simple, but note that this recipe requires to solve the time evolution equations for each $W_{m'}(\tau)$ for all possible initial conditions (167). To see how this recipe works in practice, let us revisit the simplest two-level system (see, e.g., Fig. 4.13 reproduced in Fig. 12 below in a notation more convenient for our current purposes), and calculate the correlation function of its energy fluctuations.

![Correlation function of discrete-state system](image_url)

The stationary probabilities for this system (i.e. the probabilities for $\tau \to \infty$) have been calculated in Chapter 2, and then again in Sec. 4.4. In our current notation (Fig. 12),

$$W_0(\infty) = \frac{1}{1 + e^{-\Delta/T}}, \quad W_1(\infty) = \frac{1}{e^{\Delta/T} + 1}, \quad \langle E \rangle = W_0(\infty) \times 0 + W_1(\infty) \times \Delta = \frac{\Delta}{e^{\Delta/T} + 1}. \hspace{1cm} (5.169)$$

In order to calculate the conditional probabilities $W_{m'}(\tau)$ with initial conditions (172) (according to Eq. (168), we need all 4 of them, for $m, m' = 0, 1$), we may use master equations (4.100), in our current notation reading
\[
\frac{dW_1}{d\tau} = -\frac{dW_0}{d\tau} = \Gamma_\uparrow W_0 - \Gamma_\downarrow W_1, \tag{5.170}
\]

Since Eq. (170) conserves the total probability, \(W_0 + W_1 = 1\), only one probability (say, \(W_1\)) is an independent variable, and for it, Eq. (170) gives a simple, linear differential equation,
\[
\frac{dW_1}{d\tau} = \Gamma_\uparrow - \Gamma_\downarrow W_1, \quad \text{where} \quad \Gamma_\downarrow = \Gamma_\uparrow + \Gamma_\downarrow. \tag{5.171}
\]

This equation may be readily integrated for an arbitrary initial condition:
\[
W_1(\tau) = W_1(0)e^{-\Gamma_\downarrow \tau} + W_1(\infty)(1 - e^{-\Gamma_\downarrow \tau}), \tag{5.172}
\]

where \(W_1(\infty)\) is given by the second of Eqs. (169). (It is straightforward to check that the solution for \(W_0(\tau)\) may be presented in the similar form, with the corresponding change of the state index.) Now everything is ready to calculate average \(\langle E(t)E(t+\tau) \rangle\) using Eq. (168), with \(f_{m,m'} = E_{0,1}\). Thanks to our (smart :-) choice of energy origin, of 4 terms in the double sum (168), all 3 terms that include at least one factor \(E_0 = 0\) vanish, and we have only one term left:
\[
\langle E(t)E(t+\tau) \rangle = E_1W_1(\infty)E_1W_1(\tau)\bigg|_{W_1(0)=1} = E_1^2W_1(\infty)\left[W_1(0)e^{-\Gamma_\downarrow \tau} + W_1(\infty)(1 - e^{-\Gamma_\downarrow \tau})\right]_{W_1(0)=1}
\]
\[
= \frac{\Delta^2}{e^{\Delta/T} + 1} \left[ e^{-\Gamma_\downarrow \tau} + \frac{1}{e^{\Delta/T} + 1} \left(1 - e^{-\Gamma_\downarrow \tau}\right) \right] = \frac{\Delta^2}{e^{\Delta/T} + 1} \left(1 + e^{\Delta/T} e^{-\Gamma_\downarrow \tau}\right). \tag{5.173}
\]

From here and the last of Eqs. (169), the correlation function of energy fluctuations is\(^{65}\)
\[
\begin{align*}
K_E(\tau) &\equiv \langle \tilde{E}(t)\tilde{E}(t+\tau) \rangle = \langle \left(E(t) - \langle E(t) \rangle\right)\left(E(t+\tau) - \langle E(t+\tau) \rangle\right) \rangle = \langle E(t)E(t+\tau) \rangle - \langle E(t) \rangle\langle E(t+\tau) \rangle \\
&= \langle E(t)E(t+\tau) \rangle - \langle E \rangle^2 = \frac{\Delta^2}{e^{\Delta/T} + 1} e^{-\Gamma_\downarrow \tau}. \tag{5.174}
\end{align*}
\]

Since transition rates \(\Gamma_\uparrow\) and \(\Gamma_\downarrow\) have to obey the detailed balance relation (4.103), \(\Gamma_\downarrow/\Gamma_\uparrow = \exp\{\Delta/T\}\), and hence
\[
\frac{e^{\Delta/T}}{e^{\Delta/T} + 1} = \frac{\Gamma_\downarrow/\Gamma_\uparrow}{(\Gamma_\downarrow/\Gamma_\uparrow + 1)^2} = \frac{\Gamma_\uparrow\Gamma_\downarrow}{(\Gamma_\uparrow + \Gamma_\downarrow)^2} = \frac{\Gamma_\uparrow\Gamma_\downarrow}{\Gamma_\downarrow^2}, \tag{5.175}
\]
expression (174) may be presented also in a simpler form:
\[
K_E(\tau) = \Delta^2 \frac{\Gamma_\uparrow\Gamma_\downarrow}{\Gamma_\downarrow^2} e^{-\Gamma_\downarrow \tau}. \tag{5.176}
\]

We see that the correlation function of energy decays exponentially with time, with the net rate \(\Gamma_\downarrow\), while its variance, equal to \(K_E(0)\), does not depend on the transition rates. Now using the Wiener-Khinchin theorem (58) to calculate its spectral density, we get

\(^{65}\) The transition from the first line of Eq. (174) to its second one uses the fact that the system is stationary, so that \(\langle E(t + \tau) \rangle = \langle E(t) \rangle = \langle E \rangle = \text{const.}\)
\[
S_E(\omega) = \frac{1}{\pi} \int_0^\infty \frac{\Gamma_\uparrow \Gamma_\downarrow}{\Gamma_\Sigma^2} e^{-\Gamma_\Sigma \tau} \cos \omega \tau \, d\tau = \frac{\Delta^2}{\pi \Gamma_\Sigma} \frac{\Gamma_\uparrow \Gamma_\downarrow}{\Gamma_\Sigma^2 + \omega^2}. \tag{5.177}
\]

Such dependence on frequency\textsuperscript{66} is very typical for discrete-state systems described by master equations. It is interesting that the most widely accepted explanation of the $1/f$ noise (also called the “flicker” or “excess” noise), which was mentioned in Sec. 5, is that it is a result of thermally-activated jumps between metastable states of a statistical ensemble of such two-level systems, with an exponentially-broad statistical distribution of transition rates $\Gamma_\uparrow, \downarrow$. Such a broad distribution follows from the Kramers formula (156), which is approximately valid for lifetimes of states of systems with double-well potential profiles (Fig. 13), for a statistical ensemble with a smooth statistical distribution of energy gaps $\Delta$. Such profiles are typical, in particular, for electrons in disordered (amorphous) solid-state materials that, indeed, feature high $1/f$ noise.

![Fig. 5.13. Typical double-well potential profile.](image)

Returning to the Fokker-Planck equation, we may use the evident generalization of Eq. (168) to the continuous-variable case:

\[
\int \delta(q' - q) \delta(p' - p) \, dq' \, dp' \right) = \int d^3q d^3p' f(q,p) w(q,\infty) f(q',p') w(q',p',\tau), \tag{5.178}
\]

were both probability density distributions are solutions of the equation with the delta-functional initial condition

\[
w(q',0) = \delta(q' - q) \delta(p' - p). \tag{5.179}
\]

For the Smoluchowski equation, valid in the high-damping limit, the expressions are similar, albeit with a lower dimensionality:

\[
\int \delta(q' - q) \, dq' \right) = \int d^3q f(q) w(q,\infty) f(q') w(q',\tau), \tag{5.180}
\]

\[
w(q',0) = \delta(q' - q). \tag{5.181}
\]

To see this formalism in action, let us use it to find the correlation function $K_q(\tau)$ of a linear relaxator, i.e. an overdamped 1D harmonic oscillator with $m \omega_0 \ll \eta$. In this limit, the coordinate averaged over the heat baths obeys a linear equation,

\[
\eta \langle \dot{q} \rangle + \kappa \langle q \rangle = 0, \tag{5.182}
\]

which describes its exponential relaxation from a certain initial condition $q_0$ to the equilibrium position $q = 0$, with the reciprocal time constant $\Gamma = \kappa / \eta$.

\textsuperscript{66} Regardless of the physical sense of such function of $\omega$, and of whether its maximum is situated at either zero as in Eq. (177), or at a finite frequency $\omega_0$ as in Eq. (68), it is often referred to as the Lorentzian (or “Breit-Wigner”) line.
The deterministic equation (182) corresponds to the quadratic potential energy \( U(q) = \kappa q^2/2 \), so that the 1D version of the Smoluchowski equation (122) takes the following form:

\[
\eta \frac{\partial w}{\partial t} = \kappa \frac{\partial}{\partial q} (wq) + T \frac{\partial^2 w}{\partial q^2}.
\]  

(5.184)

It is straightforward to check, by substitution, that this equation, rewritten for function \( w(q', \tau) \), with the delta-functional initial condition (181), \( w(q', 0) = \delta(q' - q) \), is satisfied by a Gaussian function,

\[
w(q', \tau) = \frac{1}{(2\pi)^{1/2} \delta q(\tau)} \exp \left\{ -\frac{(q' - \langle q \rangle(\tau))^2}{2\delta q^2(\tau)} \right\},
\]

(5.185)

with its center, \( \langle q \rangle(\tau) \), moving in accordance with Eq. (183), and the time-dependent variance

\[
\delta q^2(\tau) = \delta q^2(\infty) \left[ 1 - e^{-2\Gamma \tau} \right], \quad \text{where} \quad \delta q^2(\infty) = \langle q^2 \rangle = \frac{T}{\kappa}.
\]

(5.186)

(As a sanity check, the last equality coincides with the equipartition theorem’s result.) Finally, the first probability under the integral in Eq. (180) may be found from Eq. (185) in the limit \( \tau \to \infty \) (in which \( \langle q \rangle(\tau) \to 0 \)), by replacing \( q' \) for \( q \):

\[
w(q, \infty) = \frac{1}{(2\pi)^{1/2} \delta q(\infty)} \exp \left\{ -\frac{q^2}{2\delta q^2(\infty)} \right\}.
\]

(5.187)

The integral over \( q' \) may be worked our first, by the replacing that integration variable with \( (q'' + qe^{-\Gamma \tau}) \) and hence \( dq' \) with \( dq'' \):

\[
\langle q(t)q(t+\tau) \rangle = \frac{1}{2\pi \delta q(\tau) \delta q(\infty)} \int_{-\infty}^{\infty} dq \int_{-\infty}^{\infty} dq' q \exp \left\{ -\frac{q^2}{2\delta q^2(\infty)} \right\} q' \exp \left\{ -\frac{(q' - qe^{-\Gamma \tau})^2}{2\delta q^2(\tau)} \right\}.
\]

(5.188)

The integral of the first term in parentheses \( (q'' + qe^{\Gamma \tau}) \) equals zero (as that of an odd function in symmetric integration limits), while that with the second term is the standard Gaussian integral, giving

\[
\langle q(t)q(t+\tau) \rangle = \frac{1}{(2\pi)^{1/2} \delta q(\infty)} e^{-\Gamma \tau} \int_{-\infty}^{\infty} q^2 \exp \left\{ -\frac{q^2}{2\delta q^2(\infty)} \right\} dq = \frac{2T}{\pi^{1/2} \kappa} e^{-\Gamma \tau} \int_{-\infty}^{\infty} \xi^2 \exp \left\{ -\xi^2 \right\} d\xi.
\]

(5.190)

The last integral\(^{67}\) is just \( \pi^{1/2}/2 \), so that taking into account that for this stationary system centered at the coordinate origin, the ensemble average \( \langle q \rangle = 0 \),\(^{68}\) we finally get a very simple result,

\[\langle q(t)q(t+\tau) \rangle = \frac{2T}{\pi^{1/2} \kappa} e^{-\Gamma \tau} \int_{-\infty}^{\infty} \xi^2 \exp \left\{ -\xi^2 \right\} d\xi.\]

(5.190)

\(^{67}\) See, e.g., MA Eq. (6.9c).
As a sanity check, for $\tau = 0$ it yields $K_q(0) \equiv \langle q^2 \rangle = T/\kappa$, in accordance with Eq. (186). As $\tau$ is increased the correlation function decreases monotonically – see the solid-line sketch in Fig. 8.

So, the solution of this very simple problem has required straightforward but somewhat bulky calculations. On the other hand, the same result may be obtained literally in one line, using the Langevin formalism - namely, as the Fourier transform (59) of the spectral density (68) in the corresponding limit $m\omega << \eta$, with $S_\omega(\omega)$ given by Eq. (73):^69

$$K_q(\tau) = 2 \int_0^\infty S_q(\omega) \cos \omega \tau \, d\omega = 2 \int_0^\infty \frac{\eta T}{\pi(\kappa^2 + (\eta \omega)^2)} \cos \omega \tau \, d\omega = 2 \frac{T}{\pi} \int_0^\infty \frac{\cos \xi}{(\Gamma \tau)^2 + \xi^2} \, d\xi = \frac{T}{\kappa} e^{-\frac{\Gamma \tau}{\kappa}}. \tag{5.192}$$

This example illustrates well that for linear systems (and small fluctuations in nonlinear systems) the Langevin approach is usually much simpler that the one based on the Fokker-Planck or Smoluchowski equations. However, again, the latter approach is indispensable for the analysis of fluctuations of arbitrary intensity in nonlinear systems.

To conclude this chapter, I have to emphasize again that the Fokker-Plank and Smoluchowski equations give a quantitative description of time evolution of nonlinear Brownian systems with finite dissipation in the classical limit. The description of quantum properties of such dissipative (“open”) and nonlinear quantum systems is more complex,^70 and only a few simple problems of such theory have been solved so far,^71 typically using a particular model of the environment, e.g., as a large set of harmonic oscillators with different statistical distributions of their parameters, leading to different frequency dependence of susceptibility $\chi(\omega)$.

### 5.10 Exercise problems

#### 5.1
Considering the first 30 digits of number $\pi = 3.1415\ldots$ as a statistical ensemble of integers $k$ (equal to 3, 1, 4, 1, 5, …), calculate

(i) average $\langle k \rangle$, and
(ii) the r.m.s. fluctuation $\delta k$.

Compare the results with those for an ensemble of completely random integers 0, 1, ..,9, and comment.

#### 5.2
For a set of $N$ non-interacting Ising “spins” $s_j = \pm 1$, placed into magnetic field $h$, calculate the relative fluctuation of system’s magnetization.

*Hint:* The total magnetic moment of an Ising system is assumed to be proportional to the sum

---

^68 This fact is not in any contradiction with the nonvanishing result (183) which is only valid for a sub-ensemble with a certain (deterministic) initial condition $q_0$.

^69 The involved table integral may be found, e.g., in MA Eq. (6.11).

^70 See, e.g., QM Sec. 7.6.

\[ S \equiv \sum_{j=1}^{N} s_j , \]

so that the requested relative fluctuation may be calculated just as \( \delta S / \langle S \rangle \).

5.3. For a field-free, two-site Ising system with energy values \( E_m = -J s_1 s_2 \), in the thermal equilibrium at temperature \( T \), find the variance of energy fluctuations. Explore the low-temperature and high-temperature limits of the result.

5.4. For the 1D, three-site Ising ring with ferromagnetic coupling (and no external field), calculate the correlation coefficient \( \langle s_j s_{j'} \rangle \) for both \( j = j' \) and \( j \neq j' \).

5.5. Within the framework of Weiss’ molecular-field theory, calculate the variance of spin fluctuations in the \( d \)-dimensional Ising model. Use the result to derive the conditions of quantitative validity of the theory.

5.6. Calculate the variance of fluctuations of the energy of a quantum harmonic oscillator of frequency \( \omega \), in thermal equilibrium at temperature \( T \), and express it via the average value of the energy.

5.7. Express the r.m.s. fluctuation of the occupancy \( N_k \) of a certain energy level \( \epsilon_k \) by:

(i) a classical particle,

(ii) a fermion, and

(iii) a boson,

in the thermodynamic equilibrium, via the average occupancy \( \langle N_k \rangle \), and compare the results.

5.8. Starting from the Maxwell distribution of velocities, calculate constant \( C \) in the (approximate) expression \( K_P(\tau) = C \delta(\tau) \), for the correlation function of fluctuations of pressure \( P(t) \) of an ideal gas of \( N \) classical particles. Compare the result with that of Problem 3.2, and estimate the pressure fluctuation variance.

\( \text{Hint:} \) You may like to consider a cylindrically-shaped container of volume \( V = LA \) (see Fig. on the right) to calculate fluctuations of force acting on its plane lid of area \( A \), and then recalculate them into fluctuations of pressure \( P \).

5.9. Perhaps the simplest model of diffusion is the 1D \emph{discrete random walk}: each time interval \( \tau \), a particle leaps, with equal probability, to any of two neighboring sites of a 1D lattice with the spatial period \( a \). Prove that particle’s displacement during time interval \( t \gg \tau \), obeys Eq. (77), and calculate the corresponding diffusion coefficient \( D \).

5.10. Calculate the low-frequency spectral density of current \( I(t) \) due to random passage of charged particles between two conducting electrodes - see Fig. on the right. Assume that the particles are emitted by one of the electrodes at random times, and are fully absorbed by the counterpart electrode.
5.11. Within the rotating-wave approximation (RWA), calculate major statistical properties of fluctuations of the phase of classical self-oscillations, at:

(i) the free run of the oscillator, and
(ii) its phase locking by an external sinusoidal force,
assuming that the fluctuations are caused by a weak, broadband noise with spectral density $S(\omega)$.

5.12. Calculate the correlation function of the coordinate of a 1D harmonic oscillator with small Ohmic damping at thermal equilibrium.

5.13. Consider a very long, uniform, two-wire transmission line (see Fig. on the right), that allows the propagation of TEM waves with negligible attenuation, in thermal equilibrium with the environment at temperature $T$. Find variance $\langle \nu^2 \rangle_{\Delta \nu}$ of electromagnetic fluctuations of voltage $\nu$ between the wires within a small frequency interval $\Delta \nu$.

**Hint:** As an E&M reminder, TEM waves propagate with a frequency-independent velocity (equal to $c$ if the wires are in vacuum), with voltage $\nu$ and current $I$ (see Fig. above) related as $\nu(x,t)/I(x,t) = \pm Z$, where $Z$ is a frequency-independent constant (“wave impedance”).

5.14. Now consider a similar line terminated, at one end, with an impedance-matching resistor $R = Z$. Find variance $\langle \nu^2 \rangle_{\Delta \nu}$ of the voltage across the resistor, and discuss the relation between the result and the Nyquist theorem (81).

**Hint:** Take into account that resistor with $R = Z$ absorbs incident TEM waves without reflection.

5.15. An overdamped classical 1D particle escapes from a potential well with a smooth bottom, but a sharp edge – see Fig. on the right. Find the appropriate modification of the Kramers formula (139).

5.16. A particle may occupy any of $N$ similar sites. Particle’s interaction with environment induces its random, incoherent jumps from the occupied site to any other one with the same rate $\Gamma$. Find the correlation function and the spectral density of fluctuations of the instant occupancy $n(t)$ (equal to either 1 or 0) of any particular site.

---

72 See, e.g., CM Sec. 4.3.
73 See, e.g., EM Sec. 7.6.