Chapter 5. Fluctuations

This chapter discusses fluctuations of macroscopic variables, mostly in thermodynamic equilibrium. In particular, it describes the intimate connection between fluctuations and dissipation (damping) in dynamic systems coupled to multi-particle environments. This connection culminates in the Einstein relation between the diffusion coefficient and mobility, the Nyquist formula, and its 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

At the beginning of Chapter 2, we discussed the notion of averaging a variable *f* over a statistical ensemble – see Eqs. (2.7) and (2.10). Now, the *fluctuation* of the variable is defined simply as its deviation from the average $\langle f \rangle$:

Fluctuation

$$\boxed{\tilde{f} \equiv f - \langle f \rangle;} \qquad (5.1)$$

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

$$\langle \tilde{f} \rangle \equiv \langle f - \langle f \rangle \rangle = \langle f \rangle - \langle \langle f \rangle \rangle = \langle f \rangle - \langle f \rangle \equiv 0. \qquad (5.2)$$

As a result, such an average cannot be used to characterize fluctuations' *intensity*, and the simplest meaningful characteristic of the intensity is the *variance* (sometimes called "dispersion"):

Variance: definition

$$\boxed{\langle \tilde{f}^2 \rangle = \left\langle \left(f - \langle f \rangle \right)^2 \right\rangle.} \qquad (5.3)$$

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

$$\langle \tilde{f}^2 \rangle = \left\langle \left(f - \langle f \rangle \right)^2 \right\rangle = \left\langle f^2 - 2f\langle f \rangle + \langle f \rangle^2 \right\rangle = \langle f^2 \rangle - 2\langle f \rangle^2 + \langle f \rangle^2, \qquad (5.4\text{a})$$

so, finally:

Variance via averages

$$\boxed{\langle \tilde{f}^2 \rangle = \langle f^2 \rangle - \langle f \rangle^2.} \qquad (5.4\text{b})$$

As the simplest example, consider a variable that takes only two values, ±1, with equal probabilities $W_j = ½$. For such a variable, the basic Eq. (2.7) yields

$$\langle f \rangle = \sum_j W_j f_j = \frac{1}{2}(+1) + \frac{1}{2}(-1) = 0; \qquad (5.5\text{a})$$

however,

$$\langle f^2 \rangle = \sum_j W_j f_j^2 = ½(+1)^2 + ½(-1)^2 = 1, \quad \text{so that} \quad \langle \tilde{f}^2 \rangle = \langle f^2 \rangle - \langle f \rangle^2 = 1 \neq 0. \qquad (5.5\text{b})$$

The square root of the variance,

r.m.s. fluctuation

$$\boxed{\delta f \equiv \langle \tilde{f}^2 \rangle^{1/2},} \qquad (5.6)$$

is called the *root-mean-square* (*r.m.s.*) *fluctuation*. An advantage of this measure is that it has the same dimensionality as the variable itself, so the ratio $\delta f/\langle f\rangle$ is dimensionless, and is used to characterize the *relative intensity* of fluctuations.

As has been mentioned in Chapter 1, all results of thermodynamics are valid only if the fluctuations of thermodynamic variables (the internal energy E, entropy S, etc.) are relatively small.[1] Let us make a simple estimate of the relative intensity of fluctuations, taking as the simplest example a system of N independent, similar parts (e.g., particles), and an extensive variable

$$\mathscr{F} \equiv \sum_{k=1}^{N} f_k \ . \tag{5.7}$$

where all single-particle functions f_k are similar, besides that each of them depends on the state of only "its own" (k^{th}) part. The statistical average of such $\mathscr{F}$ is evidently

$$\langle \mathscr{F}\rangle = \sum_{k=1}^{N}\langle f\rangle = N\langle f\rangle, \tag{5.8}$$

while the variance of its fluctuations is

$$\left\langle \widetilde{\mathscr{F}}^{2}\right\rangle \equiv \left\langle \widetilde{\mathscr{F}}\widetilde{\mathscr{F}}\right\rangle \equiv \left\langle \sum_{k=1}^{N}\widetilde{f}_k \sum_{k'=1}^{N}\widetilde{f}_{k'}\right\rangle \equiv \left\langle \sum_{k,k'=1}^{N}\widetilde{f}_k\widetilde{f}_{k'}\right\rangle \equiv \sum_{k,k'=1}^{N}\left\langle \widetilde{f}_k\widetilde{f}_{k'}\right\rangle . \tag{5.9}$$

Now we may use the fact that for two independent variables:

$$\left\langle \widetilde{f}_k\widetilde{f}_{k'}\right\rangle = \left\langle \widetilde{f}_k\right\rangle\left\langle \widetilde{f}_{k'}\right\rangle = 0, \qquad \text{for } k' \neq k \,; \tag{5.10}$$

indeed, the first of these equalities may be used as the mathematical definition of their independence. Hence, only the terms with $k' = k$ make nonzero contributions to the right-hand side of Eq. (9):

$$\left\langle \widetilde{\mathscr{F}}^{2}\right\rangle = \sum_{k,k'=1}^{N}\left\langle \widetilde{f}_k^{2}\right\rangle \delta_{k,k'} = N\left\langle \widetilde{f}^{2}\right\rangle . \tag{5.11}$$

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

Relative fluctuation estimate

$$\boxed{\frac{\delta\mathscr{F}}{\langle\mathscr{F}\rangle} = \frac{1}{N^{1/2}}\frac{\delta f}{\langle f\rangle},} \tag{5.12}$$

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 $N_A \sim 10^{24}$. Nevertheless, in many situations even small fluctuations of variables are important, and in this chapter, we will calculate their basic properties, starting with the variance.

It should be comforting for the reader to notice that for one very important case, such a calculation has already been done in our course. Indeed, for any generalized coordinate q and generalized momentum p that give quadratic contributions of the type (2.46) to the system's

---

[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 (probably, bad :-) tradition, will be still called just T everywhere, besides the last part of Sec. 3, where temperature fluctuations are discussed explicitly.

Hamiltonian (as in a harmonic oscillator), we have derived the equipartition theorem (2.48), valid in the classical limit. Since the average values of these variables, in the thermodynamic equilibrium, equal zero, Eq. (6) immediately yields their r.m.s. fluctuations:

$$\delta p = (mT)^{1/2}, \qquad \delta q = \left(\frac{T}{\kappa}\right)^{1/2} \equiv \left(\frac{T}{m\omega^2}\right)^{1/2}, \qquad \text{where } \omega = \left(\frac{\kappa}{m}\right)^{1/2}. \tag{5.13}$$

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

$$\delta p = \left(\frac{\hbar m\omega}{2}\coth\frac{\hbar\omega}{2T}\right)^{1/2}, \qquad \delta q = \left(\frac{\hbar}{2m\omega}\coth\frac{\hbar\omega}{2T}\right)^{1/2}. \tag{5.14}$$

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 with a fixed number of particles, i.e. a member of a *microcanonical* ensemble, the internal energy does not fluctuate: $\delta E = 0$. However, if such a system is in thermal contact with the environment, i.e. is a member of a *canonical* ensemble (Fig. 2.6), the situation is different. Indeed, for such a system we may apply the general Eq. (2.7), with W_m given by the Gibbs distribution (2.58)-(2.59), not only to E but also to E^2. As we already know from Sec. 2.4, the former average,

$$\langle E\rangle = \sum_m W_m E_m, \qquad W_m = \frac{1}{Z}\exp\left\{-\frac{E_m}{T}\right\}, \qquad Z = \sum_m \exp\left\{-\frac{E_m}{T}\right\}, \tag{5.15}$$

yields Eq. (2.61b), which may be rewritten in the form

$$\langle E\rangle = \frac{1}{Z}\frac{\partial Z}{\partial(-\beta)}, \qquad \text{where } \beta \equiv \frac{1}{T}, \tag{5.16}$$

which is more convenient for our current purposes. Let us carry out a similar calculation for E^2:

$$\langle E^2\rangle = \sum_m W_m E_m^2 = \frac{1}{Z}\sum_m E_m^2 \exp\{-\beta E_m\}. \tag{5.17}$$

It is straightforward to verify, by double differentiation, that the last expression may be rewritten in a form similar to Eq. (16):

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

Now it is easy to use Eqs. (4) to calculate the variance of energy fluctuations:

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

---

[2] Unfortunately, even in some renowned textbooks, certain formulas pertaining to fluctuations are either incorrect or given without specifying the conditions of their applicability, so the reader's caution is advised.

Since Eqs. (15)-(19) are valid only if the system's volume V is fixed (because its change may affect the energy spectrum E_m), it is customary to rewrite this important result as follows:

$$\boxed{\left\langle \tilde{E}^2 \right\rangle = \frac{\partial \langle E \rangle}{\partial(-1/T)} = T^2 \left(\frac{\partial \langle E \rangle}{\partial T} \right)_V \equiv C_V T^2 .} \tag{5.20}$$

Fluctuations of E

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 are proportional to N, so Eq. (20) yields $\delta E \propto N^{1/2}$ 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) in thermal equilibrium.

Some corollaries of this result will be discussed in the next section, and now let us 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 its environment at will. If the chemical potential μ of the environment and its temperature T are fixed, i.e. we are dealing with the *grand canonical ensemble* (Fig. 2.13), we may use the grand canonical distribution (2.106)-(2.107):

$$W_{m,N} = \frac{1}{Z_G} \exp\left\{ \frac{\mu N - E_{m,N}}{T} \right\}, \qquad Z_G = \sum_{N,m} \exp\left\{ \frac{\mu N - E_{m,N}}{T} \right\}. \tag{5.21}$$

Acting exactly as we did above for the internal energy, we get

$$\langle N \rangle = \frac{1}{Z_G} \sum_{m,N} N \exp\left\{ \frac{\mu N - E_{m,N}}{T} \right\} = \frac{T}{Z_G} \frac{\partial Z_G}{\partial \mu}, \tag{5.22}$$

$$\langle N^2 \rangle = \frac{1}{Z_G} \sum_{m,N} N^2 \exp\left\{ \frac{\mu N - E_{m,N}}{T} \right\} = \frac{T^2}{Z_G} \frac{\partial^2 Z_G}{\partial \mu^2}, \tag{5.23}$$

so the particle number's variance is

$$\boxed{\left\langle \tilde{N}^2 \right\rangle = \left\langle N^2 \right\rangle - \left\langle N \right\rangle^2 = \frac{T^2}{Z_G} \frac{\partial Z_G}{\partial \mu} - \frac{T^2}{Z_G^2} \left(\frac{\partial Z_G}{\partial \mu} \right)^2 = T \frac{\partial}{\partial \mu} \left(\frac{T}{Z_G} \frac{\partial Z_G}{\partial \mu} \right) = T \frac{\partial \langle N \rangle}{\partial \mu},} \tag{5.24}$$

Fluctuations of N

in full analogy with Eq. (19).[3]

In particular, for an ideal classical gas, we may combine the last result with Eq. (3.32b). (As was already emphasized in Sec. 3.2, though that result has been obtained for the canonical ensemble, in which the number of particles N is fixed, at $N >> 1$ the fluctuations of N in the grand canonical ensemble should be relatively small, so the same relation should be valid for the *average* $\langle N\rangle$ in that ensemble.) Easily solving Eq. (3.32b) for $\langle N\rangle$, we get

$$\langle N \rangle = \text{const} \times \exp\left\{ \frac{\mu}{T} \right\}, \tag{5.25}$$

where "const" means a factor constant at the partial differentiation of $\langle N \rangle$ over μ, required by Eq. (24). Performing the differentiation and then using Eq. (25) again,

---

[3] Note, however, that for the grand canonical ensemble, Eq. (19) is generally invalid.

$$\frac{\partial\langle N\rangle}{\partial\mu} = \text{const}\times\frac{1}{T}\exp\left\{\frac{\mu}{T}\right\} = \frac{\langle N\rangle}{T}, \tag{5.26}$$

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

Fluctuations of N: classical gas

$$\boxed{\langle\tilde{N}^2\rangle = \langle N\rangle, \quad \text{i.e. } \delta N = \langle N\rangle^{1/2}.} \tag{5.27}$$

This relation is so important that I will also show how it may be derived differently. As a by-product of this new derivation, we will 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 its part of volume $V \leq V_0$ – see Fig. 1.

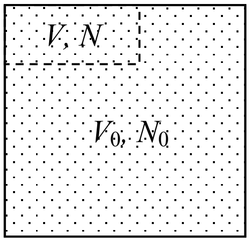


Fig. 5.1. Deriving the binomial, Poisson, and Gaussian distributions.

For one particle, such probability is obviously $W = V/V_0 = \langle N\rangle/N_0 \leq 1$, while the probability of having that particle in the remaining part of the volume is $W' = 1 - W = 1 - \langle N\rangle/N_0$. If all particles are distinct, the probability of having $N \leq N_0$ *specific* particles in volume V and $(N - N_0)$ *specific* particles in volume $(V - V_0)$ is $W^N W'^{(N_0-N)}$. However, if we do not want to distinguish the particles, we should multiply this probability by the number of possible particle combinations keeping the numbers N and N_0 constant, i.e. by the binomial coefficient $N_0!/N!(N_0 - N)!$.[4] As a result, the required probability is

Binomial distribution

$$\boxed{W_N = W^N W'^{(N_0-N)}\frac{N_0!}{N!(N_0-N)!} = \left(\frac{\langle N\rangle}{N_0}\right)^N\left(1-\frac{\langle N\rangle}{N_0}\right)^{N_0-N}\frac{N_0!}{N!(N_0-N)!}.} \tag{5.28}$$

This is the so-called *binomial probability distribution*,[5] valid for *any* $\langle N\rangle$ and N_0.

Still keeping $\langle N\rangle$ arbitrary, we can simplify the binomial distribution by assuming that the whole volume V_0, and hence N_0, are very large:

$$N_0 >> N, \tag{5.29}$$

where N means any value of interest, including $\langle N\rangle$. Indeed, 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, $(N_0 - N + 1)\,(N_0 - N + 2)\ldots(N_0 - 1)N_0$, by just N_0^N. As a result, we get

$$W_N \approx \left(\frac{\langle N\rangle}{N_0}\right)^N\left(1-\frac{\langle N\rangle}{N_0}\right)^{N_0}\frac{N_0^N}{N!} \equiv \frac{\langle N\rangle^N}{N!}\left(1-\frac{\langle N\rangle}{N_0}\right)^{N_0} = \frac{\langle N\rangle^N}{N!}\left[(1-W)^{\frac{1}{W}}\right]^{\langle N\rangle}, \tag{5.30}$$

---

[4] See, e.g., MA Eq. (2.2).

[5] It was derived by Jacob Bernoulli (1655-1705).

where, as before, $W = \langle N \rangle / N_0$. In the limit (29), $W \to 0$, so the factor inside the square brackets tends to $1/e$, the reciprocal of the natural logarithm base.[6] Thus, we get an expression independent of N_0:

Poisson distribution

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

This is the much-celebrated *Poisson distribution*[7] 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, is not necessarily an integer.

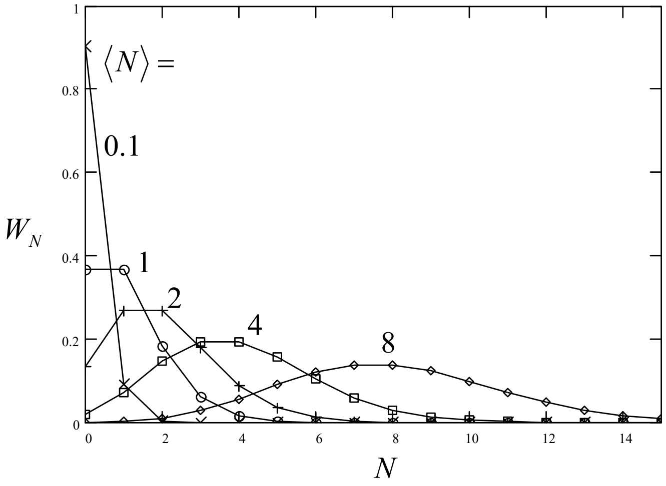


Fig. 5.2. The Poisson distribution for several values of $\langle N \rangle$. In contrast to that average, the argument N may take only integer values, so the lines in these plots are only guides for the eye.

In the limit of very small $\langle N \rangle$, the function $W_N(N)$ is close to an exponent, $W_N \approx W^N \propto \langle N \rangle^N$, while in the opposite limit, $\langle N \rangle >> 1$, it rapidly approaches the *Gaussian* (or "normal") *distribution*[8]

Gaussian distribution

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

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

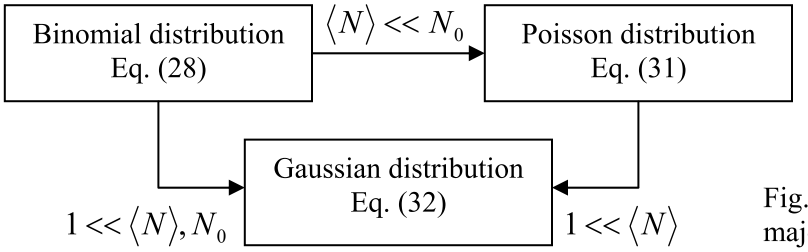


Fig. 5.3. The hierarchy of three major probability distributions.

---

[6] Indeed, this is just the most popular definition of that major mathematical constant – see, e.g., MA Eq. (1.2a) with $n = -1/W$.

[7] Named after the same Siméon Denis Poisson (1781-1840) who is also responsible for other major mathematical results and tools used in this series, including the Poisson equation – see, e.g., Sec. 6.4 below.

[8] Named after Carl Friedrich Gauss (1777-1855), even though Pierre-Simone Laplace (1749-1827) is credited for substantial contributions to its development.

A major property of the Poisson (and hence of the Gaussian) distribution is that it has the same variance as given by Eq. (27):

$$\left\langle \tilde{N}^2 \right\rangle \equiv \left\langle \left(N - \langle N \rangle\right)^2 \right\rangle = \langle N \rangle. \tag{5.33}$$

(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 specific conditions. For example, if the volume V occupied by a gas is externally fixed (say, by rigid walls), it obviously does not fluctuate at all: $\delta V = 0$. On the other hand, the volume may fluctuate in the situation when the 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 complicated by the fact that in most cases of interest, it is physically impracticable to fix its conjugate variable P, i.e. suppress its fluctuations. For example, the force $\mathscr{F}(t)$ exerted by an ideal classical gas on a container's wall (whose measure the pressure is) is the result of individual, independent hits of the wall by particles (Fig. 4), with the time scale $\tau_c \sim r_B/\langle v^2\rangle^{1/2} \sim r_B/(T/m)^{1/2} \sim 10^{-16}$ s, so its spectrum extends to very high frequencies, virtually impossible to control.

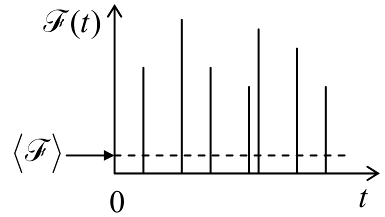


Fig. 5.4. The force exerted by gas particles on a container's wall, as a function of time (schematically).

However, we can use the following trick, very typical for the theory of fluctuations. It is almost evident that the r.m.s. fluctuations of the gas volume are independent of the shape of the container. Let us consider a particular situation similar to that shown in Fig. 1.5, with the container of a cylindrical shape, with the base area A.[9] Then the coordinate of the piston is just $q = V/A$, while the average force exerted by the gas on the cylinder is $\langle \mathscr{F} \rangle = \langle P \rangle A$ – see Fig. 5. Now if the piston is sufficiently massive, the frequency ω of its free oscillations near the equilibrium position is low enough to satisfy the following three conditions.

First, besides balancing the average force $\langle \mathscr{F} \rangle$ and thus sustaining the average pressure $\langle P \rangle$ of the gas, the interaction between the heavy piston and the relatively light particles of the gas is weak, because of a relatively short duration of the particle hits (Fig. 4). As a result, the full energy of the system may be represented as a sum of those of the particles and the piston, with a quadratic contribution to the piston's potential energy by small deviations from the equilibrium:

---

[9] As a math reminder, the term "cylinder" does not necessarily mean the "circular cylinder"; the shape of its cross-section may be arbitrary; it just should not change with height.

$$U_{\mathrm{p}} = \frac{\kappa}{2}\tilde{q}^2, \qquad \text{where } \tilde{q} \equiv q - \langle q \rangle = \frac{\tilde{V}}{A}, \tag{5.34}$$

and κ is the effective spring constant arising from the finite compressibility of the gas.

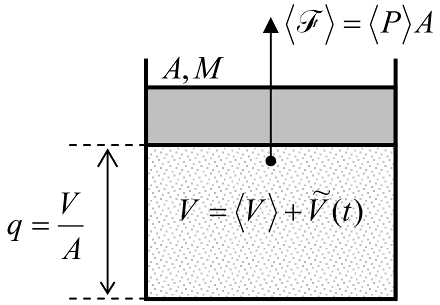


Fig. 5.5. Deriving Eq. (37).

Second, at $\omega = (\kappa/M)^{1/2} \rightarrow 0$, this spring constant may be calculated just as for constant variations of the volume, with the gas remaining in quasi-equilibrium at all times:

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

This partial derivative[10] should be calculated at whatever the thermal conditions are, e.g., with S = const for adiabatic conditions (i.e., a thermally insulated gas), or with T = const for isothermal conditions (including a good thermal contact between the gas and a heat bath), etc. With that constant denoted as X, Eqs. (34)-(35) give

$$U_{\mathrm{p}} = \frac{1}{2}\left(-A^2\frac{\partial\langle P \rangle}{\partial\langle V \rangle}\right)_X \left(\frac{\tilde{V}}{A}\right)^2 \equiv \frac{1}{2}\left(-\frac{\partial\langle P \rangle}{\partial\langle V \rangle}\right)_X \tilde{V}^2. \tag{5.36}$$

Finally, assuming that ω is also small in the sense $\hbar\omega << T$, we may apply, to the piston's fluctuations, the classical equipartition theorem: $\langle U_{\mathrm{p}} \rangle = T/2$, giving[11]

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

Fluctuations of V

Since this result is valid for any A and ω, it should not depend on the system's geometry and the piston's 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. For the

---

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

[11] One may meet statements that a similar formula,

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

is valid for pressure fluctuations. However, this equality does not take into account the different physical nature of pressure (Fig. 4), with its very broad frequency spectrum. This issue will be discussed later in this chapter.

particular case of fluctuations at constant temperature ($X = T$),[12] we may use the definition (3.58) of the isothermal bulk modulus (reciprocal compressibility) K_T of the gas to rewrite Eq. (37a) as

$$\left\langle \tilde{V}^2 \right\rangle_T = \frac{TV}{K_T}. \tag{5.37b}$$

For 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), again with $X = T$, to get

$$\left\langle \tilde{V}^2 \right\rangle_T = -T\left(-\frac{NT}{\langle P \rangle^2} \right) = \frac{\langle V \rangle^2}{N}, \qquad \text{i.e.} \quad \frac{\delta V_T}{\langle V \rangle} = \frac{1}{N^{1/2}}, \tag{5.38}$$

in full agreement with the general 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 sample we are considering is weakly coupled to a heat bath of temperature T_0, in the sense that the time τ of temperature equilibration between the two is much larger than the time of the internal equilibration, called *thermalization*. Then we may assume that, on the former time scale, temperature T describes the whole sample, though it may fluctuate:

$$T = \langle T \rangle + \tilde{T}. \tag{5.39}$$

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

$$\tilde{T} = \frac{\tilde{E}}{C_V}, \quad \text{so that} \quad \delta T = \frac{\delta E}{C_V}. \tag{5.40}$$

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

Fluctuations of T

$$\boxed{\delta T = \frac{\delta E}{C_V} = \frac{\langle T \rangle}{C_V^{1/2}}.} \tag{5.41}$$

The most straightforward application of this result is to analyses of so-called *bolometers* – broadband detectors of electromagnetic radiation in microwave and infrared frequency bands. (In particular, they are used for measurements of the CMB radiation, which was discussed in Sec. 2.6.) In such a detector (Fig. 6), the incoming radiation is focused on a small sensor (e.g., a small piece of a germanium crystal, a superconductor thin film at temperature $T \approx T_c$, etc.), which is well isolated thermally from the environment. As a result, the absorption of an even small radiation power $\mathscr{P}$ leads to a noticeable change ΔT of the sensor's average temperature $\langle T \rangle$ and hence of its electric resistance R, which is probed up by low-noise external electronics.[13] If the power does not change in time too fast, ΔT is a certain function of $\mathscr{P}$, turning to 0 at $\mathscr{P} = 0$. Hence, if ΔT is much lower than the environment temperature T_0, we may keep only the main, linear term in its Taylor expansion in small $\mathscr{P}$:

---

[12] In this case, we may also use the second of Eqs. (1.39) to rewrite Eq. (37) via the second derivative $(\partial^2 G/\partial P^2)_T$.

[13] Besides low internal electric noise, a good sensor should have a sufficiently large *temperature responsivity* dR/dT, making the noise contribution by the readout electronics insignificant – see below.

$$\Delta T \equiv \langle T \rangle - T_0 = \frac{\mathscr{P}}{\mathscr{G}}, \tag{5.42}$$

where the coefficient $\mathscr{G} \equiv \partial\mathscr{P}/\partial T$ is called the *thermal conductance* of the (perhaps unintentional but unavoidable) thermal coupling between the sensor and the heat bath – see Fig. 6.

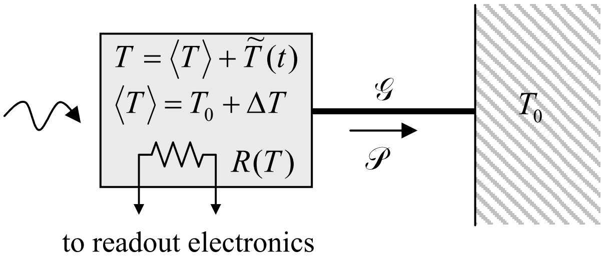


Fig. 5.6. The conceptual scheme of a bolometer.

The power may be detected if the electric signal from the sensor, which results from the change ΔT, is not drowned in spontaneous fluctuations. In practical systems, these fluctuations are contributed by several sources including electronic amplifiers. However, in modern systems, these "technical" contributions to noise are successfully suppressed,[14] and the dominating noise source is the fundamental sensor temperature fluctuations described by Eq. (41). In this case, the so-called *noise-equivalent power* ("NEP"), defined as the level of $\mathscr{P}$ that produces the signal equal to the r.m.s. value of noise, may be calculated by equating the expressions (41) (with $\langle T \rangle = T_0$) and (42):

$$\text{NEP} \equiv \mathscr{P}\Big|_{\Delta T = \delta T} = \frac{T_0 \mathscr{G}}{C_V^{1/2}}. \tag{5.43}$$

This expression shows that to decrease the NEP, i.e. improve the detector's sensitivity, both the environment temperature T_0 and the thermal conductance $\mathscr{G}$ should be reduced. In modern receivers of radiation, their typical values are of the order of 0.1 K and 10^{-10} W/K, respectively.

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

$$\tau = \frac{C_V}{\mathscr{G}}, \tag{5.44}$$

---

[14] An important modern trend in this progress [see, e.g., P. Day *et al.*, *Nature* **425**, 817 (2003)] is the replacement of the resistive temperature sensors $R(T)$ with thin and narrow superconducting strips with temperature-sensitive kinetic inductance $L_k(T)$ – see the model solution of EM Problems 6.20-6.21. Such inductive sensors have zero dc resistance and hence vanishing Johnson-Nyquist noise at typical signal pickup frequencies of a few kHz – see Eq. (81) and its discussion below. A recent example of their application is in the Prime-Cam receiver of the FYST submillimeter telescope (see, e.g., https://en.wikipedia.org/wiki/Fred_Young_Submillimeter_Telescope), to see first light in 2024.

becomes close to $1/\omega$, because at $\omega\tau >> 1$ the useful signal drops faster than noise. So, the lowest (i.e. the best) values of the NEP,

$$(\text{NEP})_{\min} = \alpha T_0 \mathscr{G}^{1/2} \nu^{1/2}, \qquad \text{with } \alpha \sim 1, \tag{5.45}$$

are reached at $\nu\tau \approx 1$. (The exact values of the optimal product $\omega\tau$, and of the numerical constant $\alpha \sim 1$ in Eq. (45), depend on the exact law of the power modulation and the readout signal processing procedure.) With the parameters cited above, this estimate yields $(\text{NEP})_{\min}/\nu^{1/2} \sim 3\times10^{-17}$ W/Hz$^{1/2}$ – a very low power indeed.

However, perhaps counter-intuitively, the power modulation allows the 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 ω, we can use the subsequent electronics stages to filter out all the noise besides its components within a very narrow band, of width $\Delta\nu << \nu$, around the modulation frequency (Fig. 7). This is the idea of a *microwave radiometer*,[15] currently used in all sensitive broadband receivers of radiation.

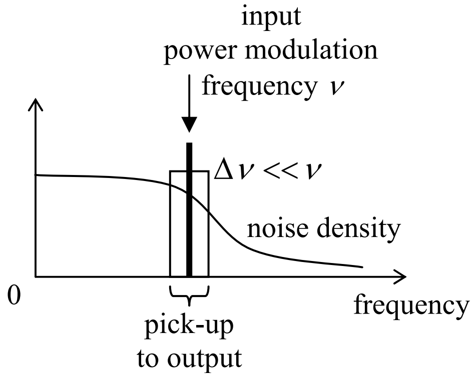


Fig. 5.7. The basic idea of the Dicke radiometer.

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 a need for analysis of variables dominated by fast (high-frequency) components, such as pressure – please have one more look at Fig. 4. Finally, during such an 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

In most discussions of the previous three sections, the averaging $\langle\ldots\rangle$ of variables was assumed to be over an appropriate statistical ensemble of many similar systems. However, as was discussed in Sec. 2.1, most physical systems of interest are ergodic. If such a system is also *stationary*, i.e. the statistical averages of its variables do not change with time, the averaging may be also understood as that over a sufficiently long time interval. In this case, we may think about fluctuations of any variable f as a random process taking place in just one particular system, but developing in time: $\tilde{f} = \tilde{f}(t)$.

---

[15] It was pioneered in the 1950s by Robert Henry Dicke, so the device is frequently called the *Dicke radiometer*. Note that the optimal strategy of using similar devices for time- and energy-resolved detection of single high-energy photons is different – though even it is essentially based on Eq. (41). For a recent brief review of such detectors see, e.g., K. Morgan, *Phys. Today* **71**, 29 (Aug. 2018), and references therein.

There are two mathematically equivalent approaches to the description of such random functions of time, called the *time-domain* picture and the *frequency-domain* picture, with their relative convenience depending on the particular problem to be solved. In the time domain, we need to characterize *random* fluctuations $\tilde{f}(t)$ by some *deterministic* function of time. Evidently, the average $\langle \tilde{f}(t) \rangle$ cannot be used for this purpose, because it equals zero – see Eq. (2). Of course, the variance (3) is not necessarily equal to zero, but if the system is stationary, that average cannot depend on time either. Because of that, let us consider the following average:

$$\left\langle \tilde{f}(t)\tilde{f}(t') \right\rangle . \tag{5.46}$$

Generally, this is a function of two arguments. However, in a stationary system, the average (46) may depend only on the difference,

$$\tau \equiv t' - t , \tag{5.47}$$

between the two observation times. In this case, the average (46) is called the *correlation function* of the variable *f*:

Correlation function

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

Again, here the averaging may be understood as that either over a statistical ensemble of macroscopically similar systems or over a sufficiently long interval of the time argument *t*, with the argument τ kept constant. The correlation function's name[16] catches the idea of this notion very well: $K_f(\tau)$ characterizes the mutual relation between the fluctuations of the variable *f* at two times separated by the given interval τ. Let us list the basic properties of this function.[17]

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

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

with $t' \equiv t - \tau$. For stationary processes, this average cannot depend on the common shift of two observation times *t* and *t'*, so the 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:

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

In the opposite limit, when τ is much larger than a certain characteristic *correlation time* τ_c of the system,[18] the correlation function has to tend to zero because the fluctuations separated by such time

---

[16] Another term, the *autocorrelation function*, is sometimes used for the average (48) to distinguish it from the *mutual correlation function*, $\langle f_1(t) f_2(t+\tau) \rangle$, of two different stationary processes.

[17] Note that this correlation function is the direct temporal analog of the spatial correlation function briefly discussed in Sec. 4.2 – see Eq. (4.30).

[18] Note that the correlation time τ_c is the direct temporal analog of the correlation radius r_c that was discussed in Sec. 4.2 – see the same Eq. (4.30).

interval are virtually independent (*uncorrelated*) – see Eq. (10). As a result, the correlation function typically looks like one of the plots sketched in Fig. 8.

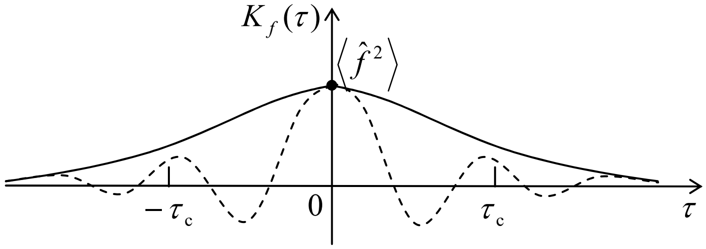


Fig. 5.8. The correlation function of fluctuations: two typical examples.

Note that on a time scale much longer than τ_c, any physically realistic correlation function may be well approximated with a delta function of τ. For example, for a process that is a sum of independent very short pulses, e.g., the gas pressure force exerted on the container wall (Fig. 4), this approximation is legitimate on time scales much longer than the single pulse duration, e.g., the time of particle's interaction with on the wall at the impact.

In the reciprocal, frequency domain, the same process $\tilde{f}(t)$ is represented as a Fourier integral,[19]

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

with the reciprocal transform being

$$f_\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{f}(t) e^{i\omega t} dt . \tag{5.53}$$

If the function $\tilde{f}(t)$ is random (as it is in the case of fluctuations), with zero average, its Fourier transform f_ω is also a random function (now of frequency), also with a vanishing statistical average. Indeed, now thinking of the operation $\langle \ldots \rangle$ as an ensemble averaging, we may write

$$\langle f_\omega \rangle = \left\langle \frac{1}{2\pi}\int_{-\infty}^{+\infty} \tilde{f}(t)\, e^{i\omega t} dt \right\rangle = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \langle \tilde{f}(t) \rangle e^{i\omega t} dt = 0 . \tag{5.54}$$

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

$$\langle f_\omega f^*_{\omega'} \rangle = \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty} dt' \int_{-\infty}^{+\infty} dt \langle \tilde{f}(t)\tilde{f}(t') \rangle e^{i(\omega' t' - \omega t)} . \tag{5.55}$$

It turns out that for a stationary process, the 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 the integration over $\tau \equiv t' - t$ (at fixed t), also in infinite limits. Replacing t' with $t + \tau$ in the 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

---

[19] The argument of the function f_ω is represented as its index with the purpose of emphasizing that this function is different from $\tilde{f}(t)$, while (very conveniently) still using the same letter for the same variable.

$$\left\langle f_\omega f^*_{\omega'} \right\rangle = \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} dt \int_{-\infty}^{+\infty} d\tau K_f(\tau) e^{i(\omega-\omega')t} e^{i\omega'\tau} \equiv \frac{1}{(2\pi)^2} \int_{-\infty}^{+\infty} K_f(\tau) e^{i\omega'\tau} d\tau \int_{-\infty}^{+\infty} e^{i(\omega-\omega')t} dt \,. \quad (5.56)$$

But the last integral is just $2\pi\delta(\omega - \omega')$,[20] so we finally get

$$\left\langle f_\omega f^*_{\omega'} \right\rangle = S_f(\omega)\delta(\omega - \omega'), \quad (5.57)$$

where the real function of frequency,

Spectral density of fluctuations

$$\boxed{S_f(\omega) \equiv \frac{1}{2\pi} \int_{-\infty}^{+\infty} K_f(\tau) e^{i\omega\tau} d\tau = \frac{1}{\pi} \int_0^{\infty} K_f(\tau) \cos\omega\tau \, d\tau \,,} \quad (5.58)$$

is called the *spectral density of fluctuations at frequency ω*. According to Eq. (58), the spectral density is just the Fourier image of the correlation function, and hence the reciprocal Fourier transform is:[21,22]

Wiener-Khinchin theorem

$$\boxed{K_f(\tau) = \int_{-\infty}^{+\infty} S_f(\omega) e^{-i\omega\tau} d\omega = 2\int_0^{\infty} S_f(\omega) \cos\omega\tau \, d\omega \,.} \quad (5.59)$$

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

$$\left\langle \tilde{f}^2 \right\rangle \equiv K_f(0) = \int_{-\infty}^{+\infty} S_f(\omega) d\omega \equiv 2\int_0^{\infty} S_f(\omega) d\omega \,. \quad (5.60)$$

The last relation shows that the term "spectral density" describes the physical sense of the function $S_f(\omega)$ very well. Indeed, if a random signal $f(t)$ had been passed through a frequency filter with a small bandwidth $\Delta\nu << \nu$ of positive cyclic frequencies, the integral in the last form of Eq. (60) could be limited to the interval $\Delta\omega = 2\pi\Delta\nu$, i.e. the variance of the filtered signal would become

$$\left\langle \tilde{f}^2 \right\rangle_{\Delta\nu} = 2S_f(\omega)\Delta\omega \equiv 4\pi S_f(\omega)\Delta\nu \,. \quad (5.61)$$

(A popular alternative definition of the spectral density is $\mathscr{S}_f(\nu) \equiv 4\pi S_f(\omega)$, making the average (61) equal to just $\mathscr{S}_f(\nu)\Delta\nu$.)

To conclude this introductory (mostly mathematical) section, let me note an important particular case. If the spectral density of some process is nearly constant within the frequency range of interest, $S_f(\omega) = \text{const} = S_f(0)$,[23] Eq. (59) shows that its correlation function may be well approximated with a delta function:

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

---

[20] See, e.g., MA Eq. (14.4).

[21] 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.

[22] Although Eqs. (58) and (59) look not much more than straightforward corollaries of the Fourier transform, they bear a special name of the *Wiener-Khinchin theorem* – after the mathematicians N. Wiener and A. Khinchin who have proved that these relations are valid even for the functions $f(t)$ that are not square-integrable, so from the point of view of standard mathematics, their Fourier transforms are not well defined.

[23] Such a process is frequently called *white noise*, because it consists of all frequency components with equal amplitudes, reminding the white light, which consists of many monochromatic components with close amplitudes.

From this relation stems another popular name of the white noise, the *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 the spectral density of a realistic, limited physical variable the approximation of constant spectral density cannot be true for *all* frequencies (otherwise, for example, the integral (60) would diverge, giving an unphysical, infinite value of its variance), and may be valid only at frequencies much lower than $1/\tau_c$.

5.5. Fluctuations and dissipation

Now we are equipped mathematically to address a major issue of statistical physics, the relation between fluctuations and dissipation This relation is especially simple for the following hierarchical situation: a relatively "heavy", slowly evolving system, weakly interacting with an environment consisting of many rapidly moving, "light" components. A popular theoretical term for such a system is the *Brownian particle*, named after botanist Robert Brown who was first to notice (in 1827) the random motion of small particles (in his case, pollen grains), caused by their random hits by fluid's molecules, under a microscope. However, the family of such systems is much broader than that of small mechanical particles. Just for a few examples, such description is valid for an atom interacting with electromagnetic field modes of the surrounding space, a macroscopic mechanical system interacting with molecules of air around it, a macroscopic electric current interacting with microscopic charge carriers, etc.[24]

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 typical mechanical pendulum – its motion does not overheat the air around it to any noticeable extent.) In this case, the averaging over a statistical ensemble of *similar environments* at a fixed, specific motion of the system of interest, may be performed assuming their thermal equilibrium.[25] I will denote such a "primary" averaging by the usual angle brackets $\langle\ldots\rangle$. At a later stage, we may carry out additional, "secondary" averaging over an ensemble of many similar systems of interest, coupled to similar environments. When we do, such secondary averaging will be denoted by double angle brackets $\langle\langle\ldots\rangle\rangle$.

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

$$m\ddot{q} + \kappa q = \mathscr{F}_{\text{det}}(t) + \mathscr{F}_{\text{env}}(t) \equiv \mathscr{F}_{\text{det}}(t) + \langle \mathscr{F} \rangle + \tilde{\mathscr{F}}(t), \qquad \text{with } \left\langle \tilde{\mathscr{F}}(t) \right\rangle = 0, \qquad (5.63)$$

where q is the (generalized) coordinate of the oscillator, $\mathscr{F}_{\text{det}}(t)$ is the deterministic external force, while both components of the force $\mathscr{F}_{\text{env}}(t)$ represent the impact of the environment on the oscillator's motion. Again, on the time scale of the fast-moving environmental components, the oscillator's motion is slow. The average component $\langle \mathscr{F} \rangle$ of the force exerted by the environment on such a slowly moving object is frequently independent of its coordinate q but does depend on its velocity $\dot{q}$. For most such systems, the Taylor expansion of the force in small velocity starts with a non-zero linear term:

---

[24] To emphasize this generality, in the forthcoming discussion of the 1D case, I will use the letter q rather than x for the system's displacement.

[25] 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 τ_c is the correlation time of the environment, while τ is the characteristic time scale of motion of our "heavy" system of interest.

$$\langle \mathcal{F} \rangle = -\eta \dot{q}\,, \tag{5.64}$$

where the constant η is usually called the *drag* (or "kinematic friction", or "damping") *coefficient*, so Eq. (63) may be rewritten as

Langevin equation for classical oscillator

$$\boxed{m\ddot{q} + \eta \dot{q} + \kappa q = \mathcal{F}_{\text{det}}(t) + \tilde{\mathcal{F}}(t)\,.} \tag{5.65}$$

This method of describing the environmental effects on an otherwise Hamiltonian system is called the *Langevin equation*.[26] Due to the linearity of the differential equation (65), its general solution may be represented as a sum of two independent parts: the deterministic motion of the damped linear oscillator due to the external force $\mathcal{F}_{\text{det}}(t)$, and its random fluctuations due to the random force $\tilde{\mathcal{F}}(t)$ exerted by the environment. The former effects are well-known from classical dynamics,[27] so let us focus on the latter part by taking $\mathcal{F}_{\text{det}}(t) = 0$. The remaining term on the right-hand side of Eq. (65) 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 the system's motion.[28]

Plugging into Eq. (65) the representation of both variables in the Fourier form similar to Eq. (52), and requiring the coefficients before the same $\exp\{-i\omega t\}$ to be equal on both sides of the equation, for their Fourier images we get the following relation:

$$-m\omega^2 q_\omega - i\omega\eta\, q_\omega + \kappa q_\omega = \mathcal{F}_\omega, \tag{5.66}$$

which immediately gives us q_ω, i.e. the (random) complex amplitude of the coordinate fluctuations:

$$q_\omega = \frac{\mathcal{F}_\omega}{(\kappa - m\omega^2) - i\eta\omega} \equiv \frac{\mathcal{F}_\omega}{m(\omega_0^2 - \omega^2) - i\eta\omega}\,. \tag{5.67}$$

Now multiplying Eq. (67) by its complex conjugate for another frequency (say, ω'), averaging both parts of the resulting equation, and using the formulas similar to Eq. (57) for each of them,[29] we get the following relation between the spectral densities of the oscillations and the random force:[30]

---

[26] Named after Paul Langevin, whose 1908 work was the first systematic development of A. Einstein's ideas on the 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.

[27] See, e.g., CM Sec. 5.1. Again, here I assume that the variable $f(t)$ is classical, with the discussion of the quantum case postponed until the end of the section.

[28] Note that the direct secondary statistical averaging of Eq. (65) with $\mathcal{F}_{\text{det}} = 0$ yields $\langle\langle q \rangle\rangle = 0$! This, perhaps a bit counter-intuitive result becomes less puzzling if we recognize that this is the averaging over a large statistical ensemble of random sinusoidal oscillations with all values of their phase and that the (equally probable) oscillations with opposite phases give mutually canceling contributions to the sum in Eq. (2.6).

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

[30] Regardless of the physical sense of such a function of ω, and of whether its maximum is situated at a finite frequency ω_0 as in Eq. (68) or at $\omega = 0$, it is often referred to as the *Lorentzian* (or "Breit-Wigner") *line*.

$$S_q(\omega) = \frac{1}{m^2(\omega_0^2 - \omega^2)^2 + \eta^2\omega^2} S_{\mathscr{F}}(\omega). \tag{5.68}$$

In the so-called low-damping limit ($\eta << m\omega_0$), the fraction on the right-hand side of Eq. (68) has a sharp peak near the oscillator's own frequency ω_0 (describing the well-known effect of high-*Q resonance*), and may be well approximated in that vicinity as

$$\frac{1}{m^2(\omega_0^2 - \omega^2)^2 + (\eta\omega)^2} \approx \frac{1}{\eta^2\omega_0^2(\xi^2 + 1)}, \qquad \text{with } \xi \equiv \frac{2m(\omega - \omega_0)}{\eta}. \tag{5.69}$$

In contrast, the spectral density of fluctuations of a typical environment is changing relatively slowly, so for the purpose of integration over frequencies near ω_0, we may replace $S_{\mathscr{F}}(\omega)$ with $S_{\mathscr{F}}(\omega_0)$. As a result, the variance of the environment-imposed random oscillations may be calculated, using Eq. (60), as[31]

$$\langle\langle \tilde{q}^2 \rangle\rangle = 2\int_0^\infty S_q(\omega)d\omega \approx 2\int_{\omega\approx\omega_0} S_q(\omega)d\omega = 2S_{\mathscr{F}}(\omega_0)\frac{1}{\eta^2\omega_0^2}\frac{\eta}{2m}\int_{-\infty}^{+\infty}\frac{d\xi}{\xi^2+1}. \tag{5.70}$$

This is a well-known table integral,[32] equal to π, so, finally:

$$\langle\langle \tilde{q}^2 \rangle\rangle = 2S_{\mathscr{F}}(\omega_0)\frac{1}{\eta^2\omega_0^2}\frac{\eta}{2m}\pi \equiv \frac{\pi}{m\omega_0^2\eta}S_{\mathscr{F}}(\omega_0) \equiv \frac{\pi}{\kappa\eta}S_{\mathscr{F}}(\omega_0). \tag{5.71}$$

On the other hand, any weak interaction with the 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

$$\frac{\kappa}{2}\langle\langle \tilde{q}^2 \rangle\rangle = \frac{T}{2}. \tag{5.72}$$

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

$$S_{\mathscr{F}}(\omega_0) = \frac{\eta}{\pi}T. \tag{5.73a}$$

Now we may argue (rather convincingly :-) that since this relation does not depend on the oscillator's parameters m and κ, and hence its eigenfrequency $\omega_0 = (\kappa/m)^{1/2}$,[33] it should be valid at any relatively low frequency ($\omega\tau_c << 1$). Using Eq. (58) with $\omega \to 0$, it may be also rewritten as a formula for the effective low-frequency drag coefficient:

---

[31] Since in this case, the process in the oscillator is entirely due to its environment, its variance should be obtained by statistical averaging over an ensemble of many similar (oscillator + environment) systems, and hence, following our convention, it is denoted by double angular brackets.

[32] See, e.g. MA Eq. (6.5a).

[33] Moreover, it does not depend on the assumption $\eta << m\omega_0$, made above only for the sake of calculation simplicity. Indeed, for a frequency-independent spectral density $S_{\mathscr{F}}$, such as the one given by Eq. (73a), the integration of both sides of Eq. (68) over all frequencies yields Eq. (71) for any η.

$$\boxed{\eta = \frac{1}{T}\int_0^\infty K_{\mathscr{F}}(\tau)\,d\tau \equiv \frac{1}{T}\int_0^\infty \left\langle \tilde{\mathscr{F}}(0)\tilde{\mathscr{F}}(\tau)\right\rangle d\tau\,.} \qquad (5.73b)$$

No dissipation without fluctuation

Formulas (73) reveal an intimate, fundamental relation between the fluctuations and the dissipation provided by a thermally-equilibrium environment. Parroting the famous political slogan, there is "no dissipation without fluctuation". This means in particular that the phenomenological description of dissipation by the drag force alone in classical mechanics[34] is (approximately) valid only when the energy scale of the considered process is much larger than T. To the best of my knowledge, this fact was first recognized in 1905 by A. Einstein,[35] for the following particular case.

Let us apply our result (73) to a free 1D Brownian particle, by taking $\mathscr{F}_{\text{det}}(t) = 0$ and $\kappa \to 0$. In this limit, both relations (71) and (72) lead to infinite coordinate variance. To understand the reason for that divergence, let us go back to the Langevin equation (65) with not only $\kappa = 0$ and $\mathscr{F}_{\text{det}}(\text{t}) = 0$, but also $m \to 0$ – just for the sake of simplicity. (The latter approximation, frequently called the *overdamping limit*, is quite appropriate, for example, for the motion of small particles in sufficiently viscous fluids – such as in R. Brown's experiments.) In this approximation, Eq. (65) is reduced to a simple equation,

$$\eta\,\dot{q} = \tilde{\mathscr{F}}(t), \quad \text{with } \left\langle \tilde{\mathscr{F}}(t)\right\rangle = 0\,, \qquad (5.74)$$

which may be readily integrated to give the particle's displacement during a finite time interval t:

$$\Delta q(t) \equiv q(t) - q(0) = \frac{1}{\eta}\int_0^t \tilde{\mathscr{F}}(t')dt'. \qquad (5.75)$$

Evidently, at the full statistical averaging of the displacement, the fluctuation effects vanish, but this does not mean that the particle does not move – just that it has equal probabilities to be shifted in either of two possible directions. To see that, let us calculate the variance of the displacement:

$$\left\langle\left\langle \Delta\tilde{q}^2(t)\right\rangle\right\rangle = \frac{1}{\eta^2}\int_0^t dt'\int_0^t dt''\left\langle \tilde{\mathscr{F}}(t')\tilde{\mathscr{F}}(t'')\right\rangle \equiv \frac{1}{\eta^2}\int_0^t dt'\int_0^t dt'' K_{\mathscr{F}}(t'-t'')\,. \qquad (5.76)$$

As we already know, at times $\tau >> \tau_c$, the correlation function may be well approximated by the delta function – see Eq. (62). In this approximation, with $S_{\mathscr{F}}(0)$ expressed by Eq. (73a), we get

$$\left\langle\left\langle \Delta\tilde{q}^2(t)\right\rangle\right\rangle = \frac{2\pi}{\eta^2}S_{\mathscr{F}}(0)\int_0^t dt'\int_0^t dt''\delta(t''-t') = \frac{2\pi}{\eta^2}\frac{\eta T}{\pi}\int_0^t dt' = \frac{2T}{\eta}t \equiv 2Dt\,, \qquad (5.77)$$

with

$$\boxed{D = \frac{T}{\eta}.} \qquad (5.78)$$

Einstein's relation

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 $(2Dt)^{1/2}$. The coefficient D in this

---

[34] See, e.g., CM Sec. 5.1.

[35] It was obtained in one of the three papers of Einstein's celebrated 1905 "triad". As a reminder, another of the papers started the relativity theory, and one more was the quantum description of the photoelectric effect, essentially starting quantum mechanics. Not too bad for one year of one young scientist's life!

relation is called the *coefficient of diffusion*, and Eq. (78) describes the extremely simple and important[36] *Einstein's relation* between that coefficient and the drag coefficient. Often this relation is rewritten, in the SI units of temperature, as $D = \mu k_B T_K$, where $\mu \equiv 1/\eta$ is the *mobility* of the particle. The physical sense of μ becomes clear from the expression for the deterministic velocity (particle's "drift"), which follows from the averaging of both sides of Eq. (74) after the restoration of the term $\mathscr{F}_{\text{det}}(t)$ in it:

$$v_{\text{drift}} \equiv \langle\langle \dot{q}(t) \rangle\rangle = \frac{1}{\eta} \mathscr{F}_{\text{det}}(t) \equiv \mu \mathscr{F}_{\text{det}}(t), \tag{5.79}$$

so the mobility is just the drift velocity given to the particle by a unit force.[37]

Another famous embodiment of the general Eq. (73) is the *thermal* (or "Johnson", or "Johnson-Nyquist", or just "Nyquist") *noise* in resistive electron devices. Let us consider a lumped,[38] two-terminal, dissipation-free "probe" circuit, playing the role of the harmonic oscillator in the analysis carried out above, connected to a resistive device (Fig. 9), playing the role of the probe circuit's environment. (The noise is generated by the thermal motion of numerous electrons, randomly moving inside the resistive device.) For this system, one convenient choice of the 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 the voltage $\mathscr{V}$ across its terminals, with the polarity shown in Fig. 9. (Indeed, the product $\mathscr{V}dQ$ is the elementary work $d\mathscr{W}$ done by the environment on the probe circuit.)

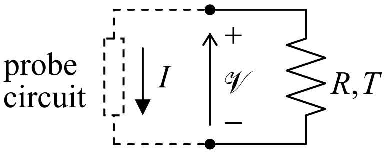


Fig. 5.9. A resistive device as a dissipative environment of a two-terminal probe circuit.

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

$$\langle \mathscr{V} \rangle = -\eta \dot{Q} \equiv -\eta I. \tag{5.80}$$

Comparing this relation with Ohm's law, $\mathscr{V} = R(-I)$,[39] we see that in this case, the coefficient η has the physical sense of the usual Ohmic resistance R of our dissipative device,[40] so Eq. (73a) becomes

$$S_{\mathscr{V}}(\omega) = \frac{R}{\pi} T. \tag{5.81a}$$

---

[36] In particular, in 1908, i.e. very soon after Einstein's publication, it was used by J. Perrin for an accurate determination of the Avogadro number N_A. (It was Perrin who graciously suggested naming this constant after A. Avogadro, honoring his pioneering studies of gases in the 1810s.)

[37] Note that in solid-state physics and electronics, the charge carrier mobility is usually defined as $|\mathbf{v}_{\text{drift}}/\mathscr{E}| = e\mathbf{v}_{\text{drift}}/|\mathscr{F}_{\text{det}}| \equiv e|\mu|$ (where $\mathscr{E}$ is the applied electric field), and is traditionally measured in cm^2/V·s.

[38] As a (good :-) student of classical electrodynamics should know, lumped (compact) electric circuits may be described by the usual Kirchhoff laws, neglecting the wave propagation effects – see, e.g., EM Sec. 6.6.

[39] The minus sign here is due to the fact that in our notation, the current flowing in the resistor, from the terminal assumed to be positive to the negative one, is ($-I$) – see Fig. 9.

[40] Due to this fact, Eq. (64) is often called the *Ohmic model* of the environment's response, even if the physical nature of the variables q and $\mathscr{F}$ is completely different from the electric charge and voltage.

Using the last equality in Eq. (61), and expressing temperature in the SI units ($T = k_B T_K$), we may bring this famous *Nyquist formula*[41] to its most popular form:

Nyquist formula

$$\boxed{\left\langle \widetilde{\mathscr{V}}^2 \right\rangle_{\Delta\nu} = 4k_B T_K R\Delta\nu .} \qquad (5.81b)$$

Note that according to Eq. (65), this result is only valid at a negligible speed of change of the coordinate q (in our current case, negligible current I), i.e. Eq. (81) describes the voltage fluctuations as would be measured by a virtually ideal *voltmeter*, with its input resistance much higher than R. On the other hand, using a different choice of generalized coordinate and force, $q \to \Phi$, $\mathscr{F} \to I$ (where $\Phi \equiv \int \mathscr{V}(t)dt$ is the generalized magnetic flux, so $d\mathscr{W} = I\mathscr{V}(t)dt \equiv Id\Phi$), we get $\eta \to 1/R$, and Eq. (73) yields the thermal fluctuations of the current through the resistive device, as would be measured by a virtually ideal *ammeter*, i.e. at $\mathscr{V} \to 0$:

$$S_I(\omega) = \frac{1}{\pi R}T, \qquad \text{i.e. } \left\langle \widetilde{I}^2 \right\rangle_{\Delta\nu} = \frac{4k_B T_K}{R}\Delta\nu . \qquad (5.81c)$$

The nature of Eqs. (81) is so fundamental that they may be used, in particular, for the so-called *Johnson noise thermometry*.[42]

Note, however, that these relations are valid for noise in thermal equilibrium only. In electric circuits that may be readily driven out of equilibrium by an applied voltage $\mathscr{V}$, other types of noise are frequently important, notably the *shot noise* that arises in short conductors, e.g., tunnel junctions, at applied voltages with $|\mathscr{V}| >> T/q$, due to the discreteness of charge carriers.[43] A straightforward analysis (left for the reader's exercise) shows that this noise may be characterized by current fluctuations with the following low-frequency spectral density:

Schottky formula

$$\boxed{S_I(\omega) = \frac{\left| q\bar{I} \right|}{2\pi}, \qquad \text{i.e. } \left\langle \widetilde{I}^2 \right\rangle_{\Delta\nu} = 2\left| q\bar{I} \right|\Delta\nu ,} \qquad (5.82)$$

where q is the electric charge of a single current carrier. This is the *Schottky formula*,[44] valid for any relation between the average I and $\mathscr{V}$. The 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 the effective temperature

$$T_{\text{ef}} = \frac{\left| q\bar{\mathscr{V}} \right|}{2} >> T . \qquad (5.83)$$

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

---

[41] It is named after Harry Nyquist who derived this formula in 1928 (independently of the prior work by A. Einstein, M. Smoluchowski, and P. Langevin) to describe the "Johnson-Nyquist" noise that had been just discovered experimentally by his Bell Labs colleague John Bertrand Johnson. The derivation of Eq. (73) and hence of Eq. (81) in these notes is essentially a twist of the derivation used by H. Nyquist.

[42] See, e.g., J. Crossno *et al.*, *Appl. Phys. Lett*. **106**, 023121 (2015), and references therein.

[43] Another practically important type of fluctuations in electronic devices is the low-frequency *1/f noise* that was already mentioned in Sec. 3 above. I will briefly discuss it in Sec. 8.

[44] It was derived by Walter Hans Schottky as early as 1918, i.e. even before Nyquist's work.

relaxation length l_e of the charge carriers.[45] (Another mechanism of shot noise suppression, which may become noticeable in highly conductive nanoscale devices, is the Fermi-Dirac statistics of electrons.[46])

Now let us return for a minute to the bolometric Dicke radiometer (see Figs. 6-7 and their discussion in Sec. 4), and use the Langevin formalism to finalize its analysis. For this system, the Langevin equation is an extension of the usual equation of heat balance:

$$C_V \frac{dT}{dt} + \mathscr{G}(T - T_0) = \mathscr{P}_{\text{det}}(t) + \widetilde{\mathscr{P}}(t), \tag{5.84}$$

where $\mathscr{P}_{\text{det}} \equiv \langle \mathscr{P} \rangle$ describes the (deterministic) power of the absorbed radiation and $\widetilde{\mathscr{P}}$ represents 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 similarly to how this was done with Eq. (65), assuming that the frequency spectrum of the fluctuation source is much broader than the intrinsic bandwidth $1/\tau = \mathscr{G}/C_V$ of the bolometer, so its spectral density at frequencies $\omega\tau \sim 1$ may be well approximated by its low-frequency value $S_{\mathscr{P}}(0)$:

$$S_T(\omega) = \left| \frac{1}{-i\omega C_V + \mathscr{G}} \right|^2 S_{\mathscr{P}}(0). \tag{5.85}$$

Then, requiring the variance of temperature fluctuations, calculated from this formula and Eq. (60),

$$\begin{aligned} (\delta T)^2 \equiv \left\langle \widetilde{T}^2 \right\rangle &= 2\int_0^\infty S_T(\omega)d\omega = 2S_{\mathscr{P}}(0)\int_0^\infty \left| \frac{1}{-i\omega C_V + \mathscr{G}} \right|^2 d\omega \\ &\equiv 2S_{\mathscr{P}}(0)\frac{1}{C_V^2}\int_0^\infty \frac{d\omega}{\omega^2 + (\mathscr{G}/C_V)^2} = \frac{\pi S_{\mathscr{P}}(0)}{\mathscr{G}C_V}, \end{aligned} \tag{5.86}$$

to coincide with our earlier "thermodynamic fluctuation" result (41), we get

$$S_{\mathscr{P}}(0) = \frac{\mathscr{G}}{\pi} T_0^2 . \tag{5.87}$$

The r.m.s. value of the "power noise" within a bandwidth $\Delta\nu << 1/\tau$ (see Fig. 7) becomes equal to the deterministic signal power $\mathscr{P}_{\text{det}}$ (or more exactly, the main harmonic of its modulation law) at

$$\mathscr{P} = \mathscr{P}_{\min} \equiv \left(\left\langle \widetilde{\mathscr{P}}^2 \right\rangle_{\Delta\nu} \right)^{1/2} = \left(2S_{\mathscr{P}}(0)\Delta\omega \right)^{1/2} = 2\left(\mathscr{G}\Delta\nu \right)^{1/2} T_0 . \tag{5.88}$$

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 ν 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 of orders of magnitude, enabling observation of extremely weak radiation. Maybe the most spectacular example is the recent measurements of the CMB radiation, which corresponds to blackbody temperature $T_K \approx 2.726$ K, with accuracy $\delta T_K \sim 10^{-6}$ K, using microwave receivers with the physical

---

[45] See, e.g., Y. Naveh *et al., Phys. Rev. B* **58**, 15371 (1998). In practically used metals, l_e is of the order of a few nanometers, so the usual "macroscopic" resistors do not exhibit the shot noise.

[46] For a review of this effect see, e.g., Ya. Blanter and M. Büttiker, *Phys. Repts.* **336**, 1 (2000).

temperature of all their components much higher than δT. The observed weak (~10^{-5} K) anisotropy of the CMB radiation is a major experimental basis of all modern cosmology.[47]

Returning to the discussion of our main result, Eq. (73), let me note that it may be readily generalized to the case when the environment's response is different from the Ohmic form (64). This opportunity is virtually evident from Eq. (66): by its derivation, the second term on its left-hand side is just the Fourier component of the average response of the environment to the system's displacement:

$$\langle \mathscr{F}_\omega \rangle = i\omega\eta\, q_\omega . \tag{5.89}$$

Now let the response be still linear, but have an arbitrary frequency dispersion,

$$\langle \mathscr{F}_\omega \rangle = \chi(\omega) q_\omega , \tag{5.90}$$

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

$$\chi(\omega) = \chi'(\omega) + i\chi''(\omega) . \tag{5.91}$$

Then Eq. (73) remains valid with the replacement $\eta \rightarrow \chi''(\omega)/\omega$:[48]

$$S_{\mathscr{F}}(\omega) = \frac{\chi''(\omega)}{\pi\omega} T . \tag{5.92}$$

This fundamental relation[49] may be used not only to calculate the fluctuation intensity from the known generalized responsibility (i.e. the deterministic response of the system to a small perturbation), but also in reverse – to calculate such linear response from the known fluctuations. The latter use is especially attractive at numerical simulations of complex systems, e.g., those based on molecular-dynamics approaches, because it circumvents the need in extracting a weak response to a small perturbation out of a 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 were based on the apparently classical equation of motion, Eq. (63). However, quantum mechanics shows[50] that a similar equation is valid for the corresponding Heisenberg-picture operators, so by repeating all the arguments that have led us to the Langevin equation (65), we may write its quantum-mechanical version

$$\boxed{m\ddot{\hat{q}} + \eta\dot{\hat{q}} + \kappa\hat{q} = \hat{\mathscr{F}}_{\text{det}} + \hat{\tilde{\mathscr{F}}} .} \tag{5.93}$$

Heisenberg-Langevin equation

---

[47] See, e.g., a concise book by A. Balbi, *The Music of the Big Bang*, Springer, 2008.

[48] Reviewing the calculations leading to Eq. (73), we may see that 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 the oscillator's frequency ω_0. This renormalization is insignificant if the oscillator-to-environment coupling is weak, i.e. if the susceptibility $\chi(\omega)$ is small – as had been assumed at the derivation of Eq. (69) and hence Eq. (73).

[49] 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 1951 paper by H. Callen and T. Welton, where a more general result (98) had been derived. Much more adequately, the Green/Kubo names are associated with Eq. (102) below.

[50] See, e.g., QM Sec. 4.6.

This is the so-called *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 side 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:

$$\hat{\tilde{f}}(t)\hat{\tilde{f}}(t') \neq \hat{\tilde{f}}(t')\hat{\tilde{f}}(t), \qquad \text{if } t' \neq t\,. \tag{5.94}$$

As a result, the function defined by Eq. (46) may not be a symmetric function of the time delay $\tau \equiv t' - t$ even for a stationary process, making it inadequate for the representation of the actual correlation function – which has to obey Eq. (50). This technical difficulty may be overcome by the introduction of the following *symmetrized correlation function*[51]

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

(where {…,…} denotes the anticommutator of the two operators), and, similarly, the symmetrical spectral density $S_f(\omega)$, defined by the following 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).

Now we may repeat all the analysis that was carried out for the classical case, and get Eq. (71) again, but now this expression has to be compared not with the equipartition theorem, but with its quantum-mechanical generalization (14), which, in our current notation, reads

$$\left\langle \left\langle \tilde{q}^2 \right\rangle \right\rangle = \frac{\hbar\omega_0}{2\kappa}\coth\frac{\hbar\omega_0}{2T}\,. \tag{5.97}$$

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

FDT

$$\boxed{S_{\mathcal{F}}(\omega) = \frac{\hbar\chi''(\omega)}{2\pi}\coth\frac{\hbar\omega}{2T}\,.} \tag{5.98}$$

This is the much-celebrated *fluctuation-dissipation theorem*, usually referred to just as the FDT, first derived in 1951 by Herbert Bernard Callen and Theodore A. Welton – in a somewhat different way.

As natural as it seems, this generalization of the relation between fluctuations and dissipation 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

Quantum noise

$$\boxed{S_{\mathcal{F}}(\omega) = \frac{\hbar\chi''(\omega)}{2\pi},} \tag{5.99}$$

---

[51] Here (and to the end of this section) the averaging ⟨…⟩ should be understood in the general quantum-statistical sense – see Eq. (2.12). As was discussed in Sec. 2.1, for the classical-mixture state of the system, this does not create any difference in either the mathematical treatment of the averages or their physical interpretation.

which describes what is frequently called *quantum noise*. According to the quantum Langevin equation (93), nothing but the random force $\tilde{\mathscr{F}}$ exerted by the environment, with the spectral density (99) proportional to the imaginary part of susceptibility (i.e. damping), is the source of the ground-state "fluctuations" of the coordinate and momentum of a quantum harmonic oscillator, with the r.m.s. values

$$\delta q \equiv \left\langle\left\langle \tilde{q}^2 \right\rangle\right\rangle^{1/2} = \left(\frac{\hbar}{2m\omega_0} \right)^{1/2}, \qquad \delta p \equiv \left\langle\left\langle \tilde{p}^2 \right\rangle\right\rangle^{1/2} = \left(\frac{\hbar m\omega_0}{2} \right)^{1/2}, \tag{5.100}$$

and the total energy $\hbar\omega_0/2$. On the other hand, 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 basic commutation relation

$$\left[\hat{q}, \hat{p}\right] = i\hbar . \tag{5.101}$$

So, what is the genuine source of the uncertainty described by Eqs. (100)?

The best resolution of this paradox I can offer is that *either* interpretation of Eqs. (100) is legitimate, with their relative convenience depending on the particular application. One may say that since the right-hand side of the quantum Langevin equation (93) is a quantum-mechanical operator rather than a classical force, it "carries the quantum uncertainty relation within itself". However, this (admittedly, opportunistic :-) view leaves the following question open: is the quantum noise (99) of an 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 ground-state uncertainty, i.e. their own quantum fluctuations, which may be readily confused with those of the system under study. Fortunately, this difficulty may be overcome, for example, by using unique frequency-mixing ("down-conversion") properties of Josephson junctions. Special low-temperature experiments using such down-conversion[52] have confirmed that the quantum noise (99) is quite real and measurable.

Finally, let me mention an alternative derivation[53] of the fluctuation-dissipation theorem (98) from the general quantum mechanics of open systems. This derivation is substantially longer than that presented above but gives an interesting sub-product, the *Green-Kubo formula*

$$\left\langle \left[\hat{\tilde{\mathscr{F}}}(t), \hat{\tilde{\mathscr{F}}}(t+\tau) \right] \right\rangle = i\hbar \mathscr{G}(\tau) , \tag{5.102}$$

where $\mathscr{G}(\tau)$ is the temporal Green's function of the environment, defined by the following relation:

$$\left\langle \mathscr{F}(t) \right\rangle = \int_0^\infty \mathscr{G}(\tau) q(t-\tau) d\tau \equiv \int_{-\infty}^t \mathscr{G}(t-t') q(t') dt' . \tag{5.103}$$

Plugging the Fourier transforms of all three functions of time participating in Eq. (103) into this relation, it is straightforward to check that this Green's function is just the Fourier image of the complex susceptibility $\chi(\omega)$ defined by Eq. (90):

$$\int_0^\infty \mathscr{G}(\tau) e^{i\omega\tau} d\tau = \chi(\omega) ; \tag{5.104}$$

---

[52] R. Koch *et al.*, *Phys. Rev. B* **26**, 74 (1982), and references therein.

[53] See, e.g., QM Sec. 7.4.

here 0 is used as the lower limit instead of $(-\infty)$ just to emphasize that due to the causality principle, Green's function has to be equal to zero for $\tau < 0$.[54]

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 similar time-domain form:

$$\left\langle \left\{ \hat{\tilde{\mathscr{F}}}(t), \hat{\tilde{\mathscr{F}}}(t+\tau) \right\} \right\rangle = 2K_{\mathscr{F}}(\tau), \qquad (5.105)$$

where the symmetrized correlation function $K_{\mathscr{F}}(\tau)$ is most simply described by its Fourier transform, which is, according to Eq. (58), equal to $\pi S_{\mathscr{F}}(\omega)$, so using the FDT, we get

$$\int_0^\infty K_{\mathscr{F}}(\tau) \cos \omega\tau \, d\tau = \frac{\hbar \chi''(\omega)}{2} \coth \frac{\hbar\omega}{2T}. \qquad (5.106)$$

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 $\chi(\omega)$ of the environment. However, the averaged anticommutator also depends on temperature, while the averaged commutator does not – at least explicitly. (The complex susceptibility of an environment may be temperature-dependent as well.)

5.6. The Kramers problem and the Smoluchowski equation

Returning to the classical case, it is evident that the Langevin equations of the type (65) provide means not only for the analysis of stationary fluctuations but also for the description of the time evolution of (classical) systems coupled to their environments – which, again, may provide both dissipation and fluctuations. However, this approach to evolution analysis suffers from two major handicaps.

First, the Langevin equation does enable a straightforward calculation of the statistical average of the variable q, and its fluctuation variance – i.e., in the common mathematical terminology, the first and second *moments* of the probability density $w(q, t)$ – as functions of time, but not of the probability distribution as such. Admittedly, this is rarely a big problem, because in most cases 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 "linear" systems – i.e., the systems whose dynamics may be 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[55]) are reduced to motion in potentials $U_{\text{ef}}(q)$ that substantially differ from quadratic parabolas, giving nonlinear equations of motion. If the energy of interaction between the system and its random environment is factorable – i.e. is a product of variables belonging to these subsystems (as it is very frequently the case), we may repeat all arguments of the last section to derive the following generalized version of the 1D Langevin equation:

---

[54] See, e.g., CM Sec. 5.1.

[55] See, e.g., CM Secs. 3.4-3.6.

$$m\ddot{q} + \eta\,\dot{q} + \frac{\partial U(q,t)}{\partial q} = \widetilde{\mathscr{F}}(t), \tag{5.107}$$

valid for an arbitrary, possibly time-dependent potential $U(q, t)$.[56] Unfortunately, the solution of this equation may be very hard. Indeed, its Fourier analysis carried out in the last section was essentially based on the linear superposition principle, which is invalid for nonlinear equations.

If the fluctuation intensity is low, $|\delta q| << \langle q \rangle$, where $\langle q \rangle(t)$ is the deterministic solution of Eq. (107) in the absence of fluctuations, this equation may be linearized[57] with respect to small fluctuations $\tilde{q} \equiv q - \langle q \rangle$ to get a linear equation of motion:

$$m\ddot{\tilde{q}} + \eta\,\dot{\tilde{q}} + \kappa(t)\tilde{q} = \widetilde{\mathscr{F}}(t), \qquad \text{with } \kappa(t) \equiv \frac{\partial^2}{\partial q^2} U\big(\langle q \rangle(t), t\big). \tag{5.108}$$

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 the fluctuations and the function $\kappa(t)$. Such calculations may be more cumbersome than those performed above, but still doable (especially if the unperturbed motion $\langle q \rangle(t)$ is periodic), and sometimes give useful analytical results.[58]

However, some important problems cannot be solved by linearization. Perhaps, the most apparent (and practically very important) example is the so-called *Kramers problem*[59] of finding the lifetime of a metastable state of a 1D classical system in a potential well separated from the region of unlimited motion with a potential barrier – see Fig. 10.

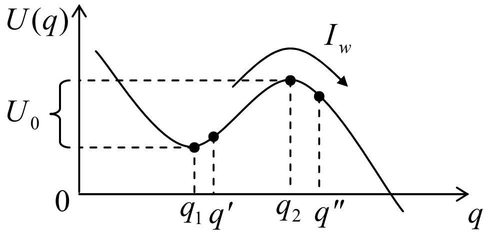


Fig. 5.10. The Kramers problem.

In the absence of fluctuations, the system, if initially at rest close to the well's bottom (in Fig. 10, at $q \approx q_1$), 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 a gradual decrease of the total probability

$$W(t) = \int_{\substack{\text{well's} \\ \text{bottom}}} w(q,t)\,dq \tag{5.109}$$

---

[56] The generalization of Eq. (107) to higher spatial dimensionality is also straightforward, with the scalar variable q replaced with a multi-dimensional vector $\mathbf{q}$, and the scalar derivative dU/dq replaced with the vector ∇U, where ∇ is the del vector-operator in the $\mathbf{q}$-space.

[57] See, e.g., CM Secs. 3.2, 5.2, and beyond.

[58] See, e.g., QM Problem 7.8, and also Chapters 5 and 6 in the monograph by W. Coffey *et al.*, cited above.

[59] It was named after Hendrik Anthony ("Hans") Kramers who, besides solving this conceptually important problem in 1940, has made several other seminal contributions to physics, including the famous Kramers-Kronig dispersion relations (see, e.g., EM Sec. 7.4) and the WKB (Wentzel-Kramers-Brillouin) approximation in quantum mechanics – see, e.g., QM Sec. 2.4.

to find the system in the well, because of a non-zero rate of its escape from it, over the potential barrier, due to thermal activation. What may be immediately expected of the situation is that if the barrier height,

$$U_0 \equiv U(q_2) - U(q_1), \tag{5.110}$$

is much larger than the temperature T,[60] the Boltzmann distribution $w \propto \exp\{-U(q)/T\}$ should be still approximately valid in most of the well, so the probability for the system to overcome the barrier in unit time should scale as $\exp\{-U_0/T\}$. From these handwaving arguments, one may reasonably expect that if the probability $W(t)$ of the system to remain in the well by time t obeys the usual "decay law"

$$\dot{W} = -\frac{W}{\tau}, \tag{5.111a}$$

then the lifetime τ has to obey the general Arrhenius law:

$$\tau = \tau_A \exp\left\{\frac{U_0}{T}\right\}. \tag{5.111b}$$

However, these relations need to be proved, and the pre-exponential coefficient τ_A (usually called the *attempt time*) needs to be calculated. This cannot be done by the linearization of Eq. (107), because this approximation 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 again.

This and other essentially nonlinear problems may be addressed using an alternative approach to fluctuations, dealing directly with the time evolution of the probability density $w(q, t)$. Due to the shortage of time/space, I will review this approach using mostly handwaving arguments, and refer the interested reader to special literature[61] for strict mathematical proofs. Let us start with the diffusion of a free classical 1D particle with inertial effects negligible in comparison with damping. It is described by the Langevin equation (74) with $\mathscr{F}_{\text{det}} = 0$. Let us assume that at all times the probability distribution stays Gaussian:

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

where q_0 is the initial position of the particle, and $\delta q(t)$ is the time-dependent distribution width, whose growth in time is described, as we already know, by Eq. (77):

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

Then it is straightforward to verify, by substitution, that this solution satisfies the following simple partial differential equation,

$$\frac{\partial w}{\partial t} = D\frac{\partial^2 w}{\partial q^2}, \tag{5.114}$$

with the delta-functional initial condition

---

[60] If U_0 is comparable with T, the system's behavior also depends substantially on the initial probability distribution, i.e., does not follow the simple law (111).

[61] See, e.g., either R. Stratonovich, *Topics in the Theory of Random Noise*, vol. 1., Gordon and Breach, 1963, or Chapter 1 in the monograph by W. Coffey *et al*., cited above.

$$w(q,0) = \delta(q - q_0)\,. \tag{5.115}$$

The simple and important *equation of diffusion* (114) may be naturally generalized to the 3D motion:[62]

$$\boxed{\frac{\partial w}{\partial t} = D\nabla^2 w\,.} \tag{5.116}$$

Equation of diffusion

Now let us compare this equation with the probability conservation law,[63]

$$\frac{\partial w}{\partial t} + \nabla \cdot \mathbf{j}_w = 0\,, \tag{5.117a}$$

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

$$\frac{d}{dt}\int_V w d^3 r + \oint_S \mathbf{j}_w \cdot d^2\mathbf{q} = 0\,, \tag{5.117b}$$

resulting from the integration of Eq. (117a) over an arbitrary time-independent volume V limited by surface S, and the application of the divergence theorem[64] to the second term of the result.)

The continuity relation (117a) coincides with Eq. (116), with D given by Eq. (78), if we take

$$\mathbf{j}_w = -D\nabla w = -\frac{T}{\eta}\nabla w\,. \tag{5.118}$$

The first form of this relation[65] allows a simple interpretation: the probability flow is proportional to the spatial gradient of the probability density (i.e., in application to $N >> 1$ 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 concentrations. This flow is the very essence of the effect of diffusion. The second form of Eq. (118) is also not very surprising: the diffusion speed scales as temperature and is inversely proportional to the viscous drag.

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

$$\mathbf{j}_w = w\mathbf{v}\,, \tag{5.119}$$

where $\mathbf{v}$ is the deterministic velocity of the particle. In the high-damping limit we are considering right now, $\mathbf{v}$ has to be just the *drift velocity*:

$$\mathbf{v} = \frac{1}{\eta}\mathscr{F}_{\text{det}} = -\frac{1}{\eta}\nabla U(\mathbf{q})\,, \tag{5.120}$$

---

[62] As will be discussed in Chapter 6, the equation of diffusion also describes several other physical phenomena – in particular, the heat propagation in a uniform, isotropic solid, and in this context is called the *heat conduction equation* or (rather inappropriately) just the "heat equation".

[63] Both forms of Eq. (117) are similar to the mass conservation law in classical dynamics (see, e.g., CM Sec. 8.2), the electric charge conservation law in electrodynamics (see, e.g., EM Sec. 4.1), and the probability conservation law in quantum mechanics (see, e.g., QM Sec. 1.4).

[64] See, e.g., MA Eq. (12.2).

[65] In application to systems of many similar, weakly-interacting particles (to be discussed in the next chapter) where w is proportional to the particle density n, this expression is sometimes called *Fick's law*, after the physiologist A. Fick who suggested it in 1855.

where $\mathscr{F}_{\text{det}}$ is the deterministic force described by the potential energy $U(\mathbf{q})$.

Now that we have descriptions of $\mathbf{j}_w$ due to both the drift *and* the diffusion separately, we may rationally assume that in the general case when both effects are present, the corresponding components (118) and (119) of the probability current just add up, so

$$\mathbf{j}_w = \frac{1}{\eta}\left[w(-\nabla U) - T\nabla w\right], \tag{5.121}$$

and Eq. (117a) takes the form

Smoluchowski equation

$$\boxed{\frac{\partial w}{\partial t} = \frac{1}{\eta}\left[\nabla \cdot (w\nabla U) + T\nabla^2 w\right].} \tag{5.122}$$

This is the *Smoluchowski equation* (or "Smoluchowski diffusion equation"),[66] which is closely related to the *drift-diffusion equation* in multi-particle kinetics – to be discussed in the next chapter.

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

$$w\nabla U + T\nabla w = 0, \qquad \text{i.e. } \frac{\nabla w}{w} = -\frac{\nabla U}{T}. \tag{5.123}$$

Since the left-hand side of the last relation is just $\nabla(\ln w)$, it may be easily integrated over $\mathbf{q}$, giving

$$\ln w = -\frac{U}{T} + \ln C, \qquad \text{i.e. } w(\mathbf{r}) = C\exp\left\{-\frac{U(\mathbf{q})}{T}\right\}, \tag{5.124}$$

where C is a normalization constant. With both sides multiplied by the number N of similar, independent systems, with the spatial density $n(\mathbf{q}) = Nw(\mathbf{q})$, this equality becomes the Boltzmann distribution (3.26).

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, $m << \eta\tau_A$, where τ_A (still to be calculated) is some time scale of the particle's motion inside the well. It is straightforward to verify that the 1D version of Eq. (121),

$$I_w = \frac{1}{\eta}\left[w\left(-\frac{\partial U}{\partial q}\right) - T\frac{\partial w}{\partial q}\right], \tag{5.125a}$$

(where I_w is the probability current at point q, rather than its density) is mathematically equivalent to

$$I_w = -\frac{T}{\eta}\exp\left\{-\frac{U(q)}{T}\right\}\frac{\partial}{\partial q}\left(w\exp\left\{\frac{U(q)}{T}\right\}\right), \tag{5.125b}$$

so we may write

$$I_w \exp\left\{\frac{U(q)}{T}\right\} = -\frac{T}{\eta}\frac{\partial}{\partial q}\left(w\exp\left\{\frac{U(q)}{T}\right\}\right). \tag{5.126}$$

---

[66] It is named after Marian Smoluchowski, who developed this formalism in 1906. Note that sometimes Eq. (122) is referred to as the *Fokker-Planck equation*, but it is more common to use that name for a more general equation discussed in the next section.

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

$$T << U_0 . \tag{5.127}$$

when the lifetime is relatively long: $\tau >> \tau_A$. Since according to Eq. (111a), the first term of the continuity equation (117b) has to be of the order of W/τ, in this limit the term, and hence the gradient of I_w, are exponentially small, so the probability current virtually does not depend on q in the potential barrier region. Let us use this fact in the integration of both sides of Eq. (126) over that region:

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

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

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

(Obviously, such selection is only possible if condition (127) is satisfied.) In this limit, the contribution from the point q'' to the right-hand side of Eq. (129) is negligible because the probability density behind the barrier is exponentially small. On the other hand, the probability at the point q' has to be close to the value given by its quasi-stationary Boltzmann distribution (124), so

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

and Eq. (128) yields

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

Patience, my reader, 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 to the region where the difference $U(q) - U(q_1)$ is much smaller than U_0 – cf. Eq. (129). According to the Taylor expansion, the shape of virtually any smooth potential $U(q)$ near the point q_1 of its minimum may be well approximated with a quadratic parabola:

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

With this approximation, Eq. (132) is reduced to the standard Gaussian integral:[67]

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

---

[67] If necessary, see MA Eq. (6.9b) again.

To complete the calculation, we may use a similar approximation for the barrier top:

$$U(q \approx q_2) - U(q_1) \approx \left[U(q_2) - \frac{\kappa_2}{2}(q-q_2)^2 \right] - U(q_1) = U_0 - \frac{\kappa_2}{2}(q-q_2)^2,$$
$$\text{where} \quad \kappa_2 \equiv -\frac{d^2U}{dq^2}\bigg|_{q=q_2} > 0, \tag{5.135}$$

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

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

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

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

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 $I_q(-\infty) = 0$, this equation is merely

$$\frac{dW}{dt} + I_w(q') = 0. \tag{5.138}$$

In our approximation, $I_w(q')$ does not depend on the exact position of the point q', and is given by Eq. (137), so plugging it into Eq. (138), we recover the exponential decay law (111a), with the lifetime τ obeying the Arrhenius law (111b), and the following attempt time:

Kramers formula for high damping

$$\boxed{\tau_A = \frac{2\pi\eta}{(\kappa_1\kappa_2)^{1/2}} \equiv 2\pi(\tau_1\tau_2)^{1/2}, \quad \text{where} \quad \tau_{1,2} \equiv \frac{\eta}{\kappa_{1,2}}.} \tag{5.139}$$

Thus the metastable state lifetime is indeed described by the Arrhenius law, with the attempt time scaling as the geometric mean of the system's "relaxation times" near the potential well bottom (τ_1) and the potential barrier top (τ_2).[68] Let me leave it for the reader's exercise to prove that if the potential profile near the well's bottom and/or top is sharp, the expression for the attempt time should be modified, but the Arrhenius decay law (111) is not affected.

5.7. The Fokker-Planck equation

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

---

[68] Actually, τ_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 the stable 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, and ∇_q (which was denoted as ∇ in that section) is the usual vector operator in the space, while $\mathbf{j}_p$ is the current density in the momentum space, and ∇_p is the similar vector operator in that space:

$$\nabla_q \equiv \sum_{j=1}^{3} \mathbf{n}_j \frac{\partial}{\partial q_j}, \qquad \nabla_p \equiv \sum_{j=1}^{3} \mathbf{n}_j \frac{\partial}{\partial p_j}. \tag{5.141}$$

At negligible fluctuations ($T \to 0$), $\mathbf{j}_p$ may be composed using the natural analogy with $\mathbf{j}_q$ – see Eq. (119). In our new notation, that relation reads,

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

so it is natural to take

$$\mathbf{j}_p = w\dot{\mathbf{p}} = w\langle \mathscr{F} \rangle, \tag{5.143a}$$

where the (statistical-ensemble) averaged force $\langle \mathscr{F} \rangle$ includes not only the contribution due to the potential's gradient but also the drag force $-\eta\mathbf{v}$ provided by the environment – see Eq. (64) and its discussion:

$$\mathbf{j}_p = w(-\nabla_q U - \eta\mathbf{v}) = -w(\nabla_q U + \eta\frac{\mathbf{p}}{m}). \tag{5.143b}$$

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

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

allows the following particular solution:

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

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

$$\langle \dot{\mathbf{q}} \rangle = \frac{\langle \mathbf{p} \rangle}{m}, \qquad \langle \dot{\mathbf{p}} \rangle = -\nabla_q U - \eta\frac{\langle \mathbf{p} \rangle}{m}, \tag{5.146}$$

describing the particle's drift, with the usual deterministic initial conditions.

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

$$m\ddot{q}_j = \widetilde{\mathscr{F}}_j(t), \qquad \text{i.e. } \dot{p}_j = \widetilde{\mathscr{F}}_j(t). \tag{5.147}$$

The last equation is identical to the high-damping 1D equation (74) (with $\mathscr{F}_{\text{det}} = 0$), with the replacement $q \to p_j/\eta$, and hence the corresponding contribution to $\partial w/\partial t$ may be described by the last term of Eq. (122), with that replacement:

$$\left.\frac{\partial w}{\partial t}\right|_{\text{diffusion}} = D\nabla^2_{p/\eta} w = \frac{T}{\eta}\eta^2\nabla^2_p w \equiv \eta T\,\nabla^2_p w\,. \tag{5.148}$$

Now the reasonable assumption that in the arbitrary case, the drift and diffusion contributions to $\partial w/\partial t$ just add up immediately leads us to the full *Fokker-Planck equation*:[69]

Fokker-Planck equation

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

As a sanity check, let us use this equation to calculate the stationary probability distribution of the momentum of particles with an arbitrary damping η but otherwise free, in the momentum space, assuming (just for simplicity) their uniform distribution in the direct space, $\nabla_q = 0$. In this case, Eq. (149) is reduced to

$$\nabla_p \cdot \left[w\left(\eta\frac{\mathbf{p}}{m}\right)\right] + \eta T\nabla^2_p w = 0, \qquad \text{i.e. } \nabla_p \cdot \left(\frac{\mathbf{p}}{m}w + T\nabla_p w\right) = 0\,. \tag{5.150}$$

The first integration over the momentum space yields

$$\frac{\mathbf{p}}{m}w + T\nabla_p w = \mathbf{j}_w\,, \qquad \text{i.e. } w\nabla_p\left(\frac{p^2}{2m}\right) + T\nabla_p\, w = \mathbf{j}_w\,, \tag{5.151}$$

where $\mathbf{j}_w$ is a vector constant describing a possible general probability flow in the system. In the absence of such flow, $\mathbf{j}_w = 0$, we get

$$\nabla_p\left(\frac{p^2}{2m}\right) + T\frac{\nabla_p w}{w} \equiv \nabla_p\left(\frac{p^2}{2m} + T\ln w\right) = 0, \qquad \text{giving } w = \text{const} \times \exp\left\{-\frac{p^2}{2mT}\right\}, \tag{5.152}$$

i.e. the Maxwell distribution (3.5). However, the result (152) is more general than that obtained in Sec. 3.1, because it shows that the distribution stays the same even at non-zero damping. It is also straightforward to verify that in the more general case of an arbitrary stationary potential $U(\mathbf{q})$, Eq. (149) is satisfied with the stationary solution (3.24), also giving $\mathbf{j}_w = 0$.

In the limit where the damping is large, i.e. the inertial effects are relatively small, the solution of the Fokker-Planck equation tends, relatively rapidly, to the following product

$$w(\mathbf{q},\mathbf{p},t) \to \exp\left\{-\frac{[\mathbf{p} - \mathbf{p}_0(\mathbf{q},t)]^2}{2mT}\right\} \times \mathscr{w}(\mathbf{q},t)\,, \tag{5.153}$$

where $\mathbf{p}_0 \equiv -(m/\eta)\nabla_q U$, followed by a much slower time evolution of the direct-space distribution $\mathscr{w}(\mathbf{q}, t)$, described by the Smoluchowski equation (122).

Another important particular case is that of a quasi-periodic motion of a particle, with low damping, in a soft potential well. In this case, the Fokker-Planck equation describes both the diffusion of

---

[69] It was first derived by Adriaan Fokker in 1913 in his PhD thesis and further elaborated by Max Planck in 1917. (Curiously, A. Fokker is more famous for his work on music theory, and the invention and construction of several new keyboard instruments, than for this and several other important contributions to theoretical physics.)

the effective phase Θ of such (generally nonlinear, "anharmonic") oscillator, and a slow relaxation of its energy. If we are only interested in the latter effect, Eq. (149) may be reduced to the so-called *energy diffusion equation*,[70] which is much easier to solve.

However, in most practically interesting cases, solutions of Eq. (149) are rather complicated. (Indeed, the reader should remember that these solutions embody, in the particular case $T = 0$, all classical dynamics of a particle.) Because of this, I will present (rather than derive) only one more of them: the Kramers' solution[71] of his problem (Fig. 10) for $\eta/m\omega_2 \sim 1$. In this general case, the metastable state's lifetime turns out to be again given by the Arrhenius formula (111b), with the following reciprocal attempt time:

$$\frac{1}{\tau_A} \equiv \frac{\omega_1}{2\pi\omega_2}\left[\left(\omega_2^2 + \frac{\eta^2}{4m^2}\right)^{1/2} - \frac{\eta}{2m}\right] \to \begin{cases} \omega_1/2\pi, & \text{for } \eta/m\omega_2 << 1, \\ m\omega_1\omega_2/2\pi\eta \equiv 1/2\pi(\tau_1\tau_2)^{1/2}, & \text{for } 1 << \eta/m\omega_2, \end{cases} \tag{5.154}$$

where $\omega_{1,2} \equiv (\kappa_{1,2}/m)^{1/2}$. Thus, in the limit $\eta/m\omega_2 << 1$, Eqs. (111b) and (154) give a very simple result

$$\tau = \frac{2\pi}{\omega_1}\exp\left\{\frac{U_0}{T}\right\}. \tag{5.155}$$

Note, however, that this result is strictly valid only if $\eta/m\omega_2 >> T/U_0$ (as a reminder, the latter ratio has to be much smaller than 1 in order for the very notion of lifetime τ to be meaningful) and at lower damping, its pre-exponential factor requires a correction, which may be calculated using the already mentioned energy diffusion equation.[72]

The Kramers' result for the classical thermal activation of a system *over* a potential barrier may be compared with that for its quantum-mechanical tunneling *through* the barrier.[73] The WKB approximation for the latter effect gives the expression

$$\tau_Q = \tau_A \exp\left\{-2\int_{\kappa^2(q)>0} \kappa(q)dq\right\}, \qquad \text{with } \frac{\hbar^2\kappa^2(q)}{2m} \equiv U(q) - E, \tag{5.156}$$

showing that generally, the classical and quantum lifetimes of a metastable state have different dependences on the barrier shape. For example, for a nearly rectangular potential barrier, the exponent that determines the classical lifetime (155) depends (linearly) only on the barrier *height* U_0, while that defining the quantum lifetime (156) is proportional to the barrier *width* and to the square root of U_0. However, in the important case of "soft" potential profiles, which are typical for the case of emerging (or nearly disappearing) quantum wells (Fig. 11), the classical and quantum results are closely related.

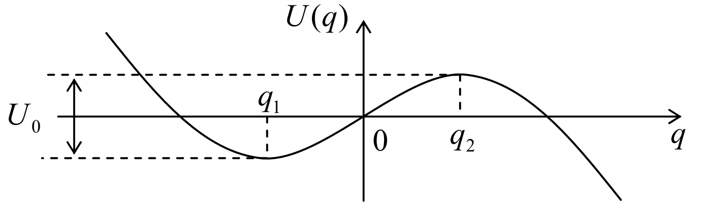


Fig. 5.11. Cubic-parabolic potential profile and its parameters.

---

[70] An example of such an equation, for the particular case of a harmonic oscillator, is given by QM Eq. (7.214). The Fokker-Planck equation, of course, can give only its classical limit, with n, $n_e >> 1$.

[71] H. Kramers, Physica 7, 284 (1940); see also the model solution of Problem 27.

[72] See, e.g., the review paper by O. Mel'nikov, *Phys. Repts*. **209**, 1 (1991), and references therein.

[73] See, e.g., QM Secs. 2.4-2.6.

Indeed, such potential profile $U(q)$ may be well approximated by four leading terms of its Taylor expansion, with the highest term proportional to $(q - q_0)^3$, near any point q_0 in the vicinity of the well. In this approximation, the second derivative d^2U/dq^2 vanishes at the inflection 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, as this is done in Fig. 11, we may reduce the approximation to just two terms:[74]

$$U(q) = aq - \frac{b}{3}q^3. \tag{5.157}$$

(For the particle's escape into the positive direction of the q-axis, we should have $a, b > 0$.) An easy calculation gives all essential parameters of this cubic parabola: the positions of its minimum and maximum:

$$q_2 = -q_1 = (a/b)^{1/2}, \tag{5.158}$$

the barrier height over the well's bottom:

$$U_0 \equiv U(q_2) - U(q_1) = \frac{4}{3}\left(\frac{a^3}{b}\right)^{1/2}, \tag{5.159}$$

and the effective spring constants at these points:

$$\kappa_1 = \kappa_2 \equiv \left|\frac{d^2U}{dq^2}\right|_{q_{1,2}} = 2(ab)^{1/2}. \tag{5.160}$$

Hence for this potential profile, Eq. (155) may be rewritten as

Soft well: thermal lifetime

$$\boxed{\tau = \frac{2\pi}{\omega_0}\exp\left\{\frac{U_0}{T}\right\}, \qquad \text{with } \omega_0^2 \equiv \frac{2(ab)^{1/2}}{m}.} \tag{5.161}$$

On the other hand, for the same profile, the WKB approximation (156) (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 lower than the barrier height U_0) yields[75]

Soft well: quantum lifetime

$$\boxed{\tau_{\mathrm{Q}} = \frac{2\pi}{\omega_0}\left(\frac{\hbar\omega_0}{864U_0}\right)^{1/2}\exp\left\{\frac{36}{5}\frac{U_0}{\hbar\omega_0}\right\}.} \tag{5.162}$$

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

$$T_{\mathrm{c}} = \frac{36}{5}\hbar\omega_0 \equiv 7.2\,\hbar\omega_0. \tag{5.163}$$

This expression for the *cubic*-parabolic barrier may be compared with a similar crossover for a *quadratic*-parabolic barrier,[76] for which $T_{\mathrm{c}} = 2\pi\,\hbar\omega_0 \approx 6.28\,\hbar\omega_0$. We see that the numerical factors for

---

[74] As a reminder, a similar approximation arises for the $P(V)$ function, at the analysis of the van der Waals model near the critical temperature – see Problem 4.6.

[75] 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-exponential factor requires taking this difference, $\hbar\omega_0/2$, into account – see, e.g., the model solution of QM Problem 2.43.

[76] See, e.g., QM Sec. 2.4.

the quantum-to-classical crossover temperature for these two different soft potential profiles are close to each other – and much larger than 1, which could result from a naïve estimate.

5.8. Back to the correlation function

Unfortunately, I will not have time/space to either derive or even 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 the solution $w(\mathbf{q}, \mathbf{p}, t)$ of the Fokker-Planck equation for a system provides full statistical information about it, one may wonder how it may be used to find its temporal characteristics that 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 clear – cf. Eq. (2.11):

$$\langle f[\mathbf{q}(t),\mathbf{p}(t)]\rangle = \int f(\mathbf{q},\mathbf{p})w(\mathbf{q},\mathbf{p},t)d^3q d^3p\,, \tag{5.164}$$

but what if the function depends on variables taken at *different* times, for example as in the correlation function $K_f(\tau)$ defined by Eq. (48)?

To answer this question, let us start from the discrete-variable case when Eq. (164) takes the form (2.7), which, for our current purposes, may be rewritten as

$$\langle f(t)\rangle = \sum_m f_m W_m(t)\,. \tag{5.165}$$

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 implies that the 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* to measure outcome m at moment t, *and* outcome m' at moment t'. The joint probability may be represented 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 may take $t = 0$, i.e. look for the conditional probability as the solution, $W_{m'}(\tau)$, of the equation describing the system's probability evolution, at time $\tau = t' - t$ (rather than t'), with the special initial condition

$$W_{m'}(0) = \delta_{m',m}\,. \tag{5.166}$$

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 $W_m(0)$ we may take the stationary probability distribution $W_m(\infty)$, independent of the initial conditions, which may be found as the same special solution, but at time $\tau \to \infty$. As a result, we get

Correlation function: discrete system

$$\boxed{\langle f(t)f(t+\tau)\rangle = \sum_{m,m'} f_m W_m(\infty) f_{m'} W_{m'}(\tau)\,.} \tag{5.167}$$

This expression looks simple, but note that this recipe requires solving the time evolution equations for each $W_{m'}(\tau)$ for all possible initial conditions (166). To see how this recipe works in practice, let us revisit the simplest two-level system (see, e.g., Fig. 4.13, which is reproduced in Fig. 12 below in a notation more convenient for our current purposes), and calculate the correlation function of its energy fluctuations.

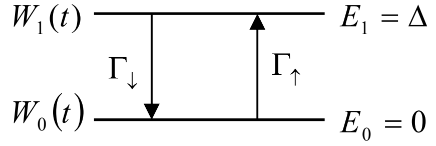


Fig. 5.12. Dynamics of a two-level system.

The stationary probabilities of the system's states (i.e. their probabilities for $\tau \to \infty$) have been calculated in problems of Chapter 2, and then again in Sec. 4.4 – see Eq. (4.68). In our current notation (Fig. 12),

$$W_0(\infty) = \frac{1}{1+e^{-\Delta/T}}, \qquad W_1(\infty) = \frac{1}{e^{\Delta/T}+1},$$
$$\text{so that} \quad \langle E(\infty)\rangle = W_0(\infty)\times 0 + W_1(\infty)\times\Delta = \frac{\Delta}{e^{\Delta/T}+1}. \tag{5.168}$$

To calculate the conditional probabilities $W_{m'}(\tau)$ with the initial conditions (166) (according to Eq. (168), we need all four of them, for $\{m, m'\} = \{0, 1\}$), we may use the master equations (4.100), in our current notation reading

$$\frac{dW_1}{d\tau} = -\frac{dW_0}{d\tau} = \Gamma_\uparrow W_0 - \Gamma_\downarrow W_1. \tag{5.169}$$

Since Eq. (169) conserves the total probability, $W_0 + W_1 = 1$, only one probability (say, W_1) is an independent variable, and for it, Eq. (169) gives a simple, linear differential equation

$$\frac{dW_1}{d\tau} = \Gamma_\uparrow(1-W_1) - \Gamma_\downarrow W_1 \equiv \Gamma_\uparrow - \Gamma_\Sigma W_1, \qquad \text{where } \Gamma_\Sigma \equiv \Gamma_\uparrow + \Gamma_\downarrow, \tag{5.170}$$

which may be readily integrated for an arbitrary initial condition:

$$W_1(\tau) = W_1(0)e^{-\Gamma_\Sigma\tau} + W_1(\infty)\left(1-e^{-\Gamma_\Sigma\tau}\right), \tag{5.171}$$

where $W_1(\infty)$ is given by the second of Eqs. (168). (It is straightforward to verify that the solution for $W_0(\tau)$ may be represented in a form similar to Eq. (171), with the corresponding replacement of the state index.)

Now everything is ready to calculate the average $\langle E(t)E(t+\tau)\rangle$ using Eq. (167), with $f_{m,m'} = E_{0,1}$. Thanks to our (smart :-) choice of the energy reference, of the four terms in the double sum (167), all three terms that include at least one factor $E_0 = 0$ vanish, and we have only one term left to calculate:

$$\langle E(t)E(t+\tau)\rangle = E_1W_1(\infty)E_1W_1(\tau)\Big|_{W_1(0)=1} = E_1^2W_1(\infty)\left[W_1(0)e^{-\Gamma_\Sigma\tau} + W_1(\infty)\left(1-e^{-\Gamma_\Sigma\tau}\right)\right]_{W_1(0)=1}$$
$$= \frac{\Delta^2}{e^{\Delta/T}+1}\left[e^{-\Gamma_\Sigma\tau} + \frac{1}{e^{\Delta/T}+1}\left(1-e^{-\Gamma_\Sigma\tau}\right)\right] \equiv \frac{\Delta^2}{\left(e^{\Delta/T}+1\right)^2}\left(1+e^{\Delta/T}e^{-\Gamma_\Sigma\tau}\right). \tag{5.172}$$

From here and the last of Eqs. (168), the correlation function of energy fluctuations is[77]

---

[77] The step from the first line of Eq. (173) to its second line utilizes the fact that our system is stationary, so $\langle E(t+\tau)\rangle = \langle E(t)\rangle = \langle E(\infty)\rangle = \text{const}$.

$$K_E(\tau) \equiv \left\langle \tilde{E}(t)\tilde{E}(t+\tau)\right\rangle = \left\langle \left(E(t) - \langle E(t)\rangle\right)\left(E(t+\tau) - \langle E(t)\rangle\right)\right\rangle$$
$$= \langle E(t)E(t+\tau)\rangle - \langle E(\infty)\rangle^2 = \Delta^2 \frac{e^{\Delta/T}}{\left(e^{\Delta/T}+1\right)^2} e^{-\Gamma_\Sigma \tau}, \qquad (5.173)$$

so its variance, equal to $K_E(0)$, does not depend on the transition rates $\Gamma_\uparrow$ and $\Gamma_\downarrow$. However, since the rates have to obey the detailed balance relation (4.103), $\Gamma_\downarrow/\Gamma_\uparrow = \exp\{\Delta/T\}$, for this variance we may formally write

$$\frac{K_E(0)}{\Delta^2} = \frac{e^{\Delta/T}}{\left(e^{\Delta/T}+1\right)^2} = \frac{\Gamma_\downarrow/\Gamma_\uparrow}{\left(\Gamma_\downarrow/\Gamma_\uparrow+1\right)^2} \equiv \frac{\Gamma_\uparrow\Gamma_\downarrow}{\left(\Gamma_\uparrow+\Gamma_\downarrow\right)^2} \equiv \frac{\Gamma_\uparrow\Gamma_\downarrow}{\Gamma_\Sigma^2}, \qquad (5.174)$$

so Eq. (173) may be represented in a simpler form:

Energy fluctuations: two-level system

$$\boxed{K_E(\tau) = \Delta^2 \frac{\Gamma_\uparrow\Gamma_\downarrow}{\Gamma_\Sigma^2} e^{-\Gamma_\Sigma\tau}.} \qquad (5.175)$$

We see that the correlation function of energy fluctuations decays exponentially with time, with the rate Γ_Σ. Now using the Wiener-Khinchin theorem (58) to calculate its spectral density, we get

$$\boxed{S_E(\omega) = \frac{1}{\pi}\int_0^\infty \Delta^2 \frac{\Gamma_\uparrow\Gamma_\downarrow}{\Gamma_\Sigma^2} e^{-\Gamma_\Sigma\tau} \cos\omega\tau \, d\tau = \frac{\Delta^2}{\pi\Gamma_\Sigma} \frac{\Gamma_\uparrow\Gamma_\downarrow}{\Gamma_\Sigma^2+\omega^2}.} \qquad (5.176)$$

Such Lorentzian dependence on frequency 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 states of two-level systems with an exponentially-broad statistical distribution of the transition rates $\Gamma_{\uparrow\downarrow}$. Such a broad distribution follows from the Kramers formula (155), which is approximately valid for the lifetimes of both states of systems with double-well potential profiles (Fig. 13), for a statistical ensemble with a smooth statistical distribution of the energy barrier heights U_0. Such profiles are typical, in particular, for electrons in disordered (amorphous) solid-state materials, which indeed feature high 1/*f* noise.

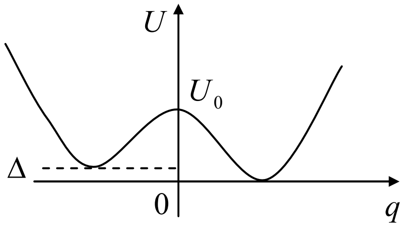


Fig. 5.13. Typical double-well potential profile.

Returning to the Fokker-Planck equation, we may use the following evident generalization of Eq. (167) to the continuous-variable case:

Correlation function: continuous system

$$\boxed{\langle f(t)f(t+\tau)\rangle = \int d^3q\,d^3p \int d^3q'd^3p' \; f(\mathbf{q},\mathbf{p})w(\mathbf{q},\mathbf{p},\infty)f(\mathbf{q}',\mathbf{p}')w(\mathbf{q}',\mathbf{p}',\tau),} \qquad (5.177)$$

were both probability densities are particular values of the equation's solution with the delta-functional initial condition

$$w(\mathbf{q}',\mathbf{p}',0)=\delta(\mathbf{q}'-\mathbf{q})\delta(\mathbf{p}'-\mathbf{p})\,. \tag{5.178}$$

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

$$\langle f(t)f(t+\tau)\rangle=\int d^3q\int d^3q'\, f(\mathbf{q})w(\mathbf{q},\infty)f(\mathbf{q}')w(\mathbf{q}',\tau)\,, \tag{5.179}$$

$$w(\mathbf{q}',0)=\delta(\mathbf{q}'-\mathbf{q})\,. \tag{5.180}$$

To see this formalism in action, let us use it to calculate the correlation function $K_q(\tau)$ of a *linear relaxator*, i.e. an overdamped 1D harmonic oscillator with $m\omega_0 << \eta$. In this limit, as Eq. (65) shows, the oscillator's coordinate averaged over the ensemble of environments obeys a linear equation,

$$\eta\langle\dot{q}\rangle+\kappa\langle q\rangle=0\,, \tag{5.181}$$

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

$$\langle q\rangle(t)=q_0e^{-\Gamma t}\,. \tag{5.182}$$

The deterministic equation (181) corresponds to the quadratic potential energy $U(q) = \kappa q^2/2$, so the 1D version of the corresponding Smoluchowski equation (122) takes the form

$$\eta\frac{\partial w}{\partial t}=\kappa\frac{\partial}{\partial q}(wq)+T\frac{\partial^2 w}{\partial q^2}\,. \tag{5.183}$$

It is straightforward to check, by substitution, that this equation, rewritten for the function $w(q',\tau)$, with the 1D version of the delta-functional initial condition (180), $w(q',0) = \delta(q'-q)$, is satisfied with a Gaussian function:

$$w(q',\tau)=\frac{1}{(2\pi)^{1/2}\delta q(\tau)}\exp\left\{-\frac{\left(q'-\langle q\rangle(\tau)\right)^2}{2\delta q^2(\tau)}\right\}, \tag{5.184}$$

with its center $\langle q\rangle(\tau)$ moving in accordance with Eq. (182), and a time-dependent variance

$$\delta q^2(\tau)=\delta q^2(\infty)\left(1-e^{-2\Gamma\tau}\right),\qquad \text{where } \delta q^2(\infty)=\langle q^2\rangle=\frac{T}{\kappa}\,. \tag{5.185}$$

(As a sanity check, the last equality coincides with the equipartition theorem's result.) Finally, the first probability under the integral in Eq. (179) may be found from Eq. (184) in the limit $\tau \to \infty$ (in which $\langle q\rangle(\tau) \to 0$), by replacing q' with q:

$$w(q,\infty)=\frac{1}{(2\pi)^{1/2}\delta q(\infty)}\exp\left\{-\frac{q^2}{2\delta q^2(\infty)}\right\}. \tag{5.186}$$

Now all ingredients of the recipe (179) are ready, and we can spell it out, for $f(q) = q$, as

$$\langle q(t)q(t+\tau)\rangle=\frac{1}{2\pi\delta q(\tau)\delta q(\infty)}\int_{-\infty}^{+\infty}dq\int_{-\infty}^{+\infty}dq'\; q\exp\left\{-\frac{q^2}{2\delta q^2(\infty)}\right\}q'\exp\left\{-\frac{\left(q'-qe^{-\Gamma\tau}\right)^2}{2\delta q^2(\tau)}\right\}. \tag{5.187}$$

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

$$\langle q(t)q(t+\tau)\rangle = \frac{1}{2\pi\delta q(\tau)\delta q(\infty)}\int_{-\infty}^{+\infty} q\exp\left\{-\frac{q^2}{2\delta q^2(\infty)}\right\}dq\int_{-\infty}^{+\infty}\left(q'' + qe^{-\Gamma\tau}\right)\exp\left\{-\frac{q''^2}{2\delta q^2(\tau)}\right\}dq''. \quad (5.188)$$

The internal integral of the first term in the parentheses equals zero (as that of an odd function in symmetric integration limits), while that with the second term is the standard Gaussian integral, so:

$$\langle q(t)q(t+\tau)\rangle = \frac{1}{(2\pi)^{1/2}\delta q(\infty)}e^{-\Gamma\tau}\int_{-\infty}^{+\infty} q^2\exp\left\{-\frac{q^2}{2\delta q^2(\infty)}\right\}dq \equiv \frac{2T}{\pi^{1/2}\kappa}e^{-\Gamma\tau}\int_{-\infty}^{+\infty}\xi^2\exp\left\{-\xi^2\right\}d\xi\,. \quad (5.189)$$

The last integral[78] equals $\pi^{1/2}/2$, so taking into account that for this stationary system centered at the coordinate origin, $\langle q(\infty)\rangle = 0$, we finally get a very simple result:

Correlation function: linear relaxator

$$\boxed{K_q(\tau) \equiv \langle\tilde{q}(t)\tilde{q}(t+\tau)\rangle = \langle q(t)q(t+\tau)\rangle - \langle q(\infty)\rangle^2 = \langle q(t)q(t+\tau)\rangle = \frac{T}{\kappa}e^{-\Gamma\tau}.} \quad (5.190)$$

As a sanity check, for $\tau = 0$, it yields $K_q(0) \equiv \langle q^2\rangle = T/\kappa$, in accordance with Eq. (185). As τ 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 $S_{\mathcal{F}}(\omega)$ given by Eq. (73a):[79]

$$K_q(\tau) = 2\int_0^\infty S_q(\omega)\cos\omega\tau\, d\omega = 2\int_0^\infty \frac{\eta T}{\pi}\frac{1}{\kappa^2 + (\eta\omega)^2}\cos\omega\tau\, d\omega \equiv 2\frac{T\Gamma}{\pi}\int_0^\infty \frac{\cos\xi}{(\Gamma\tau)^2 + \xi^2}\, d\xi = \frac{T}{\kappa}e^{-\Gamma\tau}. \quad (5.191)$$

This example illustrates the fact that for fluctuations in linear systems (and small fluctuations in nonlinear systems) the Langevin approach is usually much simpler than 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-Planck and Smoluchowski equations give a quantitative description of the time evolution of nonlinear Brownian systems with dissipation in the *classical* limit. The description of the corresponding properties of such dissipative ("open") and nonlinear *quantum* systems is more complex,[80] and only a few simple problems of their theory have been solved analytically so far,[81] typically using particular models of the environment, e.g., as a large set of harmonic oscillators with various statistical distributions of their parameters, each leading to a specific function $\chi(\omega)$ for the generalized susceptibility.

---

[78] See, e.g., MA Eq. (6.9c).

[79] The involved table integral may be found, e.g., in MA Eq. (6.11).

[80] See, e.g., QM Sec. 7.6.

[81] See, e.g., the solutions of the 1D Kramers problem for quantum systems with low damping by A. Caldeira and A. Leggett, *Phys. Rev. Lett.* **46**, 211 (1981), and with high damping by A. Larkin and Yu. Ovchinnikov, *JETP Lett.* **37**, 382 (1983).

5.10. Exercise problems

5.1. By treating 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 the average $\langle k \rangle$ and the r.m.s. fluctuation δk. Compare the results with those for the ensemble of randomly selected decimal integers 0, 1, 2…, 9.

5.2. An ideal classical gas of N similar particles fills a spherical cavity of radius R. Calculate the variance of fluctuations of the position $\mathbf{r}$ of its center of mass, in equilibrium.

5.3. Calculate the variance of fluctuations of a magnetic moment $\mathbf{m}$ placed into an external magnetic field $\mathscr{H}$, within the same two models as in Problem 2.4:

(i) a quantum spin-½ with a gyromagnetic ratio γ, and
(ii) a classical magnetic moment $\mathbf{m}$ of a fixed magnitude m_0 but an arbitrary orientation,

both in thermal equilibrium at temperature T. Compare the results.[82]

Hint: Mind all three Cartesian components of the vector $\mathbf{m}$.

5.4. For a field-free two-site Ising system with energy values $E_m = -Js_1s_2$, in thermal equilibrium at temperature T, calculate the variance of energy fluctuations. Explore the low-temperature and high-temperature limits of the result.

5.5. In a system in thermodynamic equilibrium with fixed T and μ, both the number N of particles and the internal energy E may fluctuate. Express the mutual correlation factor of these fluctuations via the average of E. Spell out the result for an ideal classical gas of $N >> 1$ particles.

5.6. As was mentioned in Sec. 2, the variance of energy fluctuations in a system with fixed T and μ (i.e. a member of a *grand canonical* ensemble) is generally different from that in a similar system in which T and N are fixed, i.e. a member of a *canonical* ensemble. Calculate and interpret the difference.

5.7. For a uniform three-site Ising ring with ferromagnetic coupling (and no external field), in thermal equilibrium at temperature T, calculate the correlation coefficients $K_s \equiv \langle s_k s_{k'} \rangle$ for both $k = k'$ and $k \neq k'$.

5.8.* For a field-free 1D Ising system of $N >> 1$ "spins", in thermal equilibrium at temperature T, calculate the correlation coefficient $K_s \equiv \langle s_l s_{l+n} \rangle$, where l and $(l + n)$ are the numbers of two specific spins in the chain.

Hint: Consider a mixed partial derivative of the statistical sum calculated in Problem 4.21 for an Ising chain with an arbitrary set of J_k, over a part of these parameters.

---

[82] Note that these two cases may be considered as the non-interacting limits of, respectively, the Ising model (4.23) and the classical limit of the Heisenberg model (4.21), whose analysis within the Weiss approximation was the subject of Problems 4.22 and 4.23.

5.9. Within the framework of the Weiss molecular-field approximation, calculate the variance of spin fluctuations in the d-dimensional Ising model. Use the result to derive the conditions of quantitative validity of the approximation.

5.10. Calculate the variance of energy fluctuations in a quantum harmonic oscillator with frequency ω, in thermal equilibrium at temperature T, and express it via the average energy.

5.11. The spontaneous electromagnetic field inside a closed volume V is in thermal equilibrium at temperature T. Assuming that V is sufficiently large, calculate the variance of fluctuations of the total energy of the field, and express the result via its average energy and temperature. How large should the volume V be for your results to be quantitatively valid? Evaluate this limitation for room temperature.

5.12. Express the r.m.s. uncertainty of the occupancy N_k of a certain energy level ε_k by non-interacting:

(i) classical particles,
(ii) fermions, and
(iii) bosons,

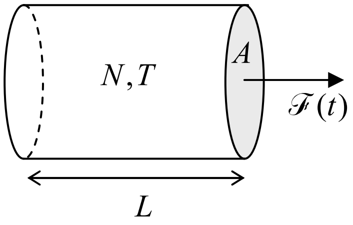
in thermodynamic equilibrium, via the level's average occupancy $\langle N_k\rangle$, and compare the results.

5.13. Write a general expression for the variance of the number of particles in the ideal gases of bosons and fermions, at fixed V, T, and μ. Spell out the result for the degenerate Fermi gas.

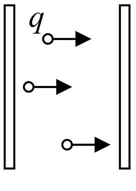
5.14. Express the variance of the number of particles, $\langle \tilde{N}^2\rangle_{V,T,\mu}$, of a single-phase system in equilibrium, via its isothermal compressibility $\kappa_T \equiv -(\partial V/\partial P)_{T,N}/V$.

5.15.* Calculate the low-frequency spectral density of fluctuations of the pressure P of an ideal classical gas, in thermal equilibrium at temperature T, and estimate their variance. Compare the former result with the solution of Problem 3.2.

Hints: You may consider a cylindrically-shaped container of volume $V = LA$ (see the figure on the right), and start by using the Maxwell distribution of velocities to calculate the spectral density of the force $\mathscr{F}(t)$ exerted by the confined particles on its plane lid of area A, approximating the force with a delta-correlated process.



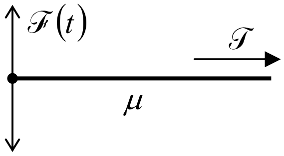
5.16. Calculate the low-frequency spectral density of fluctuations of the electric current $I(t)$ due to the random passage of charged particles between two conducting electrodes – see the figure on the right. Assume that the particles are emitted, at random times, by one of the electrodes, and are fully absorbed by the counterpart electrode. Can your result be mapped onto some aspect of the electromagnetic blackbody radiation?



Hint: For the current $I(t)$, use the same delta-correlated-process approximation as for the force $\mathscr{F}(t)$ in the previous problem.

5.17. Perhaps the simplest model of the diffusion is the 1D *discrete random walk*: each time interval τ, a particle leaps, with equal probability, to any of two adjacent sites of a 1D lattice with a spatial period a. Prove that the particle's displacement during a time interval $t >> \tau$ obeys Eq. (77), and calculate the corresponding diffusion coefficient D.

5.18.[83] A long uniform string, of mass μ per unit length, is attached to a firm support and stretched with a constant force ("tension") $\mathscr{T}$ – see the figure on the right. Calculate the spectral density of the random force $\mathscr{F}(t)$ exerted by the string on the support point, within the plane normal to its length, in thermal equilibrium at temperature T.



Hint: You may assume that the string is so long that transverse waves propagating along it from the support point never come back.

5.19.[84] Each of the two 3D isotropic harmonic oscillators, with mass m, resonance frequency ω_0, and damping coefficient $\delta > 0$, has the electric dipole moment $\mathbf{d} = q\mathbf{s}$, where $\mathbf{s}$ is the vector of the oscillator's displacement from its equilibrium position. Use the Langevin formalism to calculate the average potential of electrostatic interaction (a particular case of the so-called *London dispersion force*) of these oscillators separated by distance $r >> (T/m)^{1/2}/\omega_0$, in thermal equilibrium at temperature $T >> \hbar\omega_0$. Also, explain why the approach used to solve a very similar Problem 2.18 is not directly applicable to this case.

Hint: You may like to use the following integral: $\int_0^\infty \frac{1-\xi^2}{\left[\left(1-\xi^2\right)^2+\left(a\xi\right)^2\right]^2}d\xi = \frac{\pi}{4a}$.

5.20.* Within the van der Pol approximation,[85] calculate the major statistical properties of small fluctuations of classical self-oscillations (including their linewidth), for:

(i) a free ("autonomous") run of the oscillator, and
(ii) its phase being locked by an external sinusoidal force,

assuming that the fluctuations are caused by a noise with a smooth spectral density $S_f(\omega)$.

5.21. Calculate the correlation function of the coordinate of a 1D harmonic oscillator with low damping, in thermal equilibrium. Compare the solution with that of the previous problem.

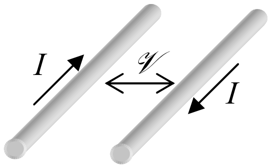
5.22. A lumped electric circuit consisting of a capacitor C shortened with an Ohmic resistor R is in thermal equilibrium at temperature T. Use two different approaches to calculate the variance of the thermal fluctuations of the capacitor's electric charge Q. Estimate the effect of quantum fluctuations.

---

[83] This problem, conceptually important for the quantum mechanics of open systems, was also given in Chapter 7 of the QM part of this series.

[84] This system, with an arbitrary temperature, was the subject of QM Problem 7.6, with QM Problem 5.15 serving as the background. However, the method used in the model solutions of those problems requires one to prescribe, to the oscillators, different frequencies ω_1 and ω_2 at first, and only after this more general problem has been solved, pursue the limit $\omega_1 \to \omega_2$, while neglecting dissipation altogether. The goal of this problem is to show that the result of that solution is valid even at non-zero damping.

[85] See, e.g., CM Secs. 5.2-5.5. Note that in quantum mechanics, a similar approach is called the *rotating-wave approximation* (RWA) – see, e.g., QM Secs. 6.5, 7.6, 9.2, and 9.4.

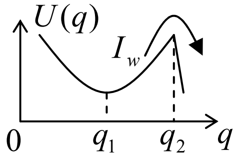


5.23. Consider a very long, uniform, two-wire transmission line (see the figure on the right) with a wave impedance $\mathscr{Z}$, which allows the propagation of TEM electromagnetic waves with negligible attenuation. Calculate the variance $\langle \mathscr{V}^2 \rangle_{\Delta\nu}$ of spontaneous fluctuations of the voltage $\mathscr{V}$ between the wires within a small interval $\Delta\nu$ of cyclic frequencies, in thermal equilibrium at temperature T.

Hint: As an E&M reminder,[86] in the absence of dispersive materials, TEM waves propagate with a frequency-independent velocity, and with the voltage $\mathscr{V}$ and current I (see the figure above) related as $\mathscr{V}(x,t)/I(x,t) = \pm\mathscr{Z}$, where $\mathscr{Z}$ is the line's wave impedance.

5.24. Now consider a similar long transmission line but terminated, at one end, with an impedance-matching Ohmic resistor $R = \mathscr{Z}$. Calculate the variance $\langle \mathscr{V}^2 \rangle_{\Delta\nu}$ of the voltage across the resistor, and discuss the relation between the result and the Nyquist formula (81b), including numerical factors.

Hint: A termination with resistance $R = \mathscr{Z}$ absorbs incident TEM waves without reflection.



5.25. An overdamped classical 1D particle escapes from a potential well with a smooth bottom but a sharp top of the barrier – see the figure on the right. Perform the necessary modification of the Kramers formula (139).

5.26.* Similar particles, whose spontaneous electric dipole moments $\mathscr{p}$ have a field-independent magnitude $\mathscr{p}_0$, are uniformly distributed in space with a density n so low that their mutual interaction is negligible. Each particle may rotate without substantial inertia but under a kinematic friction torque proportional to its angular velocity. Use the Smoluchowski equation to calculate the complex dielectric constant $\varepsilon(\omega)$ of such a medium, in thermal equilibrium at temperature T, for a weak, linearly-polarized rf electric field.

5.27.* Prove that for systems with relatively low inertia (i.e. relatively high damping), at not very high temperatures, the Fokker-Planck equation (149) reduces to the Smoluchowski equation (122) – in the sense described by Eq. (153) and the accompanying statement.

5.28.* Use the 1D version of the Fokker-Planck equation (149) to prove the solution (156) of the Kramers problem.

5.29. A constant external torque, applied to a 1D mechanical pendulum with mass m and length l, has displaced it by angle $\varphi_0 < \pi/2$ from the vertical position. Calculate the average rate of the pendulum's rotation induced by relatively small thermal fluctuations of temperature T.

5.30. A classical particle may occupy any of N similar sites. Its weak interaction with the environment induces random uncorrelated incoherent jumps from the occupied site to any other site, with the same time-independent rate Γ. Calculate the correlation function and the spectral density of fluctuations of the instant occupancy $n(t)$ (equal to either 1 or 0) of a site.

---

[86] See, e.g., EM Sec. 7.6.