

1

Thermal processes and diffusion

1.1 Boltzmann, Einstein, and molecules

During the 19th century an argument raged about the existence of atoms and molecules, and at the centre of that debate was the Austrian physicist Ludwig Boltzmann and his radical idea that temperature was manifest in the restless, random motions of atoms and molecules [1]. Boltzmann proposed that for every independent type of motion, there would be an equipartition of average thermal energy proportional to temperature. Two centuries earlier, Swiss physicist Daniel Bernoulli had used simple Newtonian mechanics to explain gas pressure in terms of particles colliding with the walls. When Boltzmann applied his equipartition idea, taking suitable statistical averages, he not only explained the gas laws, but also the laws of thermodynamics.

It is hardly surprising that he was attacked. After all, the laws of mechanics are reversible. Might that mean, given the mechanical description, that heat could flow from a colder temperature to a hotter temperature under its own volition, something we never observe? Boltzmann, in rebuttal, explained that while such a violation of the second law might be possible, it was extremely improbable, given the huge number of molecules taking part. But in systems with small numbers of molecules, reversal of the law might be seen and, indeed, modern experiments on small systems have revealed precisely that possibility [2]. Today, Boltzmann's statistical mechanics is one of the cornerstones of modern physics, and we honour his work by naming the constant of proportionality for equipartition *Boltzmann's constant*, k_B .

From the standpoint of the 21st century, the idea of a theoretical model that we test indirectly through its predictions seems entirely reasonable. But in the 19th century, theoretical physics was in its infancy, and there were those who found it distasteful to speak of atoms or molecules when no one had seen such entities. That distaste was exemplified by Ernst Mach, another Austrian, after whom we name the speed of sound. To Mach, scientific laws had to be based on what was directly measurable, not on fanciful theory—on the surface a reasonable position that might well resonate with people of common sense. But science is, as Lewis Wolpert has reminded us [3], a means of discovering truths that defy common sense. It is, after all, common sense that the sun revolves around the earth.

The philosophical disagreement was settled in 1905 by another theoretical physicist, Albert Einstein, whose work was to vindicate Boltzmann completely. That work was on Brownian motion [4], the curious random migration of pollen grains observed under the microscope in 1827 by the botanist Robert Brown [5]. Einstein boldly postulated that the pollen grain, large enough to be seen in the optical microscope, was small enough

2 Thermal processes and diffusion

to be buffeted by the random thermal motion of the surrounding water molecules. Einstein used Boltzmann's thermal energy predictions and came up with a rule by which one would be able to estimate Avogadro's number simply by observing how far the pollen grain moved over a given time. That experimental work was completed in 1908 by the French physicist Jean Perrin.¹

Mach's discomfort with the indirect observation of atoms and molecules is sharply challenged by this present book, in which we describe the measurement of molecular motion, and in particular Brownian motion, via the medium of radio waves emitted from atomic nuclei. Magnetic resonance observations of molecular translational motion draw on quantum mechanics and statistical mechanics in equal parts, reason enough for its fascinating appeal as an experimental technique, and excuse enough for digressions on these beautiful strands of physics as our story unfolds. Of course, both quantum mechanics and statistical mechanics are covered in great detail in hundreds of other texts, and one might ask therefore, what justification could there be in providing introductory material on those subjects in this book? The answer is simply this: a focus on magnetic resonance studies of molecular motion provides a particular context in which non-experts can learn about these other areas of physics. And for the physics experts, the brief descriptions given here provide a compact selection of the key ideas.

1.2 Statistical physics and ensembles

1.2.1 Temperature and entropy

In the case of gas molecules, thermal energy is principally stored as kinetic energy of translation, and in an 'ideal monatomic gas' entirely so.² Think of one gas molecule of known kinetic energy, bouncing elastically within a container. The motion, for example the time dependent velocity, is describable by simple classical physics. But for two or more molecules, where collisions are permitted, the description is immensely complicated, while for very large numbers of molecules, only the language of statistics will suffice to describe the distribution of possible velocities that our 'ensemble' of molecules will exhibit. Strangely, the resort to statistics, rather than blurring our ability to describe the behaviour of the gas, will, if the number of molecules is sufficiently large, lead to a high degree of certainty in the description of certain ensemble average properties. The enormous size of Avogadro's number means that even for quite small physical samples, these properties will obey relationships so powerful as to be described as *thermodynamic laws*.

In this section we examine two thermodynamic properties, temperature and entropy, which underpin the process of diffusion. In order to do this, and to simply illustrate the role of statistics, we choose as our example, not gas molecules with kinetic energy, but an example of particles that store energy in an even simpler manner—an example with obvious connection to magnetic resonance. This is the two-state case of quantised spin- $\frac{1}{2}$ particles, either up or down in a magnetic field. Statistical physics

¹Perrin won the Nobel Prize in 1926. Einstein won it in 1921 for another of his 1905 papers, on the photoelectric effect [6].

²Even for polyatomic gases, the other storage modes of molecular rotation and vibration are not activated at low temperatures.

rests upon the ideas of systems of *quantum states*, so this example is a particularly helpful one. The use of quantum states as a starting point for an understanding of statistical physics is central to the text by Kittel and Kroemer [7]. Sections 1.2.1 through 1.2.4 paraphrase Kittel and Kroemer's insight, and readers are referred to that remarkable book for further illumination.

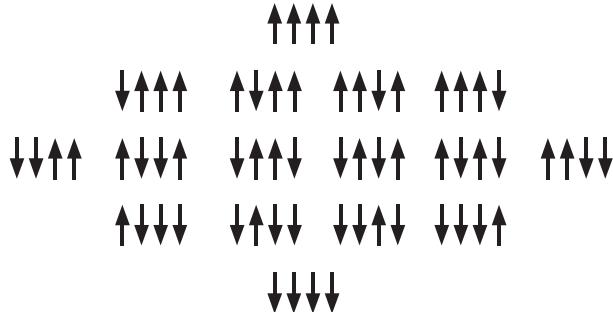


Fig. 1.1 Possible states of four independent spin- $\frac{1}{2}$ particles showing the multiplicities for different numbers of up and down spins.

Let us begin by considering probabilities without reference to energy, easily achieved by having no magnetic field present. As an example, imagine four independent spin particles that may each be either up (\uparrow) or down (\downarrow). As shown in Fig. 1.1 there is one way that all four spins can be up, four ways one can be down, six ways there can be two up and two down, four ways one can be up, and one way they can all be down. These numbers of possibilities are called *multiplicities*, and they follow a binomial pattern. In particular, if there are N spins with N_{\uparrow} up and N_{\downarrow} down, such that the difference between up and down spins is the spin excess, $2x$, then the multiplicity is

$$g(N, x) = \frac{N!}{N_{\uparrow}! N_{\downarrow}!} = \frac{N!}{(\frac{1}{2}N + x)! (\frac{1}{2}N - x)!} \quad (1.1)$$

Of course, x ranges between $-N/2$ and $N/2$. Notice that the multiplicity function is peaked about $x = 0$ and that as the total number, N , of spins grows, this peak becomes sharper and sharper in relative terms. Figure 1.2 compares the multiplicity functions of 4, 10, and 100 spins.

Underpinning the present discussion is the idea that all accessible quantum states of a system are equally likely. In the absence of a magnetic field, all spin excess values would be accessible and the probability of finding each value would be given by the multiplicity of states for each configuration. For N spins with spin excess x , that multiplicity is $g(N, x)$. It is clear that as N grows, the probability distribution associated with the spins becomes more like a certainty that $x = 0$. In fact, the spread in x values grows as $N^{1/2}$, so that the spread in x/N decreases as $N^{-1/2}$. Hence this sharpness is relative to the available range of x between $-N/2$ and $N/2$. This illustration of how, as the numbers of particles in our system increases, near certainty emerges from mere probability, lies at the heart of statistical physics, at the heart of

4 Thermal processes and diffusion

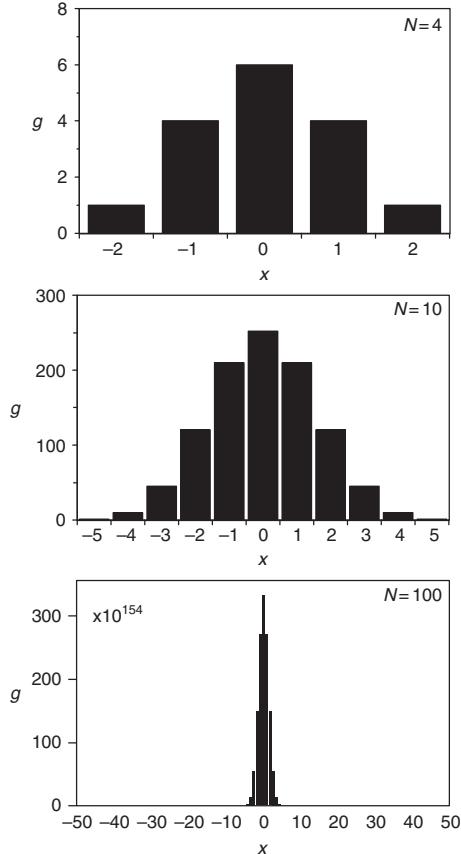


Fig. 1.2 Multiplicities $g(N, x)$ for N values of 5, 10 and 100.

Boltzmann's understanding [8]. The mathematics also simplifies. Note that large N results in another simplification in that one may approximate eqn 1.1 by

$$g(N, x) = g(N, 0) \exp\left(\frac{-2x^2}{N}\right) \quad (1.2)$$

Let us now apply the magnetic field B to the spin system. Each spin has a magnetic moment m such that for spins up along the magnetic field (low energy state) the energy is $-mB$, while for down spins (high energy state) the energy is $+mB$. If the system happened to have spin excess x , then the total system energy would be $U = -2xmB$. The interesting question concerns the rules governing how different systems of spins might exchange energy.

Take two-spin systems, one with N_1 spins and spin excess x_1 and the other with N_2 spins and spin excess x_2 . We could say that the total spin excess is $x = x_1 + x_2$. In the presence of the magnetic field, B , the total energy is $-2xmB$. Now, suppose that we bring these systems in thermal contact such that energy can be exchanged.

We will require of course that the total energy be conserved and that in turn means that, although x_1 and x_2 may change as spins exchange energy between the systems, the total value of x must remain constant. In other words, whatever x_1 value results, $x_2 = x - x_1$. Hence the multiplicity function, $g(N, x)$, that defines the combined state of the two systems involves a sum of terms representing each possible value of x_1 and of course fixed total x . Those terms will be products of the individual corresponding multiplicities g_1 and g_2 ; in other words

$$g(N, x) = \sum_{x_1} g_1(N_1, x_1) g_2(N_2, x - x_1) \quad (1.3)$$

As x_1 varies, one of these products $g_1(N_1, \hat{x}_1) g_2(N_2, x - \hat{x}_1)$ will dominate the summation. This is easily found for the case of the spin system where

$$g(N, x) = g_1(N_1, 0) g_2(N_2, 0) \sum_{x_1} \exp\left(\frac{-2x_1^2}{N_1}\right) \exp\left(\frac{-2(x - x_1)^2}{N_2}\right) \quad (1.4)$$

and the value \hat{x}_1 at which $g_1(N, x_1) g_2(N, x - x_1)$ is a maximum is found from setting the derivative, with respect to x_1 , of $\exp(-2x_1^2/N_1) \exp(-2(x-x_1)^2/N_2)$ to zero, yielding $\hat{x}_1/N_1 = (x - \hat{x}_1)/N_2$. Again, the sharpness of this maximum grows as the number of particles increases and its overwhelming dominance for large N values ensures that this is the ‘equilibrium’ combined state to which the systems in thermal contact have evolved through the process of exchanging energy. We call this state *thermal equilibrium*. Given that the starting combined multiplicity was just one member of the sum expressed by eqn 1.3, and one which is dominated by a much larger term corresponding to the maximum, it is clear that allowing the spin systems to exchange energy leads to an increased combined multiplicity.

What is the fundamental principle that determines the thermal equilibrium value \hat{x}_1 , and $\hat{x}_2 = x - \hat{x}_1$? The zeroth law of thermodynamics tells us that it is the equilibration of temperature. But what is temperature in this context? There is an easy way to derive this. We have seen that the system multiplicity in the case of the spin states is related to the number of particles N and the spin excess x . In a magnetic field, that spin excess is simply related to the system energy, U . It is for this reason that energy conservation requires the total spin excess to remain constant in the case of two systems in thermal contact. This idea can be generalised to any system by writing the multiplicity as $g(N, U)$. Then, for two systems in thermal equilibrium,

$$g(N, U) = \sum_{U_1} g_1(N_1, U_1) g_2(N_2, U - U_1) \quad (1.5)$$

Now the largest term in the sum comprising eqn 1.5 is determined by the requirement that $g_1(N_1, U_1) g_2(N_2, U - U_1)$ be a maximum, as illustrated in Fig 1.3. Hence, for an infinitesimal energy exchange dU_1 ,

$$\frac{d(g_1 g_2)}{dU_1} = \left(\frac{\partial g_1}{\partial U_1}\right)_{N_1} g_2 + g_1 \left(\frac{\partial g_2}{\partial U_1}\right)_{N_2} = 0 \quad (1.6)$$

6 Thermal processes and diffusion

Then, since by energy conservation $\partial U_2 = -\partial U_1$, dividing by $g_1 g_2$ we have

$$\frac{1}{g_1} \left(\frac{\partial g_1}{\partial U_1} \right)_{N_1} = \frac{1}{g_2} \left(\frac{\partial g_2}{\partial U_2} \right)_{N_2} \quad (1.7)$$

or

$$\left(\frac{\partial \log g_1}{\partial U_1} \right)_{N_1} = \left(\frac{\partial \log g_2}{\partial U_2} \right)_{N_2} \quad (1.8)$$

$\sigma = \log g$ is called the *entropy* of the system. And it is the derivative of the entropy with respect to energy, at fixed numbers of particles, which is equalised between two systems in thermal contact. Thus the *temperature* is defined. By historical convention $\partial \log g / \partial U$ is the inverse temperature and the natural dimension of temperature is energy. Of course, to reconcile the Kelvin units of temperature, a constant is required. Not surprisingly, this is Boltzmann's constant, so that one may define

$$\frac{1}{T} = k_B \left(\frac{\partial \log g}{\partial U} \right)_N \quad (1.9)$$

while for the same historical reason, entropy³ is defined as $S = k_B \log g$.

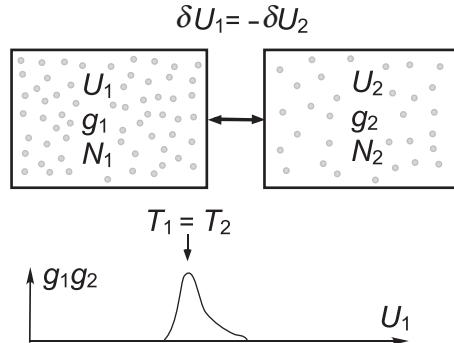


Fig. 1.3 Thermal equilibrium between two systems of N_1 and N_2 particles is attained when the combined multiplicity is a maximum, and is denoted by the temperatures T_1 and T_2 being equal, where $T^{-1} = k_B \frac{\partial \log g}{\partial U}$ at constant N .

The use of spin- $\frac{1}{2}$ particles to illustrate the statistical physics concepts that underpin entropy and temperature is merely a means to an end. Other routes are possible, for example through the more complex consideration of the kinetic energies of gas molecules. That was the path charted by Maxwell and Boltzmann [8, 9]. But the principles are indeed the same. Bring two reservoirs of gas into thermal contact and the energy will redistribute such as to maximise the combined multiplicities of possible states of the gas molecules. One consequence of the increased combined multiplicity when systems are allowed to come into thermal contact is that the energy exchange, known as heat transfer, is associated with a net entropy increase. This idea lies at

³Given this definition, it might be appropriate to refer to σ as ‘informational entropy’.

the heart the second law of thermodynamics, a law which is rooted in the statistical behaviour of immensely large ensembles.⁴

1.2.2 The Boltzmann distribution and partition function

Whether we are considering the system of nuclear spins in a nuclear magnetic resonance (NMR) sample, or the system of molecules of gas in a container, the process of coming to thermal equilibrium involves the exchange of heat with the surrounding environment or *reservoir*. Just as a water reservoir allows small removals or additions without significantly affecting its contents, so the thermal reservoir, by virtue of its much larger heat capacity or internal energy, allows the nuclear spins or the gas molecules to exchange heat without significantly affecting the reservoir temperature. By comparison with that reservoir our system is small. For the gas molecules example, the reservoir might be the laboratory environment surrounding the container, while for the nuclear spins the reservoir is the atomic and molecular environment of the sample, whose thermal motions represent a vastly larger internal energy capacity than the nuclear magnetic energy levels. In NMR we refer to this reservoir as the *lattice*.

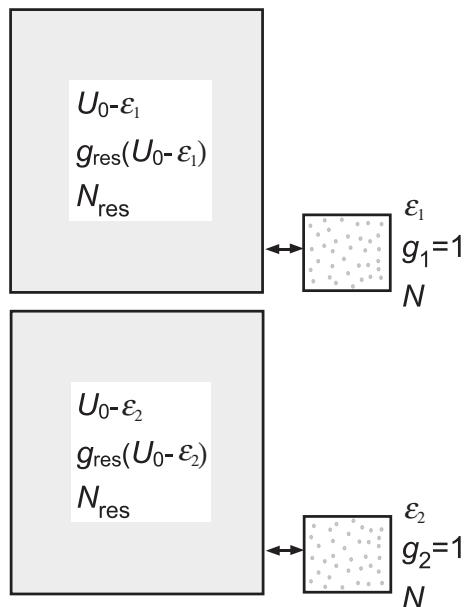


Fig. 1.4 Thermal equilibrium between a reservoir and of a system of N particles at specified states ($g_1 = 1$) of energy ϵ_1 and ϵ_2 . The relative probabilities for the two energies are determined by the ratio of reservoir multiplicities.

Suppose we specify a particular quantum state s of a small system in thermal equilibrium with a reservoir at temperature T , as illustrated in Fig 1.4. Since the

⁴By implication therefore, when the ensemble becomes sufficiently small, violations of the second law may be seen. This is indeed the case, as may be seen in reference [2].

8 Thermal processes and diffusion

system state is specified it may be assigned multiplicity 1, and the multiplicity for the combined state reduces to that of the reservoir alone. Assign to the system state the energy ϵ_s , and call the total energy of reservoir plus system U_0 . Now ponder the question, how many reservoir states, or in other words, what multiplicity of the reservoir, will permit that system state s ? More particularly, what is the relative probability of the system being in thermal equilibrium at T , at two different states $s = 1$ and $s = 2$? We might expect that relative probability to be given by the ratio of those reservoir multiplicities, i.e.

$$\frac{P(\epsilon_1)}{P(\epsilon_2)} = \frac{g_{res}(U_0 - \epsilon_1)}{g_{res}(U_0 - \epsilon_2)} = \frac{\exp(\sigma_{res}(U_0 - \epsilon_1))}{\exp(\sigma_{res}(U_0 - \epsilon_2))} \quad (1.10)$$

Only reservoir multiplicities are relevant, because the states of the system in thermal equilibrium with that reservoir are specified. Since $\epsilon_s \ll U_0$, by Taylor expansion,

$$\begin{aligned} \sigma_{res}(U_0 - \epsilon_s) &= \sigma_{res}(U_0) - \epsilon_s \left(\frac{\partial \sigma_{res}}{\partial U} \right)_N + \dots \\ &= \sigma_{res}(U_0) - \frac{\epsilon_s}{k_B T} + \dots \end{aligned} \quad (1.11)$$

This leads directly to

$$\frac{P(\epsilon_1)}{P(\epsilon_2)} = \frac{\exp\left(\frac{-\epsilon_1}{k_B T}\right)}{\exp\left(\frac{-\epsilon_2}{k_B T}\right)} \quad (1.12)$$

The term $\exp(-\epsilon_1/k_B T)$ is known as the *Boltzmann factor* for that system's energy state, and gives the relative probability of finding the system at that state when in thermal equilibrium with the reservoir at temperature T .

An extremely useful function, the *partition function*, is the sum over the Boltzmann factors for all possible states of the system,

$$Z(T) = \sum_s \exp(-\epsilon_s/k_B T) \quad (1.13)$$

Clearly, the normalised probability for finding our system in the state s will be given by

$$P(\epsilon_s) = \frac{\exp(-\epsilon_s/k_B T)}{Z} \quad (1.14)$$

while the average energy of the system will be given by

$$U = \langle \epsilon \rangle = \frac{\sum_s \epsilon_s \exp(-\epsilon_s/k_B T)}{Z} = k_B T^2 \frac{\partial \log Z}{\partial T} \quad (1.15)$$

1.2.3 Partition function, free energy, and entropy of an ideal monatomic gas

For the ideal gas, where long range interactions between molecules are neglected, the energies are described entirely in terms of particle kinetic energies, $p^2/2m$, where p is the magnitude of the molecular translational momentum vector (p_x, p_y, p_z). For molecules

in a container, quantum mechanics dictates that the allowed momenta are restricted by standing de Broglie wave states.⁵ If we imagine one cartesian direction (say x) in a cubic container of side length L (see Fig 1.5), the allowed standing waves correspond to integer multiples (n_x) of half wavelengths, such that the allowed momenta in three dimensions are $\frac{h}{2L} (n_x, n_y, n_z)$. The single particle partition function is therefore

$$Z_1 = \sum_{n_x} \sum_{n_y} \sum_{n_z} \exp \left(\frac{-h^2(n_x^2 + n_y^2 + n_z^2)}{8k_B T m L^2} \right) \quad (1.16)$$

When the spacing of the energies is small compared with $k_B T$ the sums can be replaced by integrals as

$$Z_1 = \int_0^\infty \int_0^\infty \int_0^\infty \exp(-\alpha^2(n_x^2 + n_y^2 + n_z^2)) dn_x dn_y dn_z \quad (1.17)$$

where $\alpha^2 = h^2/8k_B T m L^2$. It is easy to show that $Z_1 = (2\pi k_B T m L^2/h^2)^{3/2}$ and, from eqn 1.15, that the single particle average energy is given by $U = \frac{3}{2}k_B T$, equating to $\frac{1}{2}k_B T$, for each of the three ‘energy storing modes’ or in other words, each of the three degrees of freedom of the single particle motion. The outcome of this simple calculation, namely that there is an average thermal energy of $\frac{1}{2}k_B T$ per energy storing mode, is perfectly general for all systems in thermal equilibrium. This famous result is known as *Boltzmann’s equipartition of energy principle*.

For an ideal monatomic gas of N particles in the volume $V = L^3$, the total partition function is given by

$$Z_N = \frac{1}{N!} Z_1^N \quad (1.18)$$

The Z_1^N product of single particle exponential factors gathers up all possible single particle energy sums in the N -particle exponential, while the $1/N!$ term allows for the fact that we need to count each possible N -particle energy sum only once, and not the $N!$ times that each occurs in the product of factors.

The partition function also permits a calculation of the free energy $F = U - TS$ as $F = -k_B T \log(Z)$, and hence the entropy $S = -(\partial F / \partial T)_V$. For the ideal monatomic gas the entropy for N molecules is

$$S = Nk_B \left(\log \left(\frac{n_Q}{n} \right) + \frac{5}{2} \right) \quad (1.19)$$

where n is the number of molecules per unit volume and $n_Q = Z_1/L^3 = (2\pi k_B T m/h^2)^{3/2}$, the so-called *quantum concentration*

Equation 1.19 is known as the Sackur–Tetrode equation [10, 11]. While it provides an exact expression for the entropy of an ideal gas, it is possible to gain simpler insight in a different way. Take a container of volume V in which we successively introduce N gas molecules. If the effective volume of one molecule is V_m , then the number of

⁵The de Broglie wavelength, λ , of a particle of momentum p is h/p , where h is Planck’s constant.

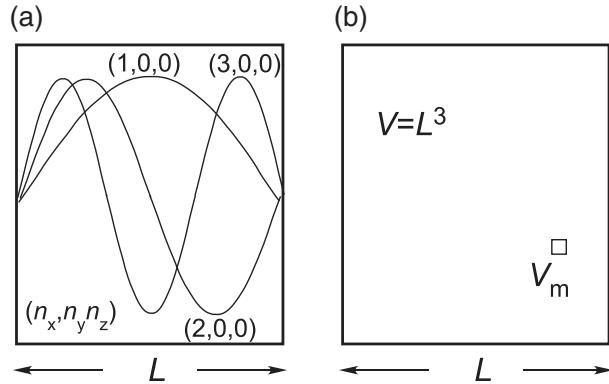


Fig. 1.5 Depictions of allowed states for a gas contained in box of side length L . In (a) the allowed (de Broglie) standing waves are shown, while in (b) a classical picture is presented in terms of the different volumes of size V_m available for each gas molecule.

possible locations for the first molecule is V/V_m , for the second $(V-V_m)/V_m \approx V/V_m$ and so on. Provided that V_m is small, that is $NV_m \ll V$, then the total multiplicity for introducing N molecules is

$$g \approx \left(\frac{V}{V_m} \right)^N \quad (1.20)$$

and so the gas entropy is

$$S \approx Nk_B \log \left(\frac{V}{V_m} \right) \quad (1.21)$$

Both eqns 1.19 and 1.21 tell us that for fixed numbers of molecules N , a volume change δV results in an entropy change $\delta S \sim N\delta V/V$. We will find this result useful in following Einstein's derivation of the self-diffusion coefficient for a Brownian particle.

1.2.4 Ensembles and averages

Statistical physics rests upon the ideas of systems of *quantum states* and *configurations*. Let us return to our system of four spin- $\frac{1}{2}$ particles. The specification ‘one spin up and three spins down’ represents a particular configuration, one in which the spin excess is -2 . But there are four particular quantum state manifestations of that configuration, namely $(\uparrow\downarrow\downarrow\downarrow)$, $(\downarrow\uparrow\downarrow\downarrow)$, $(\downarrow\downarrow\uparrow\downarrow)$, and $(\downarrow\downarrow\downarrow\uparrow)$.

The fundamental principle of statistical mechanics is that *all accessible quantum states of a system are equally likely*. If there are W quantum states of a system compatible with all the constraints on energy, volume, and any other parameter defining the physics of the system, then the probability of finding the system in each accessible quantum state is $P = W^{-1}$. To illustrate the idea of accessibility for our spin system, imagine that the spin system had a particular energy available to it when placed in a magnetic field. This in turn would specify the allowed spin excess and hence the allowed configuration. Alternatively, as we saw in the previous section, where two systems of spins are in thermal contact with total spin excess $2x$, the accessible states

belonged to configurations x_1 and $x_2 = x - x_1$, and the number of accessible states of all the combined system configurations was $g(N, x) = \sum_{x_1} g_1(N_1, x_1) g_2(N_2, x - x_1)$.

Knowing all the possible quantum states of a system allows us to construct the *ensemble*, the set of all replicas of the system representing each of the accessible quantum states. From the ensemble, we are able to calculate the average value of any physical quantity X as

$$\langle X \rangle = \sum_s P(s) X(s) \quad (1.22)$$

$\langle X \rangle$ is known as the ensemble average of X .⁶ Given the fundamental principle, we could carry out the calculation of $\langle X \rangle$ with a sum over configurations, rather than a sum over states. In this case $P(s)$ would be replaced with the configurational probability.

As a first example, consider the case of N spins in the absence of a magnetic field, such that all states are accessible. For N large, the average value of spin excess, $\langle 2x \rangle$, will be given by⁷

$$\langle 2x \rangle = \sum_{-\frac{N}{2}}^{\frac{N}{2}} 2x g(N, 0) \exp\left(\frac{-2x^2}{N}\right) = 0 \quad (1.23)$$

A second example concerns the average energy of N spins immersed in a magnetic field B and in thermal equilibrium with a reservoir at temperature T . The possible energies for the two quantum states of each spin are mB and $-mB$ with corresponding probabilities given by the normalised Boltzmann factors, $\exp(\pm mB/k_B T)/(\exp(mB/k_B T) + \exp(-mB/k_B T))$, so that

$$U = N \frac{mB \exp(mB/k_B T) - mB \exp(-mB/k_B T)}{\exp(mB/k_B T) + \exp(-mB/k_B T)} \quad (1.24)$$

Finally, let us use eqn 1.17 to find the average energy of N monatomic gas molecules in thermal equilibrium at temperature T , a result already obtained using the partition function approach. For simplicity we express the particle's momentum in spherical polar coordinates as $p_x = p \sin \theta \cos \phi$, $p_y = p \sin \theta \sin \phi$, $p_z = p \cos \theta$. The single particle Boltzmann factors. are

$$P(\epsilon) = \frac{\exp(-p^2/2mk_B T)}{4\pi \int_0^\infty p^2 \exp(-p^2/2mk_B T) dp} \quad (1.25)$$

and the kinetic energy of translation is $\epsilon = p^2/2m$. Integrating by parts, it is simple to show that, as before, the result is $U = \langle p^2/2m \rangle = \frac{3}{2} N k_B T$.

1.2.5 Fluctuations, ergodicity, and the autocorrelation function

The ideal gas example is instructive for another reason. If we were to follow the individual gas molecules we would find that as a result of random collisions, their momentum

⁶Throughout the book the notation $\langle \dots \rangle$ will be used to represent an ensemble average.

⁷By a similar, but slightly more complex calculation, it may be shown that the fractional root mean-squared spin excess, $\langle 4x^2 \rangle^{1/2}/N = N^{-1/2}$, consistent with the idea that the probability distribution becomes a progressively sharper peak centred at spin excess zero as N increases.

12 Thermal processes and diffusion

values would fluctuate. Indeed, if we followed a sufficient number of collisions for one molecule, we would find that it would sample all possible values of (p_x, p_y, p_z) . Expressed in the language of ensembles, we say that the individual molecules would, over time, sample all possible ensemble states. This particular property is known as *ergodicity*. An ergodic system is one for which the long-time average of a fluctuating variable, $A(t)$, is equal to the instantaneous ensemble average for that variable. In other words

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(t) dt = \langle A \rangle \quad (1.26)$$

Ergodicity is generally assumed for systems that are *stationary*; that is, when the probabilities associated with various ensemble states do not change with time. . Equation 1.26 will prove particularly useful in evaluating the effect of fluctuations in stationary ensembles.

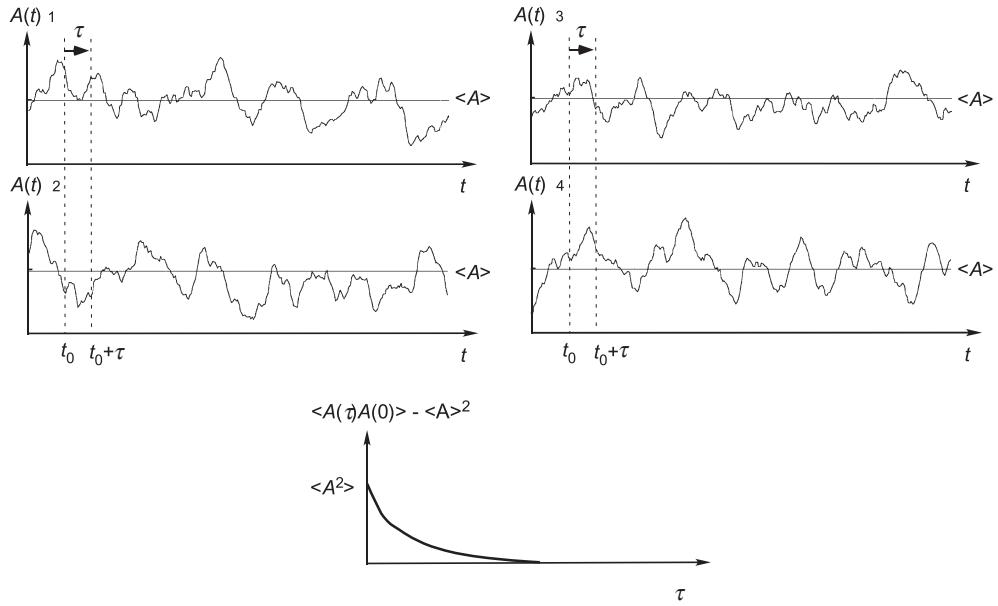


Fig. 1.6 The time-dependence of the stochastic function of time, $A(t)$, for four different members of the ensemble. For an ergodic system the time average is the same as the ensemble average. The lower graph shows the autocorrelation function obtained from the ensemble average of $A(t_0 + \tau)A(t_0)$, identical to $\langle A(\tau)A(0) \rangle$ for a stationary and ergodic system.

A convenient way to deal with a stochastic variable $A(t)$ is to define its behaviour in terms of its autocorrelation function $\langle A(t)A(0) \rangle$ [12, 13]. Figure 1.6 illustrates the noisy behaviour of $A(t)$ along with the well-defined $\langle A(t)A(0) \rangle$. When $t = 0$ the autocorrelation function is simply equal to the mean squared value $\langle A^2 \rangle$. As time increases, $\langle A(t)A(0) \rangle$ decays such that at times much longer than the ‘correlation

'time', $A(t)$ and $A(0)$ are completely uncorrelated so that we may write $\langle A(t)A(0) \rangle = \langle A(t) \rangle \langle A(0) \rangle = \langle A \rangle^2$.

We may define a normalised autocorrelation decay function $g(t)$ as

$$g(t) = \frac{\langle A(t)A(0) \rangle - \langle A \rangle^2}{\langle A^2 \rangle - \langle A \rangle^2} \quad (1.27)$$

The correlation time, τ_c , is then defined by the integral

$$\tau_c = \int_0^\infty g(t) dt \quad (1.28)$$

Note, for a stationary ensemble, the choice of the time origin is unimportant. Thus $\langle A(\tau)A(0) \rangle = \langle A(t_0 + \tau)A(t_0) \rangle$. In general, $\langle A(t)A(t') \rangle$ depends only on the time difference, $t - t'$.

1.3 Thermal energy and self-diffusion

1.3.1 Fick's law

The process of diffusion may be envisaged as a flux of particles arising from a gradient in concentration, an idea explained by Adolf Fick in 1855 [14]. Given a local concentration of particles $n(\mathbf{r}, t)$, the flux of particles may be written

$$\mathbf{J} = -D \nabla n(\mathbf{r}, t) \quad (1.29)$$

where the constant D is known as the *diffusion coefficient*. In one dimension, the component of flux along the x -axis may be written $J_x = -D \frac{\partial n}{\partial x}$. This relation is sometimes known as Fick's first law and the process is illustrated in Fig 1.7.

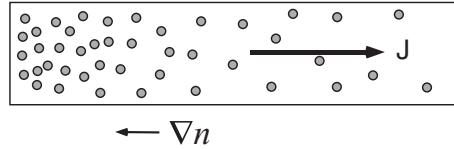


Fig. 1.7 The diffusion flux, \mathbf{J} , caused by a concentration gradient ∇n .

Of course, conservation of particles requires that the time rate of change of $n(\mathbf{r}, t)$ is simply related to the local flux divergence, $-\nabla \cdot \mathbf{J} = \frac{\partial n}{\partial t}$. This leads to the Fick's second law, also known as the diffusion equation,

$$\frac{\partial n}{\partial t} = D \nabla^2 n \quad (1.30)$$

Fick's laws were developed to describe the behaviour of solute molecules as a consequence of a non-uniform concentration—the way in which solute molecules will drift from higher to lower concentration so as to equalise concentration gradients. This

14 Thermal processes and diffusion

process is sometimes termed *mutual diffusion* since it requires a countercurrent of solute and solvent particles to maintain the overall mass density. However, the idea may also be applied even to the case of molecules in a liquid comprising a single molecular component, or to the case of solvent molecules in solution or Brownian particles in suspension, where no macroscopic concentration gradients exist. This latter process, first explained by Albert Einstein, is known as *self-diffusion* and it is driven by the random motions associated with thermal energy. Intriguingly, the language of Fick's laws still applies in the case of self-diffusion, and again we take $n(\mathbf{r}, t)$ to be the local probability of finding the solute or solvent molecule of interest. The difference now is that $n(\mathbf{r}, t)$ may be macroscopically uniform, although locally structured. In order to calculate the diffusion coefficient we need to look to the role of thermal fluctuations.

1.3.2 Brownian motion: the Einstein derivation

Einstein's explanation of Brownian motion was based on the idea of the Brownian particles being buffeted by surrounding water molecules, with a net force, K ,⁸ resulting from the imbalance of exterior collisions [4]. In so reasoning he used the idea that the particles themselves behaved like the molecules in an ideal gas, with pressure $p = k_B T \frac{N}{V}$, where $\frac{N}{V} = n$, the number of Brownian particles per unit volume. Einstein's key step was to consider a small displacement δx of the Brownian particle under the net force, such that the free energy is minimised:

$$\delta F = \delta U - T\delta S = 0 \quad (1.31)$$

The problem was then tackled by calculating δU and δS for a finite volume of particles. Taking a cuboid volume (see Fig 1.8), comprising unit area normal to the x -axis and bounded by $x = 0$ and $x = l$, Einstein expressed the energy change as the total work done by the particles contained within the volume, $-\int_0^l Kn\delta x dx$, while, for the entropy change per particle he used the ideal gas result (eqn 1.21), $\delta S = k_B \frac{\delta V}{V}$ or $\delta S = k_B \frac{\partial \delta x}{\partial x}$. Hence

$$\delta S = k_B \int_0^l n \frac{\partial \delta x}{\partial x} dx = -k_B \int_0^l \frac{\partial n}{\partial x} \delta x dx \quad (1.32)$$

The free energy relationship yields $Kn = k_B T \frac{\partial n}{\partial x}$ and, noting that the osmotic pressure is $k_B T n$, it emerges that the gradient in osmotic pressure forces provides a balance to the force K .

Next, Einstein considered the balance of mass flow needed to maintain an equilibrium concentration, n . His idea was that the drift of particles caused by K would need to be balanced by a diffusion current in the opposite direction. Using the result that a force K results in a single particle velocity (component along x) K/ζ , where ζ is the friction, he reasoned that the particle current crossing unit area per unit time is nK/ζ , and that this will be counteracted by the self-diffusive flow $-D \frac{\partial n}{\partial x}$, a balance represented by

$$\frac{Kn}{\zeta} - D \frac{\partial n}{\partial x} = 0 \quad (1.33)$$

⁸For convenience taken to be the component along one axis, labelled x in the present case.

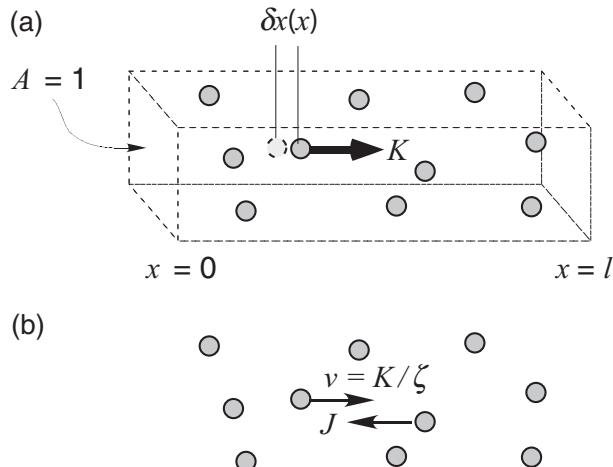


Fig. 1.8 Einstein's diffusion coefficient derivation. (a) The first step, the free energy minimisation with the force on the Brownian particle from the surrounding solvent molecules doing work ($K\delta x$) and resulting in an entropy change as the particle concentration changes; (b) the second step, the balance of the particle drag velocity with the osmotic flux, which maintains the equilibrium concentration.

Using the result $Kn = k_B T \frac{\partial n}{\partial x}$ from eqn 1.31 Einstein found the famous result

$$D = \frac{k_B T}{\zeta} \quad (1.34)$$

Einstein chose for ζ the Stokes drag [15] of a spherical particle of radius R in a medium of viscosity η , $\zeta = 6\pi\eta R$. Hence

$$D = \frac{k_B T}{6\pi\eta R} \quad (1.35)$$

In his remarkable achievement he was not alone. In 1904 an Australian physicist, William Sutherland, using the same reasoning as Einstein, derived the same diffusion relationship and reported the result at a meeting in Dunedin, New Zealand, of the Australian and New Zealand Society for the Advancement of Science [16, 17]. For that reason eqn 1.35 has become known as the Sutherland–Einstein relation.

1.3.3 The probabilistic description

Markov processes and the language of propagators

Einstein's achievement in deriving eqn 1.34 was supplemented by an even more ingenious insight, namely, to ascribe to Fick's laws for the diffusion of molecules in a concentration gradient, an interpretation based on probabilities. This interpretation made possible the description of self-diffusion of molecules or Brownian particles, with a diffusion coefficient as given by eqn 1.34, but in which the average concentration was uniform. Thus the ideas behind Fick's laws were extended to new territory, not only enabling a description of Brownian motion as a stochastic process, but also as one in which the probability densities obeyed differential equations.

16 Thermal processes and diffusion

Here we lay out some of the basic ideas. For simplicity we will describe the Brownian particle displacement, $x(t)$, over a time t in one dimension. Suppose we were to measure this displacement repeatedly in a set of measurements carried out independently, N times in succession. We would find that we would obtain different results, $x_0(t), x_1(t), x_2(t), \dots, x_N(t)$. These data represent a sample from a statistical ensemble, and, if N is made sufficiently large, we can find the probabilistic distribution obeyed by the stochastic random variable $x(t)$. This distribution is represented by a probability density $p(x, t)$, where $p(x, t) dx$ is the probability that $x(t)$ is in the range $x < x(t) \leq x + dx$. In order to describe the dynamics of Brownian motion we need to know the probabilities associated with the various possible paths of $x(t)$ as time advances. At its simplest, one can define a new probability density [12] $P(x_0, t_0; x_1, t_1)$, where $P(x_0, t_0; x_1, t_1) dx_0 dx_1$ is the joint probability that $x(t_0)$ is in the range $x_0 < x(t_0) \leq x_0 + dx_0$, while $x(t_1)$ is in the range $x_1 < x(t_1) \leq x_1 + dx_1$. Clearly, higher order joint probabilities could be described involving combinations $x_0, t_0; x_1, t_1; x_2, t_2$ and so on. However, as we shall see, the second order joint probability will be sufficient for our description of Brownian motion.

This framework enables us to describe a transition probability, the chance that a Brownian particle that certainly starts at x_0 at time t_0 will be found between x_1 and $x_1 + dx_1$ at time t_1 ; that is:

$$P(x_0, t_0 | x_1, t_1) dx_1 = \frac{P(x_0, t_0; x_1, t_1) dx_1}{p(x_0, t_0)} \quad (1.36)$$

$P(x_0, t_0 | x_1, t_1) dx$ is known as the *conditional probability*. A special feature of Brownian motion is that the probability of finding the particle between x_1 and $x_1 + dx_1$ at time t_1 , when it was certainly at x_0 at time t_0 , does not depend on where the particle happened to be before time t_0 . Such a property is called *Markovian*. For such Markov processes we can calculate the probability of various different paths in a stepwise manner. For example, the joint probability of finding the particle between x_1 and $x_1 + dx_1$ at time t_1 and between x_2 and $x_2 + dx_2$ at time t_2 when it was certainly at x_0 at time t_0 is just

$$P(x_0, t_0 | x_1, t_1; x_2, t_2) dx_1 dx_2 = P(x_0, t_0 | x_1, t_1) dx_1 P(x_1, t_1 | x_2, t_2) dx_2 \quad (1.37)$$

This enables us to find the conditional probability for any starting and finishing point by integrating over all possible intermediate paths. In other words

$$p(x_0, t_0 | x_2, t_2) = \int p(x_0, t_0 | x_1, t_1) p(x_1, t_1 | x_2, t_2) dx_1 \quad (1.38)$$

Similarly we may find the probability density $p(x, t)$ by integrating over all possible starting points

$$p(x, t) = \int p(x_0, t_0) P(x_0, t_0 | x, t) dx_0 \quad (1.39)$$

Because the diffusion process is Markovian, only the time displacement matters, not the absolute time at the origin. This enables us to simplify our notation from $P(x_0, t_0 | x_1, t_1)$ to $P(x_0 | x_1, t)$ where $t = t_1 - t_0$.

Fick's law for probability density and for propagators

Einstein's bold step was to show that the probability density $p(x, t)$ obeyed Fick's laws in the same manner as the more familiar particle concentration function, $n(x, t)$. The details of his idea are based on an analysis of simple Taylor expansions under small increments of time and space. In what follows, we will treat $p(\mathbf{r}, t)$ as a three-dimensional (3-D) probability density. We will see that just as $p(\mathbf{r}, t)$ obeys Fick's Law, so does the conditional probability density, $P(\mathbf{r}|\mathbf{r}', t)$.

Writing \mathbf{r} as our starting coordinate and \mathbf{r}' as the final coordinate,

$$p(\mathbf{r}', t) = \int p(\mathbf{r}, 0) P(\mathbf{r}|\mathbf{r}', t) d\mathbf{r} \quad (1.40)$$

since $p(\mathbf{r}', t)$ obeys the Fick's law diffusion equation for arbitrary initial conditions $p(\mathbf{r}, 0)$, it is apparent that the conditional probability also obeys the partial differential equation⁹

$$\frac{\partial}{\partial t} P(\mathbf{r}|\mathbf{r}', t) = D \nabla_{\mathbf{r}'}^2 P(\mathbf{r}|\mathbf{r}', t) \quad (1.41)$$

For an fluid of infinite extent, and given the initial condition $P(\mathbf{r}|\mathbf{r}', 0) = \delta(\mathbf{r}' - \mathbf{r})$, the Dirac delta function, the solution to eqn 1.41 is

$$P(\mathbf{r}|\mathbf{r}', t) = (4\pi Dt)^{-\frac{3}{2}} \exp\left(-\frac{(\mathbf{r}' - \mathbf{r})^2}{4Dt}\right) \quad (1.42)$$

Its dependence on time is shown in Fig 1.9. In fact, eqn 1.41 is true only for an isotropic medium, where the diffusion is indeed a simple scalar property. In anisotropic media we shall find it necessary to define a diffusion tensor and rewrite the differential equation for diffusion. This extension is covered in Chapter 2 and elsewhere in this book. The Gaussian nature of the conditional probability for self-diffusion, represented by eqn 1.41, leads to two important results,

$$\langle (x' - x)^2 \rangle = 2Dt \quad (1.43)$$

and

$$\langle (\mathbf{r}' - \mathbf{r})^2 \rangle = 6Dt \quad (1.44)$$

Finite boundaries: the eigenmode solution

Attempting a 'separation of variables' solution to Fick's law (eqn 1.30) we write $p(\mathbf{r}, t) = u(\mathbf{r})v(t)$ and obtain

$$\frac{D \nabla^2 u(\mathbf{r})}{u(\mathbf{r})} = \frac{\partial v(t)/\partial t}{v(t)} \quad (1.45)$$

and since the left- and right-hand sides are, respectively, functions of \mathbf{r} and t only, both must be a constant $-Dk^2$. This leads to an exponential decay relationship for $v(t)$.¹⁰ The $u(\mathbf{r})$ obey the Helmholtz equation

⁹Where more than one spatial coordinate is present, the subscript on the operator is used to refer to the relevant variable. For Markov systems, the operator could act on either \mathbf{r} or \mathbf{r}' and for that reason the subscript is subsequently dropped.

¹⁰Hence the choice of a negative constant $-Dk^2$. This choice is required to keep $v(t)$ finite as $t \rightarrow \infty$.

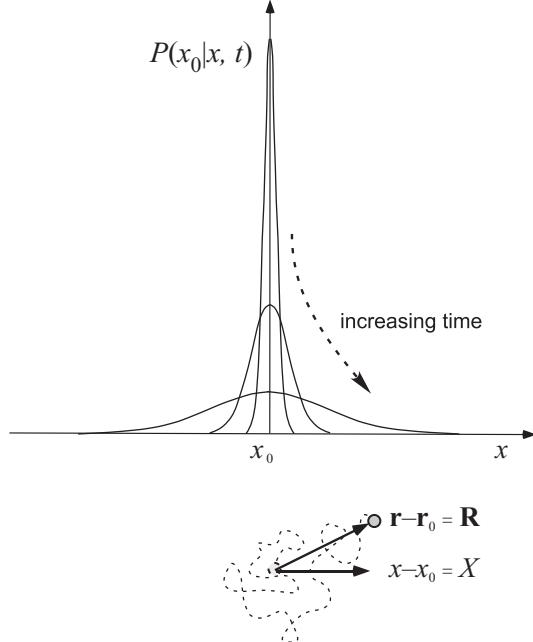


Fig. 1.9 Behaviour of the conditional probability for an ensemble of particles undergoing Brownian motion where the successive Gaussians correspond to successively increasing time.

$$\nabla^2 u(\mathbf{r}) + k^2 u(\mathbf{r}) = 0 \quad (1.46)$$

For a finite-sized sample, the $u(\mathbf{r})$ must obey some imposed boundary conditions. This constraint fixes the allowed solutions to the eigenmode set $\{u_n(\mathbf{r})\}$ with

$$\nabla^2 u_n(\mathbf{r}) = \lambda_n u_n(\mathbf{r}) \quad (1.47)$$

where the eigenvalues λ_n are $-k_n^2$. Given real eigenvalues, the $\{u_n(\mathbf{r})\}$ form an orthogonal set under $\int_V \dots d\mathbf{r}$, where V is the volume of the enclosed fluid. Thus the general eigenmode solution may be written as the linear superposition,

$$p(\mathbf{r}, t) = \sum_n A_n u_n(\mathbf{r}) \exp(-Dk_n^2 t) \quad (1.48)$$

with the A_n to be determined from the initial condition.

Suppose that the $u_n(\mathbf{r})$ are normalised under $\int_V \dots d\mathbf{r}$.¹¹ Then the initial $p(\mathbf{r}, t)$ may be written

$$p(\mathbf{r}, 0) = \sum_n A_n u_n(\mathbf{r}) \quad (1.49)$$

whence $A_m = \int_V p(\mathbf{r}', 0) u_m^*(\mathbf{r}') d\mathbf{r}'$ and

¹¹In other words, the $u_n(\mathbf{r})$ are an orthonormal set where $\int_V u_m^*(\mathbf{r}) u_n(\mathbf{r}) d\mathbf{r} = \delta_{mn}$ and we allow that the $u_n(\mathbf{r})$ may be complex.

$$p(\mathbf{r}, t) = \sum_n \left(\int_V p(\mathbf{r}', 0) u_n^*(\mathbf{r}') d\mathbf{r}' \right) u_n(\mathbf{r}) \exp(-Dk_n^2 t) \quad (1.50)$$

In the magnetic resonance examples to be considered in Chapter 6, $p(\mathbf{r}, t)$ will represent a normalised magnetisation density function. For a pore of uniform fluid density, a common situation in magnetic resonance experiments will be that the initial magnetisation across the pore will be uniform so that $p(\mathbf{r}, 0) = 1/V$ and

$$p(\mathbf{r}, t) = \sum_n \frac{\int_V u_n^*(\mathbf{r}') d\mathbf{r}'}{V} u_n(\mathbf{r}) \exp(-Dk_n^2 t) \quad (1.51)$$

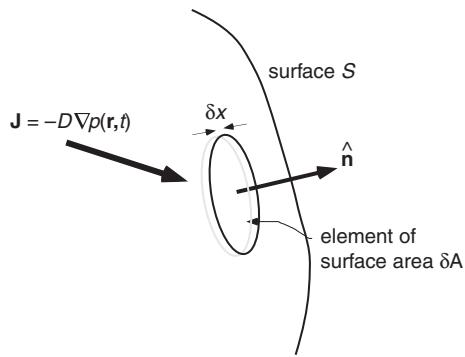


Fig. 1.10 Schematic of surface relaxation involving an element of volume adjacent to δA with layer thickness δx in which spin or particle probability may be ‘lost’ by some relaxation or permeation process.

The boundary conditions by which the eigenmodes $u_n(\mathbf{r})$ are determined relate to the degree to which molecules are reflected or absorbed at the walls. This can be easily seen by reference to Fig. 1.10, where we consider an element of volume adjacent to the surface element δA with layer thickness δx . Within this volume element, probability is ‘lost’ by some process. In the case of magnetic resonance, even under perfectly reflecting wall conditions, collision of molecules with the wall can lead to spin relaxation and hence a loss of spin magnetisation. Alternatively, for a semi-permeable wall, molecules might leak out at the boundary, with none returning. Let us call the loss rate per molecule $1/T$. Then the probability loss rate within the defined volume bounding the surface is $(\delta x \delta A p(\mathbf{r}, t)/T)|_S$. Rewriting $\bar{\rho} = \delta x/T$ and allowing that this loss rate in the volume element bounding the surface is fed by the probability current into this volume, we have at the surface $|_S$,

$$D\hat{n} \cdot \nabla p(\mathbf{r}, t) + \bar{\rho} p(\mathbf{r}, t) = 0 \quad (1.52)$$

Perfectly reflecting walls correspond to $\bar{\rho} = 0$, while perfectly permeable or absorbing walls correspond to $\bar{\rho} \rightarrow \infty$.

Conditional probability eigenmodes

The simple Gaussian form of eqn 1.42 for the conditional probability applies when diffusion is unrestricted. Where obstacles to diffusion occur, for example due to the bounding surfaces confining liquid molecules interpenetrating a porous medium, that probability will significantly deviate from a Gaussian. While Fick's law still applies within the bulk of the fluid, reflection at the boundaries imparts to the propagator $P(\mathbf{r}|\mathbf{r}', t)$ spatio-temporal properties characteristic of the boundary conditions.

Equation 1.41 may be tackled via the same eigenmode expansion employed for $p(\mathbf{r}, t)$,

$$P(\mathbf{r}|\mathbf{r}', t) = \sum_{n=0}^{\infty} u_n^*(\mathbf{r}) u_n(\mathbf{r}') \exp(-Dk_n^2 t) \quad (1.53)$$

where, again, the $u_n(\mathbf{r})$ are the orthonormal set of solutions to the Helmholtz equation parametrised by the eigenvalue $-k_n^2$. They are further subject to the identity

$$\delta(\mathbf{r}' - \mathbf{r}) = \sum_{n=0}^{\infty} u_n^*(\mathbf{r}) u_n(\mathbf{r}') \quad (1.54)$$

$P(\mathbf{r}|\mathbf{r}', t)$, thus constructed, satisfies the initial condition

$$P(\mathbf{r}|\mathbf{r}', t=0) = \delta(\mathbf{r}' - \mathbf{r}) \quad (1.55)$$

The eigenvalues λ_n depend on the boundary condition. For reflective walls,

$$\hat{\mathbf{n}} \cdot \nabla P(\mathbf{r}|\mathbf{r}', t) = 0 \quad (1.56)$$

where $\hat{\mathbf{n}}$ is the outward surface normal. For the case of permeable or absorbing walls, the boundary condition becomes

$$D\hat{\mathbf{n}} \cdot \nabla P(\mathbf{r}|\mathbf{r}', t) + \bar{\rho} P(\mathbf{r}|\mathbf{r}', t) = 0 \quad (1.57)$$

1.3.4 Relationship of diffusion to velocity autocorrelation function

Equation 1.43 provides a definition for diffusion of the form

$$D = \lim_{t \rightarrow \infty} \frac{1}{2} \frac{\partial \langle X^2(t) \rangle}{\partial t} \quad (1.58)$$

where $X(t) = x - x_0 = \int_0^t v(t') dt'$, the displacement over time t . This relation leads directly to the result

$$D = \lim_{t \rightarrow \infty} \int_0^t \langle v(\tau) v(0) \rangle d\tau \quad (1.59)$$

As an exercise, a detailed proof is given as follows:

$$\langle X^2(t) \rangle = \langle \left(\int_0^t v(t') dt' \right)^2 \rangle \quad (1.60)$$

whence¹²

$$\begin{aligned}
 \frac{\partial \langle X^2(t) \rangle}{\partial t} &= 2 \langle v(t) \int_0^t v(t') dt' \rangle \\
 &= 2 \int_0^t \langle v(t) v(t') \rangle dt' \\
 &= 2 \int_{-t}^0 \langle v(t) v(t+\tau) \rangle d\tau \\
 &= 2 \int_0^t \langle v(t) v(t-\tau) \rangle d\tau
 \end{aligned} \tag{1.61}$$

For a stationary ensemble, the correlation function is independent of starting time t and depends only on the offset τ , and so $\langle v(t) v(t-\tau) \rangle = \langle v(0) v(-\tau) \rangle$. Using that, and the fact that the correlation function is symmetric with respect to time, eqns 1.61 and 1.58 reduce to eqn 1.59.

1.3.5 The diffusion tensor

The definition of the scalar diffusion coefficient in eqn 1.59 allows for a more general description of a tensor $\underline{\underline{D}}$, the elements of which represent time integrals of the various auto- and cross-correlation functions possible when components of a velocity vector \mathbf{v} are allowed. Namely

$$\underline{\underline{D}} = \lim_{t \rightarrow \infty} \int_0^t \langle \mathbf{v}(\tau) \mathbf{v}(0) \rangle d\tau \tag{1.62}$$

with elements:

$$D_{\alpha\beta} = \lim_{t \rightarrow \infty} \int_0^t \langle v_\alpha(\tau) v_\beta(0) \rangle d\tau \tag{1.63}$$

In molecular Brownian motion, orthogonal components of velocity are uncorrelated and the diffusion tensor is diagonal. For free diffusion those diagonal elements are equal. By contrast, in an anisotropic medium with axial or biaxial symmetry, such as an oriented liquid crystal or nerve tissue, $\underline{\underline{D}}$ will be diagonal in the principal axis frame, with unequal elements D_{xx} , D_{yy} , and D_{zz} . Of course, in a rotated frame of reference, off-diagonal elements will be apparent even though there exists a similarity transformation to a diagonal representation. As a consequence of the microreversibility of non-equilibrium thermodynamics [18], in the case of self-diffusive motion, $\underline{\underline{D}}$ is symmetric, meaning $\underline{\underline{D}} = \underline{\underline{D}}^T$.

Note, however, that asymmetric and intrinsically non-diagonal properties are possible. For molecules undergoing dispersive flow (see Chapter 2) orthogonal components of velocity may indeed be correlated so that an inherently non-diagonalisable $\underline{\underline{D}}$ matrix is possible.

¹²We use the fact that both integration and differentiation are linear operations and that $\langle \dots \rangle$ is formed from a sum. Hence the ensemble averaging may be interchanged with those operations.

1.3.6 The Smoluchowski equation

Suppose that the particle is subject to an additional potential $U(\mathbf{r})$, with associated force $\mathbf{F} = -\nabla U$. Then this force would result in a velocity $\mathbf{v} = \mathbf{F}/\zeta$. Hence eqn 1.29 needs to be modified to allow for the additional particle flux $n\mathbf{v}$, so that

$$\mathbf{J} = -D\nabla n - \frac{n}{\zeta}\nabla U \quad (1.64)$$

Using the Einstein result, $D = k_B T/\zeta$ and again applying the continuity requirement $\nabla \cdot \mathbf{J} = \partial n / \partial t$, the diffusion equation may now be written

$$\frac{\partial n}{\partial t} = \frac{1}{\zeta} \nabla \cdot [k_B T \nabla n(\mathbf{r}, t) + n(\mathbf{r}, t) \nabla U] \quad (1.65)$$

Equation 1.65 is known as the Smoluchowski equation [19]. Following the same reasoning as in the previous section, we may also write this differential equation in terms of the conditional probability

$$\frac{\partial}{\partial t} P(\mathbf{r}_0 | \mathbf{r}, t) = \frac{1}{\zeta} \nabla \cdot [k_B T \nabla P(\mathbf{r}_0 | \mathbf{r}, t) + P(\mathbf{r}_0 | \mathbf{r}, t) \nabla U] \quad (1.66)$$

Note that we can use eqn 1.64 to derive the Einstein diffusion coefficient result by simply assuming a thermal equilibrium state in which the flux must vanish. In equilibrium, a time-independent Boltzmann distribution must apply, namely $n(\mathbf{r}, t) \sim \exp(-U(\mathbf{r})/k_B T)$. Setting eqn 1.64 to zero, the result $D = k_B T/\zeta$ follows.

1.3.7 The Langevin equation

A quite different approach to the problem of Brownian motion was demonstrated by Paul Langevin in 1908 [20]. Langevin's starting point was to ascribe to the collisions of neighbouring molecules on the Brownian particle a random fluctuating force, and then to attempt a solution to Newton's second law. While Einstein described Brownian motion using a familiar linear partial differential equation, albeit applied to probability densities, Langevin's approach involved a differential equation that was inherently stochastic, and which required special methods for solution.

Again, we start with a simple one-dimensional (1-D) picture of the dynamics, but finish with a generalisation to three dimensions. The fluctuating force we call $F(t)$ while, as previously, we write the particle displacement as $x(t)$. We will, however, allow for the existence of an external force, F_{ext} , noting that we may set this force to zero in the case of simple self-diffusion. Allowing for the frictional drag term, $\zeta \frac{dx}{dt}$, Newton's law gives

$$m\ddot{x} = F_{\text{ext}} - \zeta\dot{x} + F(t) \quad (1.67)$$

where we use a single dot superscript to represent the time derivative d/dt and two dots for d^2/dt^2 . Since $F(t)$ is a stochastic random variable, so will be the solution to the 1-D Langevin equation, $x(t)$. Before attempting a solution for the probability distribution

$p(x, t)$, we will attempt a simple analysis based on ensemble averages, $\langle \dots \rangle$, and in the case where the external force is zero. Multiplying eqn 1.67 by x we find

$$m x \ddot{x} = -\zeta x \dot{x} + x F(t) \quad (1.68)$$

or

$$m \left[\frac{d}{dt} (x \dot{x}) - \dot{x}^2 \right] = -\zeta x \dot{x} + x F(t) \quad (1.69)$$

We now calculate the ensemble average of eqn 1.69,

$$m \left[\frac{d}{dt} \langle x \dot{x} \rangle - \langle \dot{x}^2 \rangle \right] = -\zeta \langle x \dot{x} \rangle + \langle x F(t) \rangle \quad (1.70)$$

The term $m \langle \dot{x}^2 \rangle$ is twice the ensemble-averaged kinetic energy associated with one of the ‘energy storing modes’ available for 3-D motion, namely for that associated with motion along the x -axis. By the Boltzmann equipartition theorem, it will be twice $k_B T/2$. The term $\langle x F(t) \rangle$ is intriguing. There is no reason to expect that $x(t)$ and $F(t)$ will be correlated, although the detailed argument in support of this is delicate [21]. And if uncorrelated we may write $\langle x F(t) \rangle = \langle x(t) \rangle \langle F(t) \rangle$; given that both $\langle x(t) \rangle$ and $\langle F(t) \rangle$ are zero, the term may be neglected.

This leaves us with

$$m \frac{d}{dt} \langle x \dot{x} \rangle - k_B T = -\zeta \langle x \dot{x} \rangle \quad (1.71)$$

and its simple exponential solution

$$\langle x \dot{x} \rangle = C \exp(-t/t_\zeta) + \frac{k_B T}{\zeta} \quad (1.72)$$

where $t_\zeta = m/\zeta$. This is a curious equation, implying a subtle correlation between x and \dot{x} . Of course, at long times, if the particle has diffused along the positive x -axis, then presumably that displacement arises from a predominantly average positive value of \dot{x} . Setting $x(0) = x_0 = 0$ at the time origin requires $C = -k_B T/\zeta$, and we may write $x = X$, the displacement from origin.

Of course $\langle X \dot{X} \rangle = \frac{1}{2} \frac{d}{dt} \langle X^2 \rangle$, and so integrating eqn 1.72 from 0 to t results in

$$\langle X^2 \rangle = \frac{2k_B T}{\zeta} [t - t_\zeta (1 - \exp(-t/t_\zeta))] \quad (1.73)$$

This leads us to two limiting cases

$$\lim_{t \ll t_\zeta} \langle X^2 \rangle = \frac{k_B T}{m} t^2 \quad (1.74)$$

$$\lim_{t \gg t_\zeta} \langle X^2 \rangle = 2 \frac{k_B T}{\zeta} t \quad (1.75)$$

We can identify the time $t_\zeta = m/\zeta$ as the collision time between molecules. At short times, the motion is velocity-like (ballistic) ($X^2 \sim t^2$), while beyond the collision time, as shown in Figs 1.11 and 1.12, the motion is diffusive ($X^2 \sim t$), with the diffusion coefficient given by the Einstein result, $D = k_B T/\zeta$. This long time limit is where the friction dominates eqn 1.67 and the inertial term on the left-hand side of the equation may be neglected.

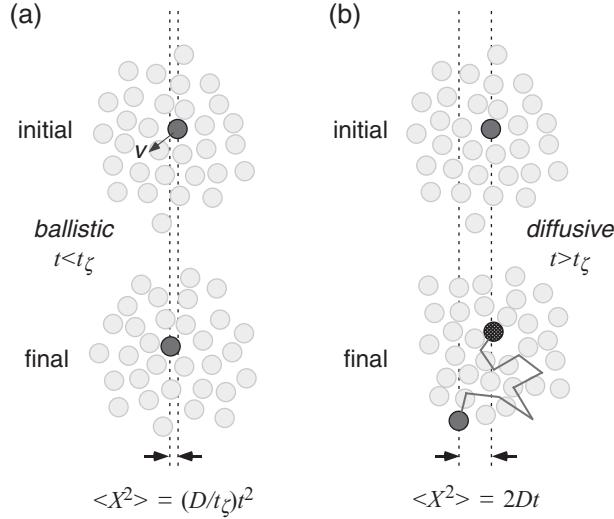


Fig. 1.11 Dense gas of particles with thermal energy, showing the transition from ballistic to diffusive motion for a labelled (black) particle. In (a), the time elapsing between the initial and final state is less than t_ζ and the particle moves with rms velocity $v = \sqrt{2k_B T/m}$ while in (b), where $t > t_\zeta$, the labelled particle suffers many collisions and executes a random walk.

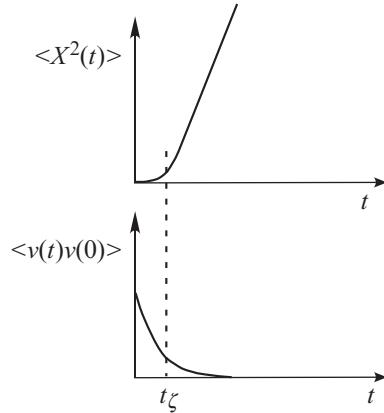


Fig. 1.12 Mean-squared displacement and velocity autocorrelation function showing the transition from ballistic to diffusive motion for colliding molecules at $t = t_\zeta$.

1.3.8 Correlation and the fluctuating force

We now return to an examination of the fluctuating force. In the case of the Langevin force its average value, $\langle F(t) \rangle$, is zero. Its temporal character may be described by the correlation function $\langle F(t) F(t') \rangle$. Restricting ourselves to the non-inertial limit, $t \gg t_\zeta$, we may treat the correlation time for fluctuations in $F(t)$ as infinitesimal and so the correlation function may be written

$$\begin{aligned}\langle F(t) F(t') \rangle &\sim \delta(t - t') \\ &= 2\zeta k_B T \delta(t - t')\end{aligned}\quad (1.76)$$

The constant multiplying the Dirac delta function is justified as follows. Setting $F_{\text{ext}} = 0$, $x(0) = x_0$, and neglecting the inertial term in eqn 1.67 we find, on integrating,

$$x(t) - x_0 = \zeta^{-1} \int_0^t F(t') dt' \quad (1.77)$$

Assuming that $F(t)$ is a Gaussian variable, then its integral, as represented by eqn 1.77, is also Gaussian,¹³ so that we may write the probability distribution of $x(t) - x_0$ as

$$\Psi(x(t) - x_0, t) = (2\pi B)^{-\frac{1}{2}} \exp\left(-\frac{(x - x_0)^2}{2B}\right) \quad (1.78)$$

where $B = \langle (x(t) - x_0)^2 \rangle$. Evaluating the integral

$$\langle (x(t) - x_0)^2 \rangle = \zeta^{-2} \left\langle \int_0^t F(t') dt' \int_0^t F(t'') dt'' \right\rangle \quad (1.79)$$

yields the Einstein result $\langle (x(t) - x_0)^2 \rangle = 2k_B T t / \zeta$ as required.

Equation 1.78 mirrors the Gaussian form of the conditional probability for free diffusion found earlier, and of course, $\Psi(x - x_0, t)$ obeys the same Fick's law differential equation. Indeed, given that we have specified the start position as x_0 , the distribution $\Psi(x - x_0, t)$ must be identical to $P(x_0|x, t)$, the conditional probability that it moved to x after time t if it started at x_0 . The distribution on $x(t)$ for a range of start points x_0 may be written

$$p(x(t)) = \int p(x_0) P(x_0|x, t) dx_0 \quad (1.80)$$

Again, specification of the starting position is equivalent to setting $p(x_0)$ as a delta function, so that $p(x(t))$ becomes identical to the conditional probability.

At this point it is helpful to introduce the probability that a particle displaces by X over a time interval t . This distribution, $\bar{P}(X, t)$, is known in magnetic resonance parlance as an 'average propagator' and, as in the case of eqn 1.80, is a sum over all possible start positions

$$\bar{P}(X, t) = \int p(x_0) P(x_0|x_0 + X, t) dx_0. \quad (1.81)$$

The use of the capital letter for X provides a means of reminding us that we are dealing with a displacement from origin, with $\bar{P}(X, t)$ being an ensemble average, $\langle \dots \rangle_{x_0}$, over all possible origins x_0 . The argument is set out in one dimension, but it can be simply extended to three, where \mathbf{R} becomes the displacement and \mathbf{r}_0 the origin vector.

¹³This assertion follows from the central limit theorem [22].

1.3.9 Ornstein–Uhlenbeck process

The problem described by eqns 1.74 and 1.75 is known as an ‘Ornstein–Uhlenbeck’ process or a ‘mean-reverting process’ [23]. Such problems involve an exponential drift towards a mean behaviour with random fluctuations superposed. In the above problem we see this in the crossover from velocity-like behaviour to diffusive behaviour on a timescale long compared with the particle collision time.

The Gaussian distribution of $\langle F(t) \rangle$ and the linear nature of the Langevin equation for free diffusion means that the probability distributions for displacements are also Gaussian at all times, and not just in the diffusion limit, $t \gg t_\zeta$. Indeed we may write the average propagator for the Brownian Ornstein–Uhlenbeck process as

$$\bar{P}(X, t) = \left(4\pi D \left[t - t_\zeta \left(1 - e^{-t/t_\zeta} \right) \right] \right)^{-\frac{1}{2}} \exp \left(-\frac{X^2}{4D \left[t - t_\zeta \left(1 - e^{-t/t_\zeta} \right) \right]} \right) \quad (1.82)$$

We will revisit these ideas when we come to consider dispersion in the next chapter.

1.3.10 Diffusion in a harmonic potential

There are many examples in nature where diffusing particles are confined in space. A very simple example, again of the Ornstein–Uhlenbeck class, is the case of a diffusion under the simultaneous influence of a harmonic potential. We may write this potential $U(x) = \frac{1}{2}kx^2$, with its associated restoring force $F = -kx$, where k is the force constant.

The Langevin equation for a particle diffusing under such a force is simply

$$m\ddot{x} = -kx - \zeta\dot{x} + F(t) \quad (1.83)$$

In the non-inertial limit, $t \gg t_\zeta$, eqn 1.83 reduces to

$$\zeta\dot{x} = -kx + F(t) \quad (1.84)$$

with solution

$$x(t) = \frac{1}{\zeta} \int_{-\infty}^t \exp \left(-\frac{(t-t')}{\tau_k} \right) F(t') dt' \quad (1.85)$$

where $\tau_k = \zeta/k$. The correlation function $\langle x(t)x(0) \rangle$ may be easily calculated using the relation $\langle F(t')F(t'') \rangle = 2\zeta k_B T \delta(t' - t'')$ as

$$\begin{aligned} \langle x(t)x(0) \rangle &= \frac{1}{\zeta^2} \int_{-\infty}^t \int_{-\infty}^0 \exp \left(-\frac{(t-t'-t'')}{\tau_k} \right) \langle F(t')F(t'') \rangle dt' dt'' \\ &= \frac{k_B T}{k} \exp \left(-\frac{t}{\tau_k} \right) \end{aligned} \quad (1.86)$$

The relative positions of the Brownian particle gradually decorrelate at a rate determined by k/ζ , the ratio of the force constant to the frictional drag.

An expression for $P(x_0|x, t)$ applicable at all times longer than the collision time t_ζ is given in reference [12]. In this non-inertial limit, but for $t \ll \tau_k$, the diffusing particle

has insufficient time to feel the effects of the harmonic potential and the conditional probability is the same as for free diffusion

$$P(x_0|x,t) = (4\pi Dt)^{-\frac{1}{2}} \exp\left(-\frac{(x-x_0)^2}{4Dt}\right) \quad (1.87)$$

In the long time limit, $t \gg \tau_k$, the probability distribution on x is time-independent and simply given by the normalised Boltzmann distribution

$$P(x) = \left(\frac{k}{2\pi k_B T}\right)^{\frac{1}{2}} \exp\left(-\frac{kx^2}{2k_B T}\right) \quad (1.88)$$

The very same function applies independently for the probability distribution of starting positions x_0 . Indeed so long as $t \gg \tau_k$, the conditional probability is independent of the starting position of the particle and as well,

$$P(x_0|x,t) = \left(\frac{k}{2\pi k_B T}\right)^{\frac{1}{2}} \exp\left(-\frac{kx^2}{2k_B T}\right) \quad (1.89)$$

Clearly $P(x)$ and $P(x_0)$ obey eqn 1.80. This time-independence of the conditional probability at long time is a feature of restricted diffusion in a bounded system. Note the form of the average propagator in this limit,

$$\bar{P}(X,t) = \int P(x_0) P(x_0 + X) dx_0 \quad (1.90)$$

This integral is the spatial autocorrelation function of the equilibrium density.

References

- [1] D. Lindley. *Boltzmann's Atom*. Free Press, New York, 2001.
- [2] G. M. Wang, E. M. Sevick, E. Mittag, D.J. Searles, and D.J. Evans. Experimental demonstration of violations of the second law of thermodynamics for small systems and short time scales. *Phys. Rev. Lett.*, 89:050601, 2002.
- [3] Lewis Wolpert. In praise of science. In Ralph Levinson and Jeff Thomas, editors, *Science Today*. Routledge, London and New York, 1997.
- [4] A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen (english: On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat). *Annalen der Physik*, 17:549, 1905.
- [5] R. Brown. A brief account of microscopical observations made in the months of June, July and August, 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Phil. Mag.*, 4:16, 1829.
- [6] A. Einstein. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt (english: On a heuristic viewpoint concerning the production and transformation of light.). *Annalen der Physik*, 17:132, 1905.

28 Thermal processes and diffusion

- [7] C. Kittel and H. Kroemer. *Thermal Physics*. W.H. Freeman, New York, 1980.
- [8] L. Boltzmann. Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie. *Wiener Berichte*, 53:195–220, 1866.
- [9] J. C. Maxwell. On the dynamical theory of gases. *Philosophical Transactions of the Royal Society*, 157:49, 1867.
- [10] S. Otto. *Lehrbuch der Thermochemie und Thermodynamik (English: A Text Book of Thermo-Chemistry and Thermodynamics)*. BiblioLife (2009), 1912. Original publisher: Springer (1912), English translation: MacMillan (1917).
- [11] H. Tetrode. Die chemische Konstante der Gase und das elementare Wirkungsquantum. *Annalen der Physik*, 343(7):434, 1912.
- [12] R. Kubo, M. Toda, and N. Hashitsume. *Statistical Physics II: Non-Equilibrium Statistical Mechanics, 2nd Edition*. Springer, New York, 1991.
- [13] M. Toda, R. Kubo, and N. Saito. *Statistical Physics I: Equilibrium Statistical Mechanics, 2nd Edition*. Springer, New York, 1992.
- [14] A. Fick. Über Diffusion. *Phil. Mag.*, 10:30, 1855.
- [15] G. G. Stokes. On the effect of the internal friction of fluids on the motion of pendulums. *Cambridge Philosophical Society Transactions*, 9:8, 1851.
- [16] W. Sutherland. The measurement of large molecular masses. *Australasian Association for the Advancement of Science, Report of Meeting*, 10:117, Dunedin, 1904.
- [17] W. Sutherland. A dynamical theory of diffusion for nonelectrolytes and the molecular mass of albumin. *Philosophical Magazine*, 6:781, 1905.
- [18] L. Onsager. Reciprocal relations in irreversible processes, i. *Phys. Rev.*, 37:405–426, 1931.
- [19] M. Smoluchowski. Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen. *Annalen der Physik*, 21:756, 1906.
- [20] P. Langevin. Sur la theorie du mouvement brownien. *C. R. Acad. Sci. Paris*, 146:530, 1908.
- [21] A. Manoliu and C. Kittel. Correlation in the Langevin theory of brownian motion. *Am. J. Phys.*, 47:678, 1979.
- [22] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 1. Wiley, 3 edition, 1968.
- [23] G. E. Uhlenbeck and L. S. Ornstein. On the theory of Brownian motion. *Phys. Rev.*, 36:823, 1930.