Part I

Historical overview and early developments

1

Limitations imposed by the second law of thermodynamics

Brownian ratchets are devices that rectify microscopic fluctuations, thus producing useful work out of a fluctuating environment. The term "Brownian ratchet" highlights two important elements required for the rectification of fluctuations to take place. First, a fluctuating environment should be present, and indeed the term "Brownian" refers to the archetype of a fluctuating environment: Brownian motion, that is the zig-zag motion of a grain of pollen in a fluid, or more generally the random motion of small particles as the result of the multiple collisions with the molecules surrounding them. The term "ratchet" highlights the second requirement for the rectification of fluctuations: the presence of appropriate asymmetries in the system, so to define a preferential direction of motion. However, there is a third important requirement for the implementation of a Brownian ratchet: the system has to be out of thermal equilibrium. In fact, for systems at equilibrium, the second law of thermodynamics prevents the generation of directed motion out of unbiased fluctuations. This is precisely the topic explored in this chapter.

The basic ideas of a Brownian ratchet goes back to the early twentieth century, when Smoluchowski analyzed a simple mechanical device involving a ratchet and a pawl. This thought experiment, later popularized by Feynman in his book *The Feynman Lectures on Physics, Vol. 1* (1962), was introduced to illustrate the limitations imposed by the second law of thermodynamics. As we shall show later, the implications of the second law in this example seem quite counter-intuitive, at first.

The operating principle of the ratchet machine is the same as that of an electrical rectifier, which was studied by Brillouin in 1950. Both are elementary examples of devices that, if they could perform as intended, harvesting thermal fluctuations from their environment to produce work, would be in violation of the second law of thermodynamics.

1.1 The second law of thermodynamics

Based on empirical evidence, the second law is a postulate of thermodynamics that limits the occurrence of many processes we know from experience do not happen, even though they are allowed by other laws of physics. For example, the water in a glass at room temperature is never seen to cool itself spontaneously to form ice cubes, releasing energy to its environment. Such transformation satisfies the law of conservation of energy, yet it is common sense it never occurs.

Though it may be expressed in several ways, the first formulation of the second law goes back to Sadi Carnot in 1824, who put a limit in the efficiency of any heat engine operating between two given temperatures. The typical efficiencies of Brownian ratchets are discussed in Chapter 6. In this chapter we are specially interested in the equivalent statement of the second law given by Lord Kelvin, which can be formulated as

There is no thermodynamic transformation whose sole effect is to extract heat from a heat reservoir and to convert it entirely into useful work.

A heat reservoir is a system so large that the exchange of heat does not change its temperature. Another common formulation involves the concept of entropy[1]:

The entropy of an isolated system never decreases.

Specializing the isolated system to a sub-system and its surroundings, the previous statement implies that the total entropy, also called the entropy of the universe, never decreases. From this formulation it is clear that, unlike the underlying molecular laws, the second law is not time-symmetric, i.e., is not invariant under a time reversal transformation, and thus displays a preferential direction (or arrow) of time.

1.2 Brillouin paradox

Diodes, the electronic components that allow the electrical current to pass mostly in one direction, were discovered in the second half of the nineteenth century. Their use experienced a boom in the 1950s due to substantial advances in the manufacture of semiconductor diodes, made today the most common type of diode.

The intrinsic asymmetry associated with diodes allows one to examine, with a simple system, important foundational questions regarding the possibility of rectifying fluctuations and the related limitations imposed by the second law of thermodynamics.

In this context, in 1950 the French physicist Léon Brillouin introduced what is now known as *Brillouin paradox*: a simple device that intuitively would lead to the

[1] see Chapter 6 for the definition of entropy in the context of Brownian motion.

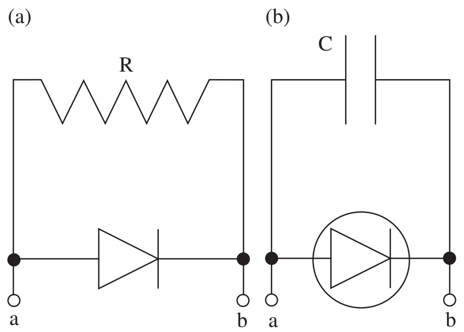

Figure 1.1 (a) The Brillouin paradox, as originally formulated: can a diode rectify the thermal fluctuations in the resistor and produce a direct voltage, like a small battery? (b) The variant of the paradox introduced by Alkemade, with the resistor replaced by a capacitor and a specific type of diode, a vacuum diode, considered.

rectification of fluctuations in a system in equilibrium, in striking violation of the second law of thermodynamics. The device is a simple circuit consisting of a diode and a resistor, as shown in Fig. 1.1(a). At a finite temperature the voltage across the resistor is fluctuating. An intuitive, though wrong, analysis leads immediately to the conclusion that the asymmetry introduced by the diode allows the establishment of a direct current in the circuit, and thus a direct voltage. This voltage would represent a source of electrical, useful energy, obtained from a single heat reservoir (the circuit) at thermal equilibrium, in violation with the second law. A more careful *microscopic* analysis reveals instead how no direct current can be induced in the circuit, thus re-estalishing the agreement with the second law.

Instead of discussing the microscopic analysis of the original Brillouin paradox, we find more illustrative to describe here an evolution of the idea, i.e. a variant of the Brillouin paradox introduced by Alkemade and van Kampen. Let us consider the circuit of Fig. 1.1(b). A capacitor is connected to a *vacuum* diode. In analogy with the original Brillouin paradox, one may wonder whether the rectifying properties of the diode allow the accumulation of a finite charge on the capacitor plates.

An initially uncharged capacitor is considered. At any finite temperature T, the capacitor will hold a small random charge Q due to thermal noise. In the absence of the diode, this charge will fluctuate in time between positive and negative values, so that the (time) average charge is zero. The voltage associated with this charge at a given instant is $V_b - V_a = Q/C$, where C is its capacitance. From the equipartition theorem, the average energy stored in the capacitor will be $\langle Q^2/2C\rangle = k_B T/2$, where $k_B = 1.38065 \times 10^{-23}$J/K is the Boltzmann constant.

However small, it is tempting to try to transform this thermal energy into useful work by connecting a diode to each end of the capacitor. Every fluctuating current with the same direction as the diode that spontaneously appears in the system will

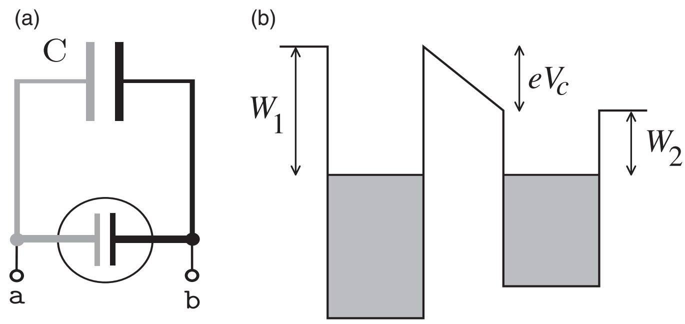

Figure 1.2 Alkemade diode: two metals with different work functions.

be allowed to cross the diode, whereas currents with the opposite directions are impeded. Thus, the asymmetry introduced by the electrical rectifier should result in a voltage offset $V_b - V_a > 0$ that could be used to perform work. This system would be able to extract heat from one heat reservoir and convert it entirely into useful work, in contradiction with the second law of thermodynamics. But then, why is the diode not able to rectify thermal fluctuations? To answer this question, the system has to be examined in more detail, i.e. a *microscopic* analysis of the system is required.

The essential rectifying element of the considered circuit is the vacuum diode, as illustrated in Fig. 1.2. It consists of two parallel electrodes, separated by a short distance, short enough to allow a flux of electrons between the two electrodes. The electrodes are in equilibrium at temperature T, but have different work functions W_1 and W_2. The vacuum diode is connected to an ideal capacitor, of capacitance C. We assume that the capacitance of the diode can be neglected against that of the capacitor.

The natural tendency to equilibrium will produce a net flow of electrons from the electrode with higher Fermi level (i.e., higher chemical potential or lower voltage) to the one with lower Fermi level (lower chemical potential or higher voltage). Thermal equilibrium is obtained when the two Fermi levels are equal, and the net flux of electrons ceases. However, this does not imply that at equilibrium the electric potential in the vacuum between the electrodes is flat. On the contrary, in general a potential difference inside the vacuum between the two plates, and thus also an electric field, is established at equilibrium. This electric field is responsible for the diode's characteristic asymmetry, acting as an effective resistance that slows down electrons in the direction that defines the diode, as we discuss later. It can be traced back to the different work functions of the two metal plates.

The work function W is the minimum energy required to remove an electron from the interior of the metal. Then, the requirement of equilibrium leads to the establisment of a potential difference $V_c = (W_1 - W_2)/e$, as indicated in Fig. 1.2, where $e = 1.60217657 \times 10^{-19}$ C is the electron charge. In metals V_c is commonly referred to as the *Volta potential*, and it is usually a good approximation to assume that the work function is not affected by the presence of other metals. More generally, the potential difference V_c between the surfaces of two arbitrary materials is known as the *contact potential*, and it usually depends on factors such as the distance between the materials and surface contamination.

At equilibrium, the contact potential V_c is actually generated by the charge distribution induced in the circuit due to the flow of electrons. Indeed, during the process to reach equilibrium, electrons will move from the right electrode to the left, going from the diode to the capacitor's plates. This transfer of electrons will charge the capacitor up to the equilibrium charge CV_c. Note that we are neglecting any charge accumulation in the surfaces of the diode's plates because the capacitance of the capacitor is assumed to be much larger than that of the diode. In contrast, in semiconductor diodes, which are made from the union of two different semiconductors, such as a p-n junction, the contact potential is usually created by charges accumulated at the surfaces in the junction of the two materials, the so-called *depletion layer*. A similar situation would happen here if we bypassed both capacitor plates.

It is now possible to examine the role of thermal fluctuations around such an equilibrium situation. In thermal equilibrium, the density probability for a given energy fluctuation E is proportional to the Bolztmann factor $\exp(-E/k_B T)$. An electron at the left electrode needs to overcome a barrier of W_1 to be able to cross to the right electrode. Thus, at equilibrium the probability for an electron at the left to cross is precisely $\exp(-W_1/k_B T)$. On the other hand, an electron at the right electrode needs a minimum energy of $W_2 + eV_c = W_1$ to cross to the left electrode. Thus, the probability is exactly the same to cross to the right as to cross to the left, implying no average current at the diode, and no rectification of thermal fluctuations. This restores the agreement with the second law of thermodynamics, which indeed prevents the extraction of work from a single bath at a single temperature.

There are of course situations in which direct currents occur, without any violation of the second principle of thermodynamics. The first situation corresponds to the standard use of a diode in a circuit, where it allows flow of charge in the presence of a bias potential. A direct current can be forced through the diode by applying an external voltage bias V to the diode, such as to reduce the potential barrier in the vacuum between both electrodes. If $V < V_c$, the energy barrier for a jump from the left electrode is the same as before, but the barrier an electron faces at the right electrode is reduced by $-eV$, implying an increased probability $\exp[-(W_1 - eV)/k_B T]$ for a jump to the left. Since the flow in each direction should be proportional to each probability, at equilibrium we should expect a net current

$$I = A\left(\exp(eV/k_B T) - 1\right), \tag{1.1}$$

Cambridge University Press
978-1-107-06352-5 - Brownian Ratchets: From Statistical Physics to Bio and Nano-motors
David Cubero and Ferruccio Renzoni
Excerpt
More information

where A is a constant. Equation (1.1) is the typical $I - V$ characteristic of most diodes such as a metal-oxide or a p-n junction. For a forward bias $V > 0$, the current grows exponentially, while for a reverse bias $V < 0$ the current goes backward, eventually approaching a (considerably smaller) constant value. Therefore, the external bias is thus acting by reducing ($V > 0$) or increasing ($V < 0$) the effective resistance associated with the contact potential. Furthermore, the constant energy power supplied externally by the voltage bias prevents the formation of a surface charge that would stop the current.

A direct current can also be produced without an external voltage bias by keeping both electrodes at different temperatures, say $T_1 < T_2$. The system is in this way out of equilibrium, and no violation of the second law takes place. Assuming a situation of local equilibrium at each electrode, the current in each direction should be approximately proportional to the corresponding equilibrium probability of crossing. Since $T_2 > T_1$, the probability to jump to the left $\exp(-W_1/k_B T_2)$ is larger than to the right $\exp(-W_1/k_B T_1)$, and a forward current should be expected. In this situation, the diode is actually acting as a small battery, with the energy power coming from thermal fluctations via the temperature gradient. This phenomenon is known as the themoelectric or Seebeck effect, being discovered by Seebeck in 1821. Today it is the operating mechanism of applications such as thermocouples, a widely used type of temperature sensor.

The analysis carried out so far showed that intuition may lead to incorrect predictions – the so called Brillouin paradox – in clear violation of the second law of thermodynamics. A correct microscopic analysis leads instead to predictions in agreement with the second law. It is interesting to revisit this *intuitive* approach, which leads to incorrect predictions, from a microscopic point of view, so to highlight which assumptions lead to results in contradiction to the second law. To this end, we re-examine the circuit of Fig. 1.2.

An appreciable defect of charge in the capacitor, $\langle \tilde{Q} \rangle = CV_c - \langle Q \rangle$, will create a forward potential bias $V = \langle \tilde{Q} \rangle / C$. Thus, from (1.1) we obtain an equation for the average charge at the capacitor,

$$\frac{d\langle \tilde{Q} \rangle}{dt} = -A \left[\exp\left(\frac{e\langle \tilde{Q} \rangle}{Ck_B T} \right) - 1 \right], \tag{1.2}$$

valid to describe the dynamics of the macroscopic fluctuation $\langle \tilde{Q} \rangle$ in out-of-equilibrium situations. It could be tempting to use this equation to model bare fluctuations, for example by considering it valid for $\tilde{Q}$ and adding a noise term $\xi(t)$,

$$\frac{d\tilde{Q}}{dt} = -A \left[\exp\left(\frac{e\tilde{Q}}{Ck_B T} \right) - 1 \right] + \xi, \tag{1.3}$$

where $\langle \xi(t) \rangle = 0$. The fluctuations $\tilde{Q}$ should be small, so we could expand the exponential in (1.3) to

$$\frac{d\tilde{Q}}{dt} = -A\left(\frac{e}{Ck_BT}\tilde{Q} + \frac{e^2}{2(Ck_BT)^2}\tilde{Q}^2\right) + \xi, \tag{1.4}$$

However, this stochastic equation does not admit (nor does (1.3)) the expected thermal equilibrium solution, with $\langle \tilde{Q} \rangle_{\text{eq}} = 0$ and $\langle \tilde{Q}^2 \rangle_{\text{eq}} = Ck_BT$. This is obvious by taking averages in (1.4). Rather, if nonzero values of $\langle \tilde{Q} \rangle_{\text{eq}}$ are considered, equation (1.4) gives

$$\langle \tilde{Q} \rangle_{\text{eq}} = -\frac{1}{2}\frac{e}{Ck_BT}\langle \tilde{Q}^2 \rangle_{\text{eq}} = -e/2\,, \tag{1.5}$$

a defect charge that would violate the second law. Such an unphysical result is nothing else than the Brillouin paradox, as derived via an incorrect microscopic analysis. Both equations (1.4) and (1.3) are incorrect. It is not legitimate to generalize a nonlinear phenomenological equation valid for macroscopic quantities, such as (1.2), to study fluctuating quantities at the microscopic scale. This is a procedure that generally leads to contradictions, as shown in the earlier derivation. It is worth highlighting that it is the nonlinearity of equation (1.2) that leads to paradoxical results when its validity is extended to microscopic fluctuations. The use of the linearized version of the phenomenological equation (1.2) would not have led to paradoxical results when applied to microscopic fluctuations. In other words, (1.4) would be correct if the quadratic term in $\tilde{Q}$ appearing in the right hand side of the equation is neglected, removing also the apparent paradox.

1.3 Feynman ratchet

The ratchet and pawl system of Smoluchowski and Feynman is the mechanical version of the elecrical rectifier examined in Section 1.2. In fact, most of the quantitative analysis carried out for the diode is the same Feynman did for the ratchet system in his famous lectures.

The mechanical rectifier is shown in Fig. 1.3. An axle has some vanes attached at one side. The collisions of the molecules in the surrounding air with the vanes will produce a Brownian motion in the vanes that will be transmitted to the rotational motion of the axle. To simplify matters, all solid parts in this system are assumed to be perfectly rigid. At the other end of the axle there is a ratchet wheel, i.e., an asymmetric gear specifically designed to allow rotation in one direction only. Motion of the ratchet in the opposite direction is forbidden by the pawl. As it is usual in these mechanical systems, the pawl is loaded with an internal spring in order to engage the ratchet's teeth, to prevent movement in the backward direction.

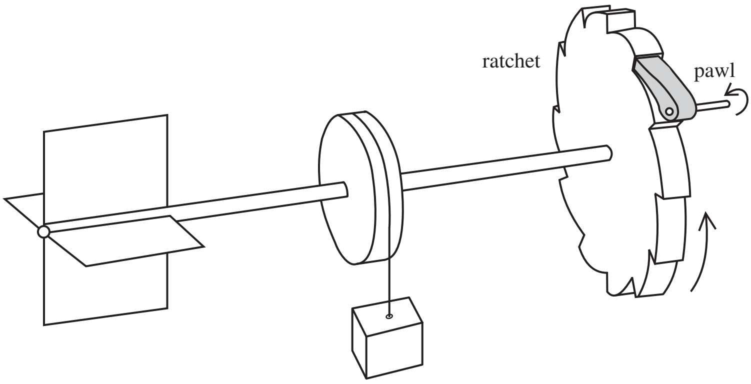

Figure 1.3 Smoluchowski and Feynman's ratchet and pawl system.

Though not explicitly drawn, the effect of this spring is indicated as a small curved arrow in Fig. 1.3.

The molecular collisions taking place on the side of the vanes that would rotate the wheel in the backward direction are absorbed by the pawl. In contrast, the molecular collisions in the right side of the vanes are allowed to contribute to the rotation of the wheel in the forward direction. This motion could be used to do work on other systems, such as lifting a weight against gravity, as shown in Fig. 1.3. Since everything is at just one temperature, this machine would be obviously violating the second law of thermodynamics. Clearly, the microscopic analysis described earlier must be incorrect. Just as in the Brillouin paradox, the problem stems from the extension of nonlinear averaged mechanisms, valid at macroscopic level, to the realm of microscopic fluctuations.

The previous analysis did not take into account the fact that the pawl must also be subject to thermal fluctuations. Since it has a spring, there must also be a damping mechanism to avoid the pawl to indefinitely bounce with the ratchet wheel. This implies a fluid surrounding the pawl, and thus a fluctuating pawl. Nevertheless, the quantitative analysis at thermal equilibrium is fairly independent of the specific details of the molecular interaction. If ϵ is the minimum energy required to lift the pawl (the amount necessary for the wheel to advance one notch), then $\exp(-\epsilon/k_B T)$ will be both the probability that enough energy is gathered at the vanes to rotate the wheel one notch, and the probability that the pawl is accidentally lifted by thermal fluctuations at the wheel, allowing backward motion. Like in the electrical rectifier, the net result should be no motion.

A similar analysis shows that if the temperature at the vanes T_1 is kept at a different value than at the wheel T_2, say $T_1 > T_2$ (or $T_1 < T_2$), then the ratchet should be rotating at constant speed in the forward (backward) direction, working