1

A primer on electron transport

1.1 Nanoscale systems

Let us briefly discuss the systems I will consider in this book, those of *nanoscale* dimensions (1 nm $= 10^{-9}$ m). The phenomena and theoretical approaches I will present are particularly relevant for these structures rather than those with much larger dimensions.

So, what is a nanoscale system? The simplest – and most natural – answer is that it is a structure with at least one dimension at the nanoscale, meaning that such dimension is anywhere in between a few tens of nanometers and the size of an atom (Di Ventra *et al.*, 2004a). One can then define structures with larger – but still not yet macroscopic – dimensions as *mesoscopic*. This separation of scales is arguably fuzzy. Mesoscopic structures share some of the transport properties of nanostructures; the theoretical description of both classes of systems is often similar; and in certain literature no distinction between them is indeed made.

Is there then, in the context of electrical conduction, another key quantity that characterizes nanoscale systems? As I will emphasize several times in this book, this key quantity is the *current density* – current per unit area – they can carry. This can be extremely large.

As an example, consider a wire made using a mechanically controllable break junction (Muller *et al.*, 1992), a junction that is created by mechanically breaking a metal wire. Such a structure – a type of metallic *quantum point contact* – may result in a single atom in between two large chunks of the same material (see schematic in Fig. 1.1). If a typical current of 1μA is set to flow across the system, at the atom position, considering a cross section of 10 Å^2, we would expect a current density of about 1×10^9 A/cm^2! These current densities are typically orders of magnitude larger than those found in mesoscopic/macroscopic systems.
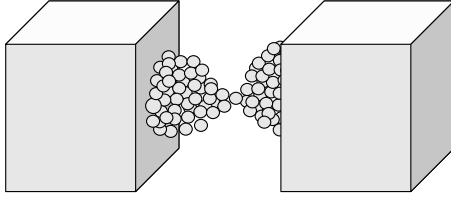
1

A primer on electron transport

Fig. 1.1. Schematic of an atomic metallic quantum point contact.

A large current density implies a large number of scattering events per unit time and unit volume. This means that interactions among electrons, or among electrons and ions are particularly important.

Note that I have said nothing about how fast a single electron "crosses" nanoscale structure. This *transit time* may be extremely short. However, due to the large current density the *cumulative* effect of all electrons is to amplify electron-electron and electron-ion interactions locally in the nanostructure. For instance, as I will discuss later in the book, both ions and electrons heat up locally in the junction above their nominal background temperature, thus affecting its structural stability under current flow.

nanotubes/nanowires   molecular wires/quantum dots
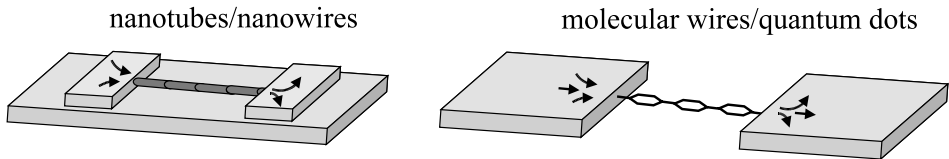
Fig. 1.2. Left panel: A quasi-1D wire laid on top of a surface and in between two bulk electrodes. Right panel: A single molecule between bulk electrodes. The arrows indicate the direction of charge flow.

I mention here a few other nanoscale systems of present interest. These include *nanotubes* or long *atomic wires* in contact with metal electrodes (see Fig. 1.2, left panel), and small *molecules* sandwiched between bulk metals (Fig. 1.2, right panel). These systems may sometimes be referred to as *quantum dots* if their bonding to the electrodes is very weak (these concepts will become clearer as we go along with the book).

Many other structures – and their combinations – that confine electrons in one or more dimensions can be fabricated. These systems represent ideal test beds to understand electron and ion dynamics at these length scales, and may find application in the broadly defined field of optoelectronics, or even in biotechnology and medicine. The latter point is particularly relevant nowadays as the conducting properties of DNA and its single units – called ucleotides – are being studied for possible use in sequencing technology (Zwolak and Di Ventra, 2008).
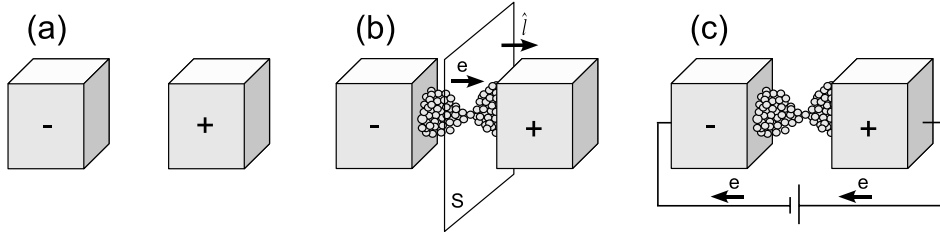
Fig. 1.3. (a) Two finite electrodes charged differently. (b) The electrodes are connected via a junction. (c) Symbol used to represent a battery.

The above list of examples can go on and on. To illustrate the phenomena discussed in this book, I will refer to selected nanoscale systems whose properties have been studied experimentally. The choice of these specific examples reflects both their pedagogical appeal and, of course, the author's taste.

1.2 Generating currents

Before considering the different approaches one can employ to formulate the transport problem in nanoscale systems, let us ask a basic question that will guide us in developing theories of charge transport, namely

How do we generate electrical currents?

There are several answers to this question suggesting different ways to describe the corresponding transport problem.

Let us start from the simplest experimental realization possible. Consider two large but *finite* electrodes and charge them differently: one has more electrical charge than the other, or equivalently one is charged negatively, the other positively. How we charge these electrodes is irrelevant. For instance, we can do it by simply rubbing them with some other material!

I label the one with more electrons with a "−" sign, the other with a "+" sign (see Fig. 1.3(a)). At equilibrium, and due to the conducting nature of each electrode, the extra charge (whether negative or positive) is found on the surface of the electrodes. The electrons in the "−" electrode have higher potential energy (and lower electrostatic potential) than the electrons in the "+" electrode. A potential difference, or *bias*, between the two electrodes has been created so that if we put them in contact by means of some other conducting material, electrons will flow from the "−" region to the "+" region (see Fig. 1.3(b)). *Electrical current* has been thus established.

A primer on electron transport

We can now take an arbitrary surface S cutting the conductor in between the two electrodes and determine the rate at which electrons cross this surface (see Fig. 1.3(b)). Let us indicate with $\hat{\mathbf{l}}$ the unit vector perpendicular to an infinitesimal element dS of the surface S. At every instant of time, and at every point in space, we can assign to the electrons a velocity

$$\mathbf{v}_{\text{tot}}(\mathbf{r}, t) = \mathbf{v}_{th}(\mathbf{r}, t) + \mathbf{v}_{drift}(\mathbf{r}, t) \tag{1.1}$$

which is sum of a randomly oriented[1] velocity, $\mathbf{v}_{th}$, due to thermal fluctuations, and a component, $\mathbf{v}_{drift}$, we call *drift velocity*, that, on average, points in the direction of global electron flow.[2] If we average the above velocity ver all particles, only the average drift velocity is different from zero,

$$\mathbf{v}(\mathbf{r}, t) \equiv \langle \mathbf{v}_{th}(\mathbf{r}, t) \rangle + \langle \mathbf{v}_{drift}(\mathbf{r}, t) \rangle = \langle \mathbf{v}_{drift}(\mathbf{r}, t) \rangle, \tag{1.2}$$

where the bracket operation $\langle \cdots \rangle$ means average over the *ensemble* of particles. If $n(\mathbf{r}, t)$ is the *number density* of carriers with charge e at any given oint in space and any given instant of time,[3] the amount of charge dQ that crosses the surface dS in an infinitesimal time dt is

$$dQ = en\mathbf{v} \cdot \hat{\mathbf{l}}\, dS\, dt. \tag{1.3}$$

The current across the surface dS is thus

$$dI = \frac{dQ}{dt} = en\mathbf{v} \cdot \hat{\mathbf{l}}\, dS \equiv \mathbf{j} \cdot d\mathbf{S}, \tag{1.4}$$

where I have defined the *current density* vector

$$\mathbf{j}(\mathbf{r}, t) = en(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t), \tag{1.5}$$

and the surface vector $d\mathbf{S} = \hat{\mathbf{l}}\, dS$. The total average current across the surface S is then the integral

$$I_S(t) = \int_S d\mathbf{S} \cdot \mathbf{j}(\mathbf{r}, t), \tag{1.6}$$

where the subscript S is to remind us that, in general, the total current depends on the chosen surface.[4]

Randomly oriented means that the thermal velocity has an isotropic spatial distribution.
In the next chapter I will discuss physical reasons why the drift velocity does not increase indefinitely in time.
For electrons, I choose the convention $e = -|e|$. Note that the standard convention for the direction of current is opposite to the electron flow direction.
In strictly 2D systems the surface integral in the definition 1.6 is replaced by a line integral, and the current density 1.5 is defined with the 2D number density. In 1D the current and the current density are the same quantity, with the definition 1.5 containing the 1D number density.

Polarization and magnetization

In addition to "free" charges, which are able to move across macroscopic regions of the system, there may be bound charges and localized currents, i.e., charges and currents localized to microscopic regions of the sample. Bound charges give rise to *polarization*, and localized currents to *magnetization*. Let us call q_i these bound charges at position $\mathbf{r}_i$, and $\mathbf{v}_i$ their velocities. From classical electrodynamics we know that the polarization is (Jackson, 1975)

$$\mathbf{P}(\mathbf{r}) = \frac{1}{V} \sum_i q_i \, \mathbf{r}_i, \tag{1.7}$$

i.e., the average dipole moment per unit volume, and the magnetization

$$\mathbf{M}(\mathbf{r}) = \frac{1}{2cV} \sum_i q_i \, (\mathbf{r}_i \times \mathbf{v}_i), \tag{1.8}$$

is the average magnetic moment per unit volume, where the sums extend over all charges in a volume V centered at position $\mathbf{r}$, and c is the speed of light. The total current density is (Jackson, 1975)

$$\mathbf{j}_{\text{tot}}(\mathbf{r}, t) = en(\mathbf{r}, t)\mathbf{v}(\mathbf{r}, t) + \frac{\partial \mathbf{P}}{\partial t} + c\nabla \times \mathbf{M}, \tag{1.9}$$

and the total density

$$en_{\text{tot}} = en - \nabla \cdot \mathbf{P}. \tag{1.10}$$

The total current is given again by Eq. 1.6, the integral of the total current density over a surface S. In the following, when discussing densities and current densities, I will always refer to the total density 1.10 and current density 1.9, even though the contributions to the current from polarization and magnetization are, in most of the systems and conditions discussed in this book, small.[5]

In addition, the current density 1.9 generates a magnetic field, which "acts back" on the current, and thus modifies it. This *self-consistent* effect is known as *magnetic screening*. It is generally a small effect for the systems I consider in this book, for the same reasons that the magnetic current

---

[5] For instance, at steady state the polarization current is zero; for non-magnetic materials and small magnetic fields, the magnetization is small. Indeed, for paramagnetic and diamagnetic materials $|\mathbf{M}| \propto \chi|\mathbf{B}|$, with $\mathbf{B}$ an external magnetic field, and χ is the magnetic susceptibility. χ is of the order of $10^{-3} - 10^{-5}$ cm^3/mol (Ashcroft and Mermin, 1975) so that the magnetization is a very small fraction of the field. We will see in Chapter 7 that stochastic time-dependent current density functional theory provides a formally – in principle – exact way to calculate all contributions (from free and bound charges and localized currents) to the current density of a many-body system using effective single-particle equations.
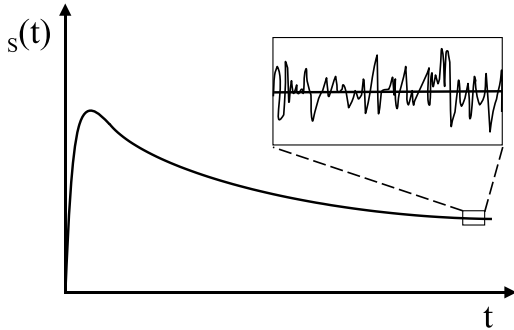
Fig. 1.4. The current as a function of time during the discharge of a capacitor. The current fluctuates around some average value as shown schematically in the inset.

contribution in Eq. 1.9 is small. I thus either neglect it altogether,[6] or assume the current has been determined self-consistently with this magnetic screening effect included.

To the above we need to add the *continuity equation*[7]

$$e\frac{\partial n(\mathbf{r},t)}{\partial t} = -\nabla \cdot \mathbf{j}(\mathbf{r},t) \qquad (1.11)$$

that states the conservation of charge.

Fluctuations and reservoirs

The *gedanken* experiment I have outlined describes the discharge of a capacitor across a conductor. We then know from experiments that as time goes on the current $I_S(t)$ decays (see Fig. 1.4) so that, if we wait enough time, no current (in a time-averaged sense) will flow across the surface S: all negative charges on the "$-$" electrode have neutralized all positive charges on the "$+$" electrode, and the electrode-conductor-electrode system ends up in *global equilibrium*.

We can however make this decay time longer and longer by increasing the size of the electrodes while keeping the conductor in between them unchanged. This corresponds to increasing the number of carriers that are stored in the two electrodes. In classical circuit theory this corresponds to increasing the *capacitance* of the capacitor while keeping the electron *resis-*

This is clearly a theoretical statement: in experiments one cannot eliminate this screening effect.

Since the divergence of the curl of a vector is zero, the continuity equation is identically satisfied between the total current 1.9 and the total density 1.10:

$$e\frac{\partial n_{\text{tot}}(\mathbf{r},t)}{\partial t} = -\nabla \cdot \mathbf{j}_{\text{tot}}(\mathbf{r},t).$$

tance to flow from one plate of the capacitor to the other constant. Intu-
itively, one may guess that if the dimensions of the electrodes were set to go
to infinity while keeping the conductor and its contact with the electrodes
unchanged, the average current would not change in time, and its value
would be independent of the chosen surface S used to evaluate Eq. 1.6.[8]

In reality, the electrical current continually *fluctuates* about an average
value (see inset of Fig. 1.4). If we take a time interval T large enough that
many current fluctuations occur in that interval, but not so large that the
decay of the current is appreciable, we can define

$$\langle I_S \rangle = \frac{1}{T} \int_{-T/2}^{+T/2} dt \, I_S(t) \tag{1.12}$$

as the average current in that interval.[9] If we take the limit of infinite
electrodes first, we can then take the limit of $T \to +\infty$. We thus realize the
current-carrying steady state condition

$$\frac{d\langle I_S \rangle}{dt} = 0 \qquad \text{stationarity condition,} \tag{1.13}$$

in which the time derivative of the *average* total current is zero.[10]

The above limit of infinite electrodes is also the first encounter with the
theoretical concept of *reservoir*, which I will develop further in Chapter 3.
Here, I give the following definition

Reservoir: *An ideal "system" that can supply and receive an arbitrary
amount of carriers and energy without changing its internal
state.*

If this "ideal system" exchanges only energy, and not particles, with some
other physical system – whose dynamics we are interested in – I shall call it
bath and not reservoir. Both the bath and the reservoir embody the notion
that physical systems are never truly *closed*; they always have some degree
of interaction with an external *environment*.[11]

---

[8] Note, however, that the infinite electrode size is no guarantee that the stationary condition 1.13
is satisfied (see discussion in Sec. 7.6).

[9] Note that an analogous averaging process is always performed by the actual apparatus measur-
ing the current. This is due to the fact that such apparatus has a finite frequency bandwidth
$\Delta \nu$. Here I am assuming that $1/\Delta \nu \ll T$.

[10] I anticipate here that the time average 1.12 in the limit of $T \to +\infty$ may be replaced by
the average over all possible electron state configurations (*ensemble-averaged* current). This is
known as the *ergodic hypothesis*. I will come back to it in Chapter 2 where I will also discuss
the conditions for its validity.

[11] Consider for instance electrons in a solid kept at a finite temperature. The finite temperature
is that of an external bath. For this bath to be called as such, the dynamics of the electrons
cannot affect its properties, namely its temperature.

1.2.1 Finite versus infinite systems

Indeed, the experimental fact that the discharge of a *finite* capacitor ends up in a state of global equilibrium (no average current flows) is a consequence of the fact that the electrons in the capacitor are coupled to some external environment.[12] They thus lose energy during the discharge. If the electron system is truly *closed*, namely electrons do not interact with any external environment – and they are not subject to any external field[13] – their total energy is conserved, and their dynamics is *deterministic* in the sense that e can (in principle) solve exactly their equation of motion at all times – the Schrödinger equation 1.16 in the quantum case. We then know that the following *Poincaré recurrence theorem* applies: there exists a time – called the *recurrence time* – after which the system will return to the neighborhood of its initial conditions. "Neighborhood" means as close as possible to the initial conditions.

For the case at hand, this implies that a finite, closed and isolated capacitor will eventually return to its initial state: it will recharge itself! This apparent paradox, however, does not pose any practical worry. Even if the electron system were truly closed, its recurrence time would be extremely long.

For clarity, consider the quantum case. The many-body states of the capacitor are eigenstates of some Hamiltonian $\hat{H}$. Since we are dealing with large number N of *interacting* electrons (of the order of an Avogadro's number $N_A = 6 \times 10^{23}$ particles/mol) the spectrum of $\hat{H}$ is extremely *dense*, namely the separation between the many-body energy states dE is extremely small. If ΔE is a representative energy interval of the system,[14] and M the number of states in that interval, it is indeed easy to argue that the separation between many-body energy states is of the order of (Landau and Lifshitz, 1959a)

$$dE \sim \frac{\Delta E}{M} \sim \Delta E \, e^{-S/k_B} \sim \Delta E \, e^{-N}, \qquad (1.14)$$

where S is the statistical entropy I will discuss in Sec. 2.8, $k_B = 1.38 \times 10^{-23}$ J/K is the Boltzmann constant. In the above I have used the fact that the

The ionic vibrations of the underlying lattice may be such an environment if they form a *dense* energy spectrum, and no correlated state between electrons and vibrations forms. Otherwise, energy will flow back and forth between electrons and ionic vibrations, and the following considerations apply to this combined system.

Which means the system is also *isolated*. In this book, I will choose the convention of calling a system *closed* when it is not interacting *dynamically* with an environment, and *isolated* when it is not subject to external *deterministic* forces, like a classical electromagnetic field.

Statistically, this interval represents the width of the energy distribution of the system, like the bandwidth in the single-particle case.

number of states is exponentially related to the statistical entropy[15] and the latter is an *extensive variable*, namely it increases with the dimensions of the system, or, equivalently, with the number of particles.

The recurrence time is thus of the order of

$$\Delta t_{rec} \sim \frac{\hbar}{dE} \sim \frac{\hbar}{\Delta E}\, e^N, \tag{1.15}$$

which for a macroscopically large number of particles may be much larger than the age of the universe![16] For all practical purposes, there is no way for us to observe the reverse process, the recharge of a closed capacitor, and the dynamics of the system is thus practically *irreversible*.[17]

The above point is interesting for another reason. If I keep the electron system closed, but I make it truly *infinite*, namely with an infinite number of particles, the spectrum of its Hamiltonian becomes a continuum, $dE \to 0$, and the recurrence time tends to infinity, $\Delta t_{rec} \to +\infty$. This means that an infinite system will never return to its initial state, and therefore we do not need to worry about recurrence times: *the dynamics of an infinite system is intrinsically irreversible.* In this case, the energy is *diluted* into an infinite number of *degrees of freedom* (property of a bath, as I have discussed above), and if we added a particle to it we would not be able to modify the internal state of the system (property of a reservoir), and that particle would be "lost" into the reservoir, in the sense that, almost immediately, we would not be able to follow its dynamics.[18]

1.2.2 Electron sources

In real life there is clearly no such a thing as an infinite system and in order to set current flowing one generally relies on batteries attached to the electrodes. The batteries are devices that, via internal chemical processes, continually charge the surfaces of the electrodes so that a constant potential difference can be applied at the two contacts between the electrodes and

---

[15] I will demonstrate this explicitly in Sec. 2.8, Eq. 2.144. Here I just need to use this property to show that an infinite capacitor has an infinite recurrence time.

[16] As an example, consider a system with only 100 electrons, whose many-body spectrum is distributed over $\Delta E = 1$ eV of energy. From Eq. 1.14, $dE \approx 10^{-44}$ eV, and the recurrence time, Eq. 1.15, is $\Delta t_{rec} \sim \hbar/dE \approx 10^{28}$ s, or about 10^{20} years! This number has to be compared with the age of the universe which is about 10^{18} s.

[17] This result applies even if we could solve *exactly* the equation of motion of a macroscopic number of particles. Needless to say, this is a practically impossible task. Approximations to the exact dynamics introduce *loss of information*, and thus provide another source of irreversibility of dynamical systems (see Sec. 2.8).

[18] This statement is obvious in the quantum case where electrons are *indistinguishable*. It is also true in classical mechanics where, for all practical purposes, it is impossible to follow the dynamics of a single particle in an ensemble of infinite identical particles.

the battery, with consequent current flow across the electrode-conductor-electrode structure. In other words, the battery creates a capacitor-like situation as discussed above, and, in addition, *maintains* it till the chemical processes run out. I will use the conventional symbol for a battery as represented in Fig. 1.3(c).

Electrons from the negative side of the battery cross the whole length of the structure to be collected at the other positive end of the battery. On verage, at the same time electrons are injected from the negative side into the structure thanks to the electrochemical reaction inside the battery. If for moment we put quantum mechanics aside and could label single electrons, e could easily argue that the ones that get into the positive side of the battery are not necessarily the same ones that get out of the battery from its negative side. Quite generally, we would need *inelastic* processes to carry electrons *inside* the battery from its positive terminal to the negative one. Inelastic means that any one electron changes its energy in going from the ositive terminal to the negative terminal inside the battery.

Most importantly, once an electron is collected at the positive end of the battery, its subsequent dynamics is practically impossible to follow, and this electron is "lost" into the battery, most likely without changing the internal state of the latter. Similarly, we can assume that an electron leaving the battery does not modify its state considerably. In other words, the battery effectively acts as a reservoir of electrons.

1.2.3 Intrinsic nature of the transport problem

Are these the only ways we can generate current? Not quite. For instance, we can go back to our finite electrode-conductor-electrode system of Fig. 1.3(b) after the discharge is complete, i.e., the system is at equilibrium. We can then immerse it in an oscillating electric field. Electric harges would then respond to the field and would start moving from one electrode to the other, generating current. One can also generate current in metallic ring threaded by a magnetic field (see, e.g., Kamenev and Kohn, 2001), and so on and so forth.

All this seems awfully complicated to describe theoretically. There appear to be so many ways we can generate currents, and so many processes we need to take into account (e.g, the chemical processes in the battery). And have not even discussed what happens in the conductor itself, or how we *probe* these currents!

This leads us to the following question: What is our main concern when e want to study transport properties of a given system? For instance, are