Nanophysics

Lecture notes for the course Kyl-0.108

Fall 2005

Tero Heikkilä

Compiled on May 15, 2007

The most recent update of this text can be downloaded at
http://ltl.tkk.fi/opetus/nanophysics/notes/nanophysicsfall05.pdf. Use the account
"nano" and the password "elektroni".

Contents

Chapter 1

Introduction

These lecture notes deal with the electron transport phenomena in small, "mesoscopic" systems. The word "mesoscopic" comes from the Greek word "mesos", which means "middle", indicating that the mesoscopic world resides between the microscopic and the macroscopic. Typically this means that the phenomena take place in systems consisting of a large number of atoms and electrons, but they only occur if the system is "small enough", i.e., the size of the system is smaller than some length scale characterizing the transition from the microscopic to the macroscopic world.

Let us take a few examples that are discussed in the remainder of this course. In a perfect metallic crystal, electron transport is essentially dissipationless, i.e., the conductivity is infinite. But there are no everyday-scale perfect crystals. The dislocations, vacancies and other impurities distort the crystal periodicity and lead to a finite conductivity. Also the presence of lattice vibrations, phonons, induce scattering that again decreases conductivity. These processes come with a length scale, a mean free path, that essentially characterizes the average distance between the impurities, or between subsequent scatterings of conduction electrons from them, or from the phonons. A metallic wire which is smaller than this mean free path is said to be ballistic. In a ballistic wire, the resistance does not depend on the length of the wire (up to the mean free path, of course), and therefore the conductivity can be said to be infinite. As explained below, however, even with this infinite conductivity, the measured resistance of such a wire is nonzero, as the finite number of quantum channels in this wire leads to a finite resistance. Anyway, in this example the crossover from mesoscopic (where a "macroscopic" observable such as resistance is well defined, but it is determined by the quantum nature of the wire) to macroscopic (where resistance scales with the length of the wire) is characterized by the mean free path.

Another example concerns the definition of electron temperature. From the statistical physics one learns that temperature characterizes the width of the electron energy distribution function turning from one (below the Fermi energy E_F) to zero (above E_F). This temperature is a parameter in the Fermi-Dirac distribution function. In a nonequilibrium setting, an applied voltage V making

rise to the current I leads to the power $P = VI$ applied to the sample, and the sample heats up. But as shown in Chap. 4, the resulting distribution function may not always be of the Fermi-Dirac form, and one can have many different types of definitions for the electron temperature, leading in general to different numerical values depending on the definition. This nonequilibrium form persists in the wire as long as the electrons relax into the usual equilibrium shape. Such a relaxation is again characterized by a relaxation length, and this again describes a crossover between mesoscopic and macroscopic regimes.

Especially the length scales concerning energy relaxation are strongly temperature dependent. For example, for a typical copper wire (see details in Sec. II.C.2 of the review on thermal effects by Giazotto, *et al.* [1] – in this particular case the electron-phonon scattering length is proportional to $T^{-3/2}$), the relaxation length at room temperature is of the order of a few nanometers, at T =10 K, it is 700 nm, at T =1 K it already reaches 20 μm, and for the minimum typically achievable temperature, T =10 mK, it is a "macroscopic" length, 2 cm. This shows that mesoscopic effects are best seen at low temperatures. Typically such low temperatures are reached with different liquid Helium refrigerators, whose operation is an everyday task in the experimental mesoscopic physics research groups.

In general, the studied object can be almost anything that conducts electrons. In mesoscopics, there are three main types of systems depending on the main materials used for their construction:

1. Perhaps the simplest studied systems are small metallic wires and metal-to-metal contacts fabricated with a help of a lithographically patterned mask through which the metals are evaporated. Typically used metallic materials are based on aluminum, copper, gold, niobium, silver or palladium, but also other metals are used. As aluminum and niobium turn superconducting at sub-kelvin temperatures (the critical temperature of Al is 1.1 K and for Nb it is 9.3 K), a combination of superconducting and normal-metallic (nonsuperconducting) effects can be studied using them. Most single-electron transistors (SETs) and superconducting Josephson junctions are fabricated with metals. For an example of a system consisting of two metallic wires, see Figs. 1.1,1.2.

2. With semiconducting materials, one has to use more massive and expensive devices than with metals, but the resulting systems are also generally more controllable. One can, for example, realize low-dimensional wires and 0-dimensional dots by imposing gates on systems fabricated typically on a 2-dimensional electron gas (2DEG) formed in a layer between two semiconductors with different conduction and valence band energies (see Fig. 1.3). This means that at low temperatures, there are extended electronic states only in one or two dimensions, i.e., electrons can move only in those directions. With semiconductors, one can realize quantum

[1] F. Giazotto, T. T. Heikkilä, A. Luukanen, A. M. Savin, and J. P. Pekola, Rev. Mod. Phys. **78**, 217 (2006).

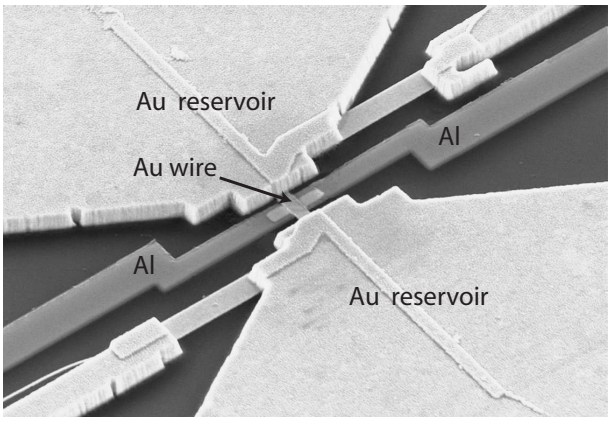


Figure 1.1: Scanning electron micrograph (SEM) of a nonequilibrium Josephson transistor, containing a normal-metal (Au) weak link between two Al superconductors. The weak link is further connected to large reservoirs to ensure proper thermalization and control of the electron distribution function. The thickness of the wires is a few tens of nanometers, width of the order of 100 nm, and length of the order of a micrometer. See J. J. A. Baselmans, Nature **397**, 43 (1999). Courtesy of Jochem Baselmans.

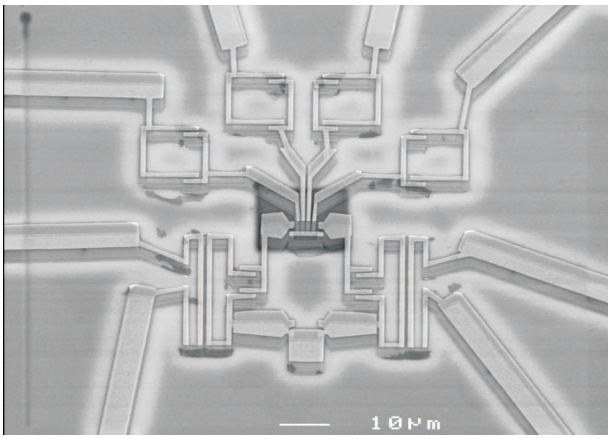


Figure 1.2: SEM picture of a setup used in the ongoing research (2005) in the Low Temperature Laboratory to study the coupling between the electrons and the photon field. The bright areas denote Cu wires and the darker regions underneath them Al wires. Both are evaporated on a Si substrate as is the case with majority of mesoscopic samples. Courtesy of Matthias Meschke.

dots (QDs), quantum point contacts (QPCs), ballistic quantum wires, etc. (See Fig. 1.4 for an example.) These systems are described during this course. On the other hand, heavily doped semiconductors can also be used to replace metals in the study of metallic properties with slightly different (and slightly more tunable) parameters than in real metals. Typically used semiconductors are based on silicon or gallium arsenide (GaAs).

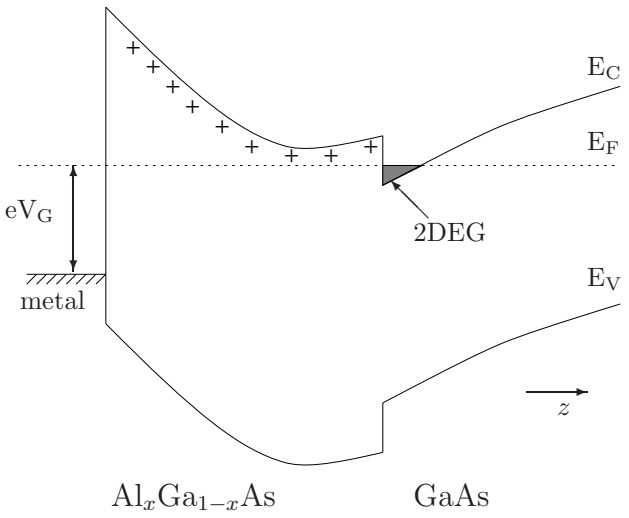


Figure 1.3: Band-bending diagram of modulation doped GaAs-$Al_xGa_{1-x}As$ heterostructure. A 2DEG is formed in the undoped GaAs at the interface with the p-type doped AlGaAs.

3. There are ways to contact molecules between metallic leads and study transport through them. Often these molecules are based on the carbon structures found in the beginning of 1990's; fullerenes (C_{60}) or especially carbon nanotubes (see Fig. 1.5), but also the electric current through a single Hydrogen molecule has recently been measured.[2] The fullerenes and other fairly small molecules form essentially quantum dots, and the nanotubes can be used to construct single-electron transistors or quantum wires.

For the study of electron transport through a given nanostructure, one also has to fabricate the contact of this structure to the macroscopic measuring devices. Typically this is done by structuring much wider electrodes than the studied wires in contact to them (see Figs. 1.1-1.5 for examples). Such an electrode ideally works as a *reservoir* (heat bath) of electrons: once the electrons

[2] R. H. M. Smit, *et al.*, "Measurement of the conductance of a Hydrogen molecule", Nature **419**, 906 (2002).

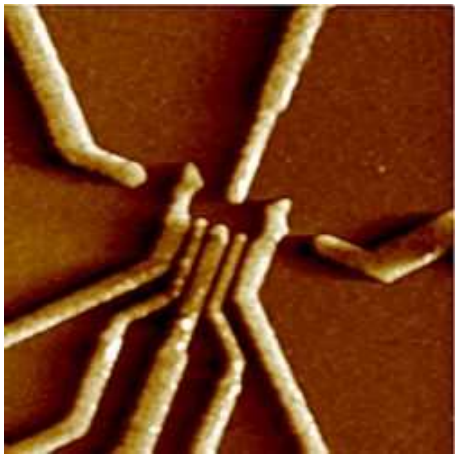


Figure 1.4: Atomic force microscope (AFM) image of a quantum dot defined with electrostatic gates patterned on top of a 2-dimensional electron gas which forms in a layer between different materials with different conductance and valence band energies (see Fig. 1.3). The middle "fingers" can be used to accurately tune the number of electrons on the dot. This device also has two quantum point contacts patterned in the sides of the dot. The typical size of the dot ranges from some ten nanometers to microns. From http://qt.tn.tudelft.nl/research/qdots/.

enter the electrode, they quickly thermalize with the lattice. In practice, this means that their energy distribution function obtains the Fermi-Dirac form

$$f(E) = \frac{1}{\exp[(E-\mu)/(k_B T)] + 1}, \tag{1.1}$$

where μ is the chemical potential and T the lattice temperature of the reservoir. In the ideal case, the reservoir is then unperturbed by what happens in the nanostructure. These wide electrodes continue for some hundreds of microns and connect to contact pads, the latter with dimension of the order of a millimeter. To these one solders the external wires, which then contact the sample to the measurement apparatus. At this point there are often a few amplification stages, either in the Helium cryostat or in room temperature. As the experiments are done at low temperatures, the thermal noise in the measurement apparatus (residing at room temperature) and the electronic heat current through the wires aim to heat up the electrons in the sample. Therefore, one typically needs a good thermal contact of the wires with the low-temperature equipment, and several electronic filters for the noise between the sample and the room-temperature equipment.

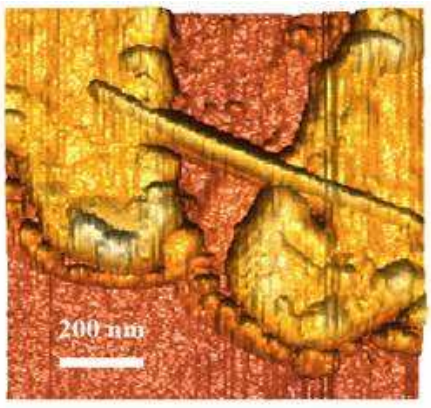


Figure 1.5: AFM picture of a free-standing multi-wall carbon nanotube, acting as a single-electron transistor (SET). See L. Roschier, et al., Microelectric Engineering **61-62**, 687 (2002). Courtesy of P. Hakonen, Low Temperature Laboratory, TKK.

Length	Symbol	order of magnitude
Fermi wavelength (metals)	λ_F	0.1-1 nm
Fermi wavelength (2DEG)	λ_F	10-100 nm
Elastic scattering length (metals)	$\ell_{\rm el}$	10-100 nm
Elastic scattering length (2DEG)	$\ell_{\rm el}$	$\ldots 10$ μm
Energy relaxation length at $T = 1$ K	$\ell_{\rm en}$	$1 \ldots$ few tens μm
Dephasing length at $T = 1$ K	ℓ_φ	$0.1 \ldots$ 10 μm
Thickness $\times$ width $\times$ length (metals)		30-50 $\times$ 100-500 $\times$ 500- $(\text{nm})^3$

Table 1.1: Orders of magnitude for some length scales in typically studied mesoscopic systems.

1.1 Classical vs. quantum transport

In the forthcoming lectures, we will discuss quantum-mechanical transport properties of nanostructures. These show deviations from the classical Ohm's law $I = U/R$, stating that the current I through a given sample is a linear function of the voltage U, the coefficient R being the resistance. In the classical case, this resistance scales linearly with the wire length,

$$R = \frac{L}{\sigma A}, \tag{1.2}$$

A being the cross section of the wire and σ its conductivity. In Chap. 4, we derive this law using semiclassical arguments: ignoring the interference effects, which often are weak, and single-electron effects, which would arise in strongly interacting systems (discussed in Chap. 5).

1.1.1 Drude formula and Einstein relation

The Boltzmann-equation analysis yields the Einstein relation for the conductivity (see Chap. 4),

$$\sigma = e^2 N_0 D = e^2 N_0 v_F \ell_{\rm el}/d, \tag{1.3}$$

where e is the electric charge, N_0 is the density of states at the Fermi level, D is the diffusion constant, v_F is the Fermi velocity, $\ell_{\rm el}$ is the elastic mean free path and d is the dimensionality of the wire.

There is a classical estimate for the conductivity that predates the Einstein relation, namely the Drude formula derived by Paul Drude in 1900, before the microscopic theories of solid state. It goes as follows. Assume an electron in a solid where one has applied an electric field $\vec{E}$. The electric field produces a Lorenz force $-e\vec{E}$, accelerating the electron. Assume that a time t has passed since the electron last collided with the lattice. The velocity of the electron has increased to $\vec{v}_0 - e\vec{E}t/m$ due to the Lorenz force, $\vec{v}_0$ being the initial velocity after the last collision. Assume that the electron collides in a random direction from the impurities. In this case the initial velocity will have no contribution to the average velocity of the electron, which must therefore be given as the average of $-e\vec{E}t/m$. Denote the average time between collisions as τ. This implies the average electron velocity given by[3]

$$\vec{v}_a = -\frac{e\vec{E}\tau}{m}. \tag{1.4}$$

The average current carried by the electrons is given by $\vec{j} = -en\vec{v}_a$, where n is the electron density. Therefore, we get $\vec{j} = \sigma_D \vec{E}$ with the Drude conductivity

$$\sigma_D = \frac{ne^2\tau}{m}. \tag{1.5}$$

This is the same as Eq. (1.3) (up to a prefactor) after one identifies $n \sim N_0 v_F^2 m/2 = N_0 E_F$.

1.1.2 Quantum effects

Typical quantum effects encountered in nanoelectronic systems arise due to the energy and/or charge quantization effects, tunneling, and interference effects. A large-scale quantum phenomenon is the transition to the superconducting state taking place in many materials at low temperatures. This transition is not a mesoscopic effect, but also in nanoelectronic circuits it makes rise to many types of phenomena, which would not be present in simple normal-metallic systems.

There are many possible reasons for breaking Ohm's law, i.e., either linearity between current and voltage or the scaling of the resistance with the length of the wire. Typically both of them are broken, but there are exceptions to this. Such non-classical effects are, for example,

[3] In fact, the average would be half this, as the average velocity is between 0 and that obtained on average when colliding with the next particle. But as τ is not so well defined here, such prefactors of order unity do not matter much.

- **Tunneling** through a thin insulating region or a vacuum gap between two metals. Such *tunnel barriers* are an important part of mesoscopic electronics and they can be used to realize many types of structures. A single tunnel barrier fabricated between two normal (nonsuperconducting) metals is typically a linear object, i.e., the resistance is independent of voltage up to very large voltages. However, the scaling of the resistance with the thickness of the tunnel barrier is exponential, not linear. Calculating the current due to tunneling through a single insulating barrier is illustrated below.

- In low-dimensional systems, such as quantum dots, electron **energy level quantization** within the system shows up in the current through the system.

- **Single-electron effects:** In the case when tunnel barriers have a small capacitance C, the energy required for charging the capacitance with a single electron, $E_C = e^2/(2C)$, may become relevant. In this case, to get a finite current through the tunnel contact, the external circuit must provide this *charging energy* E_C to the electrons crossing the barrier. It turns out that in a double-junction system ("single-electron transistor" SET, see Ch. 5), this energy has to be provided by the bias voltage or the temperature. Otherwise no current can flow, and the system is in the state of a *Coulomb blockade*.

- In metal or heavily-doped semiconductor wires whose size is smaller than the *phase relaxation length* ℓ_φ, **interference effects** between different electron paths within the wire may alter the conductivity. Such effects are Aharonov-Bohm effects in multiply connected systems under an applied magnetic field, and localization and universal conductance fluctuations in disordered systems. These effects are discussed in Ch. 3.

- In **ballistic systems**, i.e., systems whose size is smaller than the mean free path, the resistance no longer scales with the length of the wire, but it becomes quantized, depending on the ratio between the width of the wire and the Fermi wavelength λ_F of electrons (see Ch. 2).

- At a high magnetic field, *Landau levels* are formed within the conductor, and the resistance through the system is strongly field dependent. The resulting **quantum Hall effects** are not discussed during this course.

- In addition to all the above, in **superconducting** wires and junctions, and superconductor–normal metal heterostructures, quantum-mechanical phase coherent effects show up, and Ohm's law ceases to be valid.

As an example of calculating the current through a low-dimensional object, consider the case of a tunnel junction depicted in Fig. 1.6. The proper microscopic calculation of the current in this case is done in Ch. 5, but the formula for the current can also be argued based on a simple picture. With an applied voltage V through the junction, the Fermi levels of the electron systems

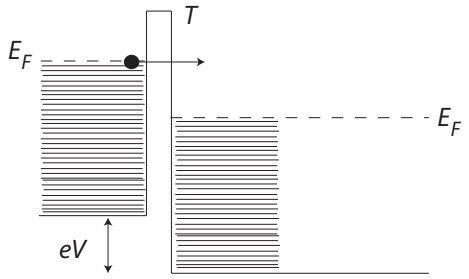


Figure 1.6: Schematic model of a tunnel junction with transparency T.

in the two sides of the junctions are shifted by eV. Assume that the average transmission probability for an electron on the left hand side of the junction to enter the right side is T (assumed independent of energy and not to be confused with the temperature!). This is essentially of the form $T \sim \exp(-2d\sqrt{2mU}/\hbar)$, where d is the width and U is the height (work function difference between the metal and the tunnel barrier) of the tunnel barrier, and m is the electron mass. At a vanishing temperature, the current from left to right at energy E is then $eAcTN_L(E)N_R(E+eV)$, where A is the area of the junction, $N_{L/R}(E)$ is the density of states in the left/right side of the junction and c is a constant fixing the dimensions of this expression to (current/energy). In what follows, we identify the prefactor as the resistance, and get rid of c. For a finite temperature, one has to include the occupation numbers (distribution functions) $f_i(E)$ of electrons in reservoir i: the initial state has to be filled and the final state has to be empty. Therefore, the total current (current from left to right minus that from right to left) is

$$\begin{aligned} I &= ecAT \int_{-\infty}^{\infty} dE N_L(E) N_R(E+eV) \{ f(E;\mu_L,T_L)[1-f(E;\mu_R,T_R)] \\ &\qquad - f(E;\mu_R,T_R)[1-f(E;\mu_L,T_L)] \} \\ &= ecAT \int_{-\infty}^{\infty} dE N_L(E) N_R(E+eV)[f(E;\mu_L,T_L) - f(E;\mu_R,T_R)]. \end{aligned} \tag{1.6}$$

Here $f(E;\mu,T) = (\exp((E-\mu)/(k_BT))+1)^{-1}$ and $\mu_R = \mu_L + eV$. The integrand is non-vanishing within the window of width $\sim \max(k_BT, eV)$ around μ_L and μ_R, i.e., around the Fermi energies E_F of the two leads.[4] If the reservoirs are normal metals and the voltage is much smaller than E_F/e, the densities of states are almost constant within this window, and can be replaced by their values at E_F, $N_{L/R}(E) \approx N_{L/R}(E_F) \equiv N_0^{L/R}$. Then the current is given by

$$I = ecAT N_0^L N_0^R \int_{-\infty}^{\infty} dE [f(E;\mu_L,T_L) - f(E;\mu_R,T_R)]. \tag{1.7}$$

[4] In the limit $k_BT \ll E_F$, E_F is the same as the chemical potential.

It is left as an exercise for the students to calculate this remaining integral and to show that the conductance of the tunnel junctions is independent of temperature and voltage. In the forthcoming chapters, the prefactor of this expression is often written in terms of the resulting resistance, $R_T = 1/(e^2 cAT N_0^L N_0^R)$. As the absolute zero point of energy is irrelevant, one often fixes one of the Fermi energies as the zero point. Therefore, below the point $E = 0$ refers to the Fermi energy.

Similarly, one may write for the heat current carried by the electrons from the left reservoir

$$\dot{Q}_L = \frac{1}{e^2 R_T} \int_{-\infty}^{\infty} dE(E - \mu_L)[f(E; \mu_L, T_L) - f(E; \mu_R, T_R)] \tag{1.8}$$

and from the right reservoir

$$\dot{Q}_R = \frac{1}{e^2 R_T} \int_{-\infty}^{\infty} dE(E - \mu_R)[f(E; \mu_L, T_L) - f(E; \mu_R, T_R)]. \tag{1.9}$$

The difference in $\dot{Q}_L$ and $\dot{Q}_R$ results from the Joule power $\dot{Q} = IV$ induced in the reservoirs.

1.2 Second quantization: short introduction

In various parts of this course, we will employ the formalism of second quantization. This short introduction follows the one given in the book "Theory of Nonequilibrium Superconductivity" (by N. Kopnin). The formalism is introduced separately for bosons and fermions.

1.2.1 Bosons

Consider a system of N Bose particles occupying quantum states with normalized wave functions $\phi_i(x)$, $i = 1, \ldots$. The state of such a system is described by a set of numbers N_i, which show how many particles are in state i. The normalized total wave function for the occupation numbers $N_1, N_2, \ldots$ is given by

$$\Phi_{N_1,N_2,\ldots} = \left(\frac{N_1! N_2! \ldots}{N!} \right)^{1/2} \sum_P \phi_{j_1}(x_1)\phi_{j_2}(x_2) \ldots \phi_{j_N}(x_N). \tag{1.10}$$

Here j_m are the labels of the states, for example $j_1, \ldots j_{N_1} = 1, j_{N_1+1}, \ldots j_{N_1+N_2} = 2$, etc. The sum goes over all possible permutations of the labels and the prefactor takes care of the normalization. This wave function is symmetric on interchanging the labels.

For what follows, use the bra-ket notation for this state,

$$|N_1, N_2, \ldots\rangle = \Phi_{N_1,N_2,\ldots},$$

i.e., labeling the different many-particle states simply by the occupation numbers of the different states. For example, for a three-particle system the state with $N_1 = 2$ and $N_2 = 1$ is

$$|2,1\rangle = \frac{1}{\sqrt{3}}[\phi_1(x_1)\phi_1(x_2)\phi_2(x_3) + \phi_1(x_1)\phi_2(x_2)\phi_1(x_3) + \phi_2(x_1)\phi_1(x_2)\phi_1(x_3)].$$

Normalization then corresponds to integrating over the coordinates x_i.

Introduce now the operator a_i which decreases the occupation number in the state i by 1. Define it such that

$$\langle N_1, N_2, \ldots, N_i - 1, \ldots | a_i | N_1, N_2, \ldots, N_i, \ldots \rangle = \sqrt{N_i}. \tag{1.11}$$

The matrix element is naturally finite only for this type of combinations of states. The conjugate operator is defined such that

$$\langle N_1, N_2, \ldots, N_i, \ldots | a_i^\dagger | N_1, N_2, \ldots, N_i - 1, \ldots \rangle = \sqrt{N_i}. \tag{1.12}$$

Combining the two operators then yields the properties

$$a_i^\dagger a_i = N_i, \quad a_i a_i^\dagger = N_i + 1, \tag{1.13}$$

i.e., the combined operator is proportional to the unit operator. For different states $i \neq j$ we also get that any diagonal matrix element of $a_i a_j^\dagger$ (i.e., bracketed between the same states) has to vanish. Similarly also the diagonal matrix elements of $a_i a_j$ and $a_i^\dagger a_j^\dagger$ vanish. Therefore, we get

$$[a_i, a_j^\dagger] \equiv a_i a_j^\dagger - a_j^\dagger a_i = \delta_{ij} \tag{1.14a}$$

$$a_i a_j - a_j a_i = a_i^\dagger a_j^\dagger - a_j^\dagger a_i^\dagger = 0 \tag{1.14b}$$

This is the usual bosonic commutation relation.

1.2.2 Fermions

For fermions, the occupation number of each state can only be 0 or 1, and the total wave function has to be antisymmetric with respect to interchanging any two particles,

$$|N_1, N_2, \ldots\rangle = \frac{1}{\sqrt{N!}} \sum_P \delta_P \phi_{j_1}(x_1)\phi_{j_2}(x_2)\cdots\phi_{j_N}(x_N). \tag{1.15}$$

Here δ_P is +1 or -1 depending on whether the state $j_1, j_2, \ldots$ is obtained after even or odd number of transpositions from some initial configuration. For example, we have the three-particle state

$$\begin{aligned}|1,0,1,1\rangle =& \frac{1}{\sqrt{6}}[\phi_1(x_1)\phi_3(x_2)\phi_4(x_3) - \phi_1(x_1)\phi_4(x_2)\phi_3(x_3) - \phi_3(x_1)\phi_1(x_2)\phi_4(x_3) \\ & - \phi_4(x_1)\phi_3(x_2)\phi_1(x_3) + \phi_3(x_1)\phi_4(x_2)\phi_1(x_3) + \phi_4(x_1)\phi_1(x_2)\phi_3(x_3)].\end{aligned} \tag{1.16}$$

Now the creation and annihilation operators are defined through the matrix elements

$$\langle N_1, N_2, \ldots, N_i + 1, \ldots | a_i^\dagger | N_1, N_2, \ldots, N_i, \ldots \rangle = (-1)^{\sum_{k=1}^{i-1} N_k} \delta_{N_i,0} \quad (1.17)$$

and

$$\langle N_1, N_2, \ldots, N_i - 1, \ldots | a_i | N_1, N_2, \ldots, N_i, \ldots \rangle = (-1)^{\sum_{k=1}^{i-1} N_k} \delta_{N_i,1}. \quad (1.18)$$

These satisfy (again for diagonal matrix elements) (exercise)

$$a_i^\dagger a_i = N_i, \quad a_i a_i^\dagger = 1 - N_i \quad (1.19)$$

and

$$a_i^\dagger a_j + a_j a_i^\dagger = 0 \quad (1.20)$$

for $i \neq j$. Thus, we find the fermionic commutation relations

$$\{a_i, a_j^\dagger\} \equiv a_i a_j^\dagger + a_j^\dagger a_i = \delta_{ij} \quad (1.21a)$$

$$a_i a_j + a_j a_i = a_i^\dagger a_j^\dagger + a_j^\dagger a_i^\dagger = 0. \quad (1.21b)$$

During the course, we will essentially only need the commutation relations (Eqs. (1.14,1.21)) and the properties (1.13,1.19).

1.3 Heisenberg vs. Schrödinger pictures

In the following, many of the phenomena are treated in the Heisenberg picture of quantum mechanics, where the operators rather than the wave functions are time dependent. The two pictures are connected as follows. Let the Hamiltonian $\hat{H}$ be time independent. The wave function $\Psi(t)$ in the Schrödinger picture obeys the Schrödinger equation

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi. \quad (1.22)$$

Its solution can be formally written as

$$\Psi(t) = e^{-i\hat{H}t/\hbar} \Psi_H, \quad (1.23)$$

where Ψ_H is time independent. The time dependence of an observable O would be calculated as a function of the matrix elements of the corresponding operator $\hat{O}$,

$$O_{mn} = \langle \Psi_m^*(t) | \hat{O} | \Psi_n(t) \rangle = \langle \Psi_{Hm}^* | e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar} | \Psi_{Hn} \rangle. \quad (1.24)$$

This is thus the same as the matrix element of a time-dependent (Heisenberg) operator

$$\tilde{O}(t) \equiv e^{i\hat{H}t/\hbar} \hat{O} e^{-i\hat{H}t/\hbar}, \quad (1.25)$$

calculated over the time-independent basis Ψ_{Hm}. This operator satisfies the Heisenberg equation

$$\frac{\partial \tilde{O}}{\partial t} = \frac{i}{\hbar} [\hat{H}, \tilde{O}]. \quad (1.26)$$

Below, if not stated otherwise, any time-dependent operator refers to the Heisenberg operator.

1.4 Problems

1. The charging energy required to bring a single electron between two capacitor plates is $E_C = e^2/(2C)$. Let us assume a simple parallel-plate model for a capacitor, $C = \epsilon A/d$, where $\epsilon = \epsilon_r \epsilon_0$, typical $\epsilon_r = 10$ and $\epsilon_0 = 8.85$ pF/m, A is the area of the plates and d their separation. We can also set $d = 1$ nm, a typical thickness of the oxide layer ("tunnel contact") formed between two metals. Assuming a square plate, $A = w^2$, estimate a width w of the junction which would correspond to the charging energy E_C being equal to 1 K. Estimate also the corresponding capacitances. How should these scales change so that charging effects would be observable at room temperature, i.e., $E_C \approx 300K/k_B$? Remember that $e = 1.6\ 10^{-19}$ C and $k_B = 1.38\ 10^{-23}$ J/K.

2. Show that the resistance of a tunnel barrier is independent of temperature or voltage.

3. Show that the heat current through a tunnel barrier in the linear response regime obeys a Wiedemann-Franz law, i.e., $\dot{Q}/\Delta T \propto T/R_T$. Find also the prefactor of this expression. Hint: Assume a vanishing voltage and that the temperatures $T_{L/R}$ of the left/right reservoirs are $T_{L/R} = T \pm \Delta T/2$. Finally, take the linear order in ΔT.

4. Prove Eqs. (1.19) and (1.21) by using the definitions (1.17) and (1.18).

5. The Hamiltonian for a Harmonic oscillator is

$$H = \hbar\omega(\hat{a}^\dagger \hat{a} + \frac{1}{2}), \tag{1.27}$$

where $\omega = \sqrt{k/m}$ is the resonance frequency for a mass m and spring constant k characterizing the potential, and $\hat{a}$ is the bosonic annihilation operator. With these operators, the (Heisenberg) operators for position $\hat{x}$ and and momentum $\hat{p}$ are $\hat{x} = \sqrt{\hbar/(2m\omega)}(\hat{a} + \hat{a}^\dagger)$ and $\hat{p} = i\sqrt{\hbar m\omega/2}(\hat{a}^\dagger - \hat{a})$. From the Heisenberg equation of motion,

$$i\hbar \frac{d\hat{O}}{dt} = [\hat{O}, H], \tag{1.28}$$

valid for any operator $\hat{O}$ without external time dependence, derive the equations of motion for $\hat{x}$ and $\hat{p}$.

Chapter 2

Scattering approach to quantum transport

It can be said that the field of theoretical mesoscopic electron transport started when it was noticed that the current through low-dimensional objects can be directly related to their scattering properties. The original idea is due to Rolf Landauer and dates back to his paper in the year 1957.[1] This concept was refined by Marcus Büttiker,[2] who essentially provided the framework of the scattering formalism explained below. This theory is often called by their fathers the Landauer-Büttiker formalism. It is described in detail also for example in the book by S. Datta.[3]

2.1 Scattering region, leads and reservoirs

The scattering formalism describes a structure with M reservoirs (here assumed non-superconducting), connected to each other via the *scattering region* (see Fig. 2.1). Between the reservoirs and the scattering region, the theory assumes semi-infinite *leads*. These leads are in a way a mathematical construction: they simply allow one to define the scattering states, and they mimic the physical properties of the reservoirs. The quasiparticle reservoirs are characterised by their chemical potentials μ_i and temperatures T_i. It is assumed that the quasiparticles undergo inelastic scatterings inside the reservoirs such that the outgoing particles are uncorrelated and their energy distribution may be described by the Fermi function $f(E;\mu,T)$. Furthermore, it is assumed that the number of particles in the reservoirs is infinite, such that they can act as constant sources of particles and the chemical potential is not affected by a removal

[1] R. Landauer, IBM J. Res. Dev. **1**, 223 (1957).
[2] M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986).
[3] S. Datta, "Electronic Transport in Mesoscopic Systems", Cambridge University Press, 1995.

of a single particle. The leads are assumed semi-infinite and clean, so that there is no back-scattering from them.

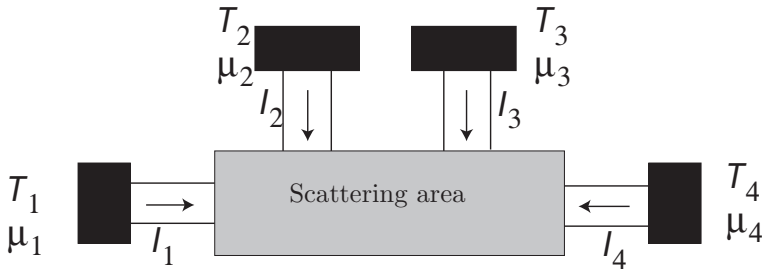


Figure 2.1: Scattering formalism. Quasiparticle reservoirs are connected to the scattering region via the ballistic (disorder-free) leads.

2.1.1 Transverse modes in semi-infinite leads

Now let us consider one of the semi-infinite leads. For simplicity, assume that it is a waveguide, a perfect rectangular 3d wire with a finite width and thickness (say, x and y directions), but with an infinite length (in the z-direction). The Schrödinger equation for the electrons in this wire in the nearly-free electron approach is

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + U(x,y,z)\right]\psi(x,y,z) = E\psi(x,y,z). \tag{2.1}$$

To make the back-scattering within the leads vanish, assume that the potential $U(x,y,z)$ is independent of the coordinate z. In this case, the equation separates in the transverse and longitudinal directions and the solution is of the form $\psi(x,y,z) = \chi(x,y)e^{ik_z z}$. The transverse functions obey the equation

$$\left[-\frac{\hbar^2}{2m}(\partial_x^2 + \partial_y^2) + U(x,y)\right]\chi_n(x,y) = \epsilon_n\chi_n(x,y). \tag{2.2}$$

After solving this, the wave number in the longitudinal direction is obtained from

$$k_z^n = \pm\sqrt{\frac{2m}{\hbar^2}(E-\epsilon_n)}. \tag{2.3}$$

The wave numbers k_z with finite imaginary parts correspond to evanescent waves which do not carry a current, and the real k_z describe currents carried in the two directions in the wire, depending on the sign of k_z. This can be seen by considering the group velocity of the state n in the direction of transport:

$$v_n = \frac{1}{\hbar}\frac{dE}{dk_z} = \frac{\hbar k_z}{m}, \tag{2.4}$$

which has the same sign as k_z. Below, we refer to the states with $k_z > 0$ as "incoming" states and to those with $k_z < 0$ as "outgoing" states.

For the transverse modes, one could envisage many different types of potentials describing a wire with finite width and thickness. Generally which model is considered does not have a big effect on the resulting physics. The easiest model is a wire with hard walls, i.e., $U(x, y) = 0$ inside the wire and infinity elsewhere. This yields the boundary conditions

$$\chi(0, y) = \chi(L_x, y) = \chi(x, 0) = \chi(x, L_y) = 0, \tag{2.5}$$

where L_x and L_y are the dimensions of the wire in the two directions. The resulting transverse mode wave functions are (see Fig. 2.2)

$$\chi_n(x, y) = A_n \sin(k_{n_x} x) \sin(k_{n_y} y), \tag{2.6}$$

where $k_{n_{x/y}} = n\pi/L_{x/y}$. The corresponding energies are

$$\epsilon_n = \frac{\hbar^2}{2m}(k_{n_x}^2 + k_{n_y}^2). \tag{2.7}$$

Typically the normalization of the state is chosen $A_n = \sqrt{v_n}$, such that each state carries a unit probability flux. Below, we refer to the transverse modes mostly only by their index n, referring to one particular combination of $\{n_x, n_y\}$.

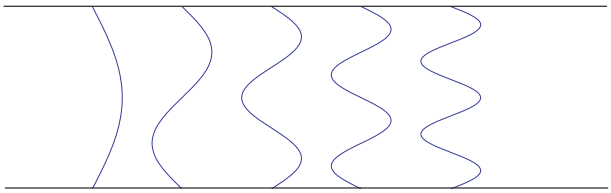


Figure 2.2: Wave functions for the lowest-energy transverse modes in a two-dimensional wire (i.e., $L_y \to 0$) with hard-wall potentials.

Due to the assumption of an infinite length of the wire, in the z-direction the different modes have different dispersion relations, (see Fig. 2.3)

$$\epsilon_{k_z^n} = \epsilon_n + \frac{\hbar^2 k_z^2}{2m}. \tag{2.8}$$

This means that a particular transverse state n can contribute to transport only if the transverse energy ϵ_n is below the Fermi energy (or more precisely, the chemical potential) of the reservoir that feeds the particles into the states. We can thus define a function $M(E)$ that counts the number of modes with ϵ_n below some energy E:

$$M(E) \equiv \sum_n \theta(E - \epsilon_n), \tag{2.9}$$

where $\theta(x)$ is the Heaviside step function.

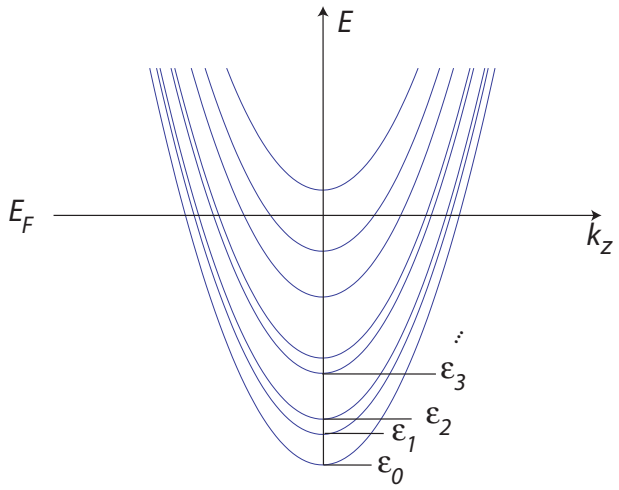


Figure 2.3: Dispersion relation for the longitudinal k-vector in the different transverse modes. Only the modes with a cutoff energy $\epsilon_n < E_F$ can contribute to electric current.

Quasi-continuous k_z-values

In what follows, we aim to sum the contributions to the current from all the occupied k_z-states. In the case of an infinite wire, the spectrum is continuous and we need to find a way to convert the k_z-sums into integrals over k_z. Therefore, consider a finite 1-dimensional wire of length L, and assume periodic boundary conditions, $\psi_z(z) = \psi_z(z + L)$. The Schrödinger equation is now solved with $\psi_z(z) = e^{ik_z z}$, $k_z = 2n\pi/L$. Therefore, the level spacing is $2\pi/L$, and hence any sum over the k_z-states below can be converted with the prescription

$$\sum_k f(k) \to 2 \text{ (for spin)} \times \frac{L}{2\pi} \int dk f(k). \tag{2.10}$$

This holds provided the function $f(k)$ is smooth on the scale of the level spacing.

2.1.2 Current carried by a transverse mode

Consider now a single incoming transverse mode, which is occupied with some probability $f_L(E_k)$, $f_L(E_k)$ being the energy distribution function inside the reservoir feeding the lead. A system with n electrons per unit length carries a charge current env, where v is the velocity of the electrons. The electron density associated with a single electron in one k_z-state in a conductor of length L is $1/L$, and thus the incoming current in the wire (towards the scattering region) is

$$I^n_{\text{in}} = \frac{e}{L} \sum_{k^n_z} v_{k^n_z} f_L(E_{k^n_z}) = \frac{2e}{2\pi} \int dk \frac{1}{\hbar} \frac{dE}{dk} f_L(E_k). \tag{2.11}$$

Here we used the above scheme of converting sums over k into integrals over them. It is convenient to convert this integral into an integral over energy, defining the density of states for the 1-d case, $N(E) \equiv \frac{dk}{dE} = 1/(dE/dk)$. We find that the velocity and density of states cancel, yielding

$$I_{\rm in}^n = \frac{2e}{h}\int_{\epsilon_n}^{\infty} dE f_L(E). \tag{2.12}$$

The integral is carried out from the mode cutoff ϵ_n, below which this mode does not carry a current. Now summing over the modes yields the total incoming current,

$$I_{\rm in} = \frac{2e}{h}\int_{-\infty}^{\infty} dE f_L(E) M(E), \tag{2.13}$$

where $M(E)$ is the function counting the number of propagating modes.

2.1.3 Wire between two reservoirs

But there is also a current carried in the opposite direction. First, assume that each mode is backscattered with the same probability $1 - T(E)$. Thus, the outgoing current is a sum of back-scattered incoming particles and the current from the other reservoirs, transmitted with the probability $T(E)$. In the case of two reservoirs, we hence get

$$I_{\rm out} = \frac{2e}{h}\int_{-\infty}^{\infty} dE((1-T(E))f_L + T(E)f_R(E)M(E), \tag{2.14}$$

where we assumed for simplicity that the two leads have the same function $M(E)$. The two distribution functions $f_L(E)$ and $f_R(E)$ can easily be identified as the distribution functions of the two reservoirs. The total current is thus

$$I = I_{\rm in} - I_{\rm out} = \frac{2e}{h}\int dE T(E)(f_L(E) - f_R(E))M(E) \stackrel{T=0}{=} \frac{2e}{h}\int_{\mu_R}^{\mu_L} T(E)M(E). \tag{2.15}$$

The last equality is obtained in the limit of a vanishing temperature (not vanishing transmission!). Consider now the case where $M(E)$ and $T(E)$ are constant for energies between μ_R and $\mu_L = \mu_R + eV$, or within the window of a few k_BT around $\mu_{L/R}$. Then the resulting current is[4]

$$I = \frac{2e^2}{h} MTV. \tag{2.16}$$

In the case of a ballistic wire, $T = 1$. In this case each mode adds a contribution $G_0 = 2e^2/h = (12.9\text{k}\Omega)^{-1}$ to the total conductance of this ballistic wire.

But in the case $T = 1$ there was no regard of the actual length of the wire! This is due to the lack of backscattering inside the wire. Such a property indeed leads to an infinite conductivity (conductance per unit length), but not to an

[4]This result is independent of temperature, as long as the above assumption is satisfied.

infinite conductance! The finite resistance is due to the contacts of the larger reservoirs to the ballistic wires: also the dissipation takes place at these contacts.

Let us estimate the number of modes for some typical structures. This in principle depends on the assumption of the potential $U(x, y)$ describing the transverse modes, but we get a fair estimate already from the analysis with infinite walls. At a vanishing temperature, all the states up to the Fermi energy in principle contribute to the current. Thus, the requirement is $\epsilon_n < E_F$ or

$$\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} < \frac{k_F^2}{\pi^2} = \frac{4}{\lambda_F^2}, \tag{2.17}$$

where k_F is the Fermi wave vector and λ_F is the Fermi wave length. We can thus see that if the structure is fabricated on a 2-dimensional electron gas, and thus one dimension is absent (say, y), there are $M = \mathrm{int}(2L_x/\lambda_F)$ propagating modes. Here $\mathrm{int}(x)$ denotes the closest integer smaller than x. In a square wire with $L_x = L_y$, the number of propagating modes is $\frac{1}{2}\mathrm{int}(2L_x/\lambda_F)(\mathrm{int}(2L_x/\lambda_F) - 1)$. In a metal, the typical λ_F is of the order of a few angstroms, and thus a 100 nm $\times$ 100 nm contact has some hundred thousand modes, indicating a contact resistance of the order of 100 mΩ. In semiconductors, however, λ_F can be a few tens of nanometers, and thus the number of propagating modes can be reduced to order one. Moreover, with gates it is typically possible to control the local potential profile of the 2DEG. With such control, L_x and therefore the number of channels can be controlled.

The conductance quantization was first observed in 1988[5] in a 2DEG formed between GaAs and AlGaAs and patterned with gates (see Fig. 2.4). Recently it has been a challenge to see a similar conductance quantization in carbon nanotubes.[6]

2.2 Scattering matrix

In most cases the assumption of mode-independent transmission is not very realistic. For such a case, it is useful to define a *scattering matrix* $s_{nm}^{\alpha\beta}$ that connects the amplitude of an incoming transverse mode m in lead β to the outgoing mode n in lead α. For an N-probe system with M_α propagating modes, this scattering matrix is thus $M_T \times M_T \equiv \sum_{\alpha=1}^{N} M_\alpha \times \sum_{\alpha=1}^{N} M_\alpha$ dimensional.

To be specific, consider a two-probe system with $N_1 = 2$ and $N_2 = 1$. In this case, denote the amplitudes of the incoming modes by $a_{\alpha,n}$ and of the outgoing modes by $b_{\alpha,n}$. Then we have

$$\begin{pmatrix} b_{11} \\ b_{12} \\ b_{21} \end{pmatrix} = \begin{pmatrix} s_{11}^{11} & s_{12}^{11} & s_{11}^{12} \\ s_{21}^{11} & s_{22}^{11} & s_{21}^{12} \\ s_{11}^{21} & s_{12}^{21} & s_{11}^{22} \end{pmatrix} \begin{pmatrix} a_{11} \\ a_{12} \\ a_{21} \end{pmatrix}. \tag{2.18}$$

[5] B. J. van Wees, *et al.*, Phys. Rev. Lett. **60**, 848 (1988), and D. Wharam, *et al.*, J. Phys. C, **21**, L209 (1988).

[6] See for example M. J. Biercuk, *et al.*, Phys. Rev. Lett. **94**, 026801 (2005).

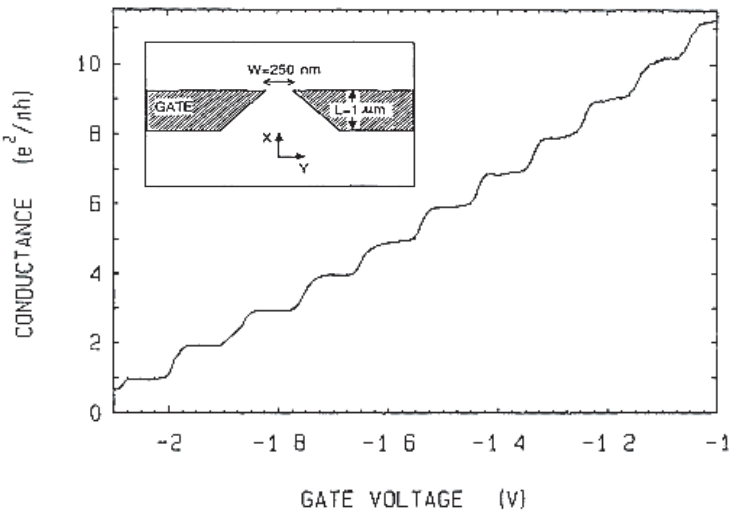


Figure 2.4: Quantized conductance of a quantum point contact, as a function of the gate voltage that deforms the potential in the contact. Revised from B. J. van Wees, *et al.*, Phys. Rev. Lett. **60**, 848 (1988).

In most cases, of course, the dimension of the scattering matrix is huge and it is better to refer to it by simply $s_{nm}^{\alpha\beta}$. Below, the lead indices are sometimes dropped for simplicity. In all cases, the first index refers to the transverse mode n in lead α and the second to the transverse mode m in lead β.

The scattering matrix can of course be calculated if one knows the potential profile $U(x, y, z)$ and the vector potential $\mathbf{A}$ within the total scattering region. An example for such a calculation is indicated below. However, for now it is enough to know that the scattering matrix for a given system exists.

2.2.1 Some properties of the scattering matrix

Transmission and reflection probabilities

From the scattering matrix, it is easy to find the transmission and reflection probabilities T_{nm} and R_{nm} between modes n and m. They are given by

$$\begin{aligned} T_{nm} &\equiv |s_{nm}^{\alpha\beta}|^2, \quad (\alpha \neq \beta) \\ R_{nm} &\equiv |s_{nm}^{\alpha\alpha}|^2. \end{aligned} \tag{2.19}$$

As T_{nm} and R_{nm} are probabilities, they must lie between 0 and 1.

Unitarity

Another property of the scattering matrices can be shown from the requirement that the total amplitude for outgoing probability current is the same as that

for the incoming probability current. For this, let us write the amplitudes of the incoming and outgoing states in the vector notation, $\{a\}$ and $\{b\} = S\{a\}$, where S is the total scattering matrix. The total amplitude for the probability currents hence is

$$\sum_\beta \sum_{m\in\beta} |a_{\beta m}|^2 = \{a\}^\dagger\{a\} = \{b\}^\dagger\{b\} = \{a\}^\dagger S^\dagger S\{a\}. \tag{2.20}$$

Hence we get

$$S^\dagger S = SS^\dagger = I, \tag{2.21}$$

a unit matrix. Hence, any proper scattering matrix has to be unitary. Writing in terms of the scattering states, this implies

$$\sum_n |s_{nm}|^2 = 1 = \sum_m |s_{nm}|^2. \tag{2.22}$$

The first equality implies that the total probability for scattering from a given incoming state to one of the outgoing states is unity. The latter is the reverse: it tells that for any outgoing state, the total probability that it scattered from one of the incoming states is one.

Time reversal symmetry

Assume we obtained the scattering matrix S from the solution of the Schrödinger equation

$$\left[\frac{(i\hbar\nabla + e\mathbf{A})^2}{2m} + U(\vec{r})\right]\psi(\vec{r}) = E\psi(\vec{r}). \tag{2.23}$$

describing our setup. From this we obtained $\{b\} = S\{a\}$. Now take the complex conjugate of Eq. (2.23) and invert the magnetic field, $\mathbf{A} \to -\mathbf{A}$. We thus get

$$\left[\frac{(i\hbar\nabla + e\mathbf{A})^2}{2m} + U(\vec{r})\right]\psi^*(\vec{r}) = E\psi^*(\vec{r}). \tag{2.24}$$

As the differential operators for the two cases are the same, we find that $[\psi(\vec{r})^*]_{-B} = [\psi(\vec{r})]_B$. But taking the complex conjugate makes an outgoing wave an incoming one and vice versa. Thus, we have

$$\{b\} = [S]_B\{a\}, \text{ and } \{a\} = [S^*]_{-B}\{b\}, \tag{2.25}$$

or in other words,

$$[S^{-1}]_B = [S^*]_{-B}. \tag{2.26}$$

Combining this with unitarity, $S^{-1} = S^\dagger$ yields

$$[S]_B = [S^T]_{-B}, \tag{2.27}$$

where the superscript T indicates a transpose. For the transmission probabilities, this implies

$$[T_{nm}]_B = [T_{mn}]_{-B}. \tag{2.28}$$

Thus, in the absence of a magnetic field, we have to have $T_{nm} = T_{mn}$. Such symmetry is called a *time reversal symmetry*: for the system it is not relevant whether we calculate particles from modes m travelling to modes n in the forward or backward time direction.

This discussion assumed that the electrostatic potential $U(\vec{r})$ stays constant upon reversing the magnetic field. Such an assumption breaks down at a large bias voltage in the case when we have Landau levels in a quantum Hall effect, but for low field or voltages the symmetry between the transmission probabilities is perfectly valid.

Two-probe scattering matrix

If we only have two reservoirs in the system, the scattering matrix has the form

$$S = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix}. \tag{2.29}$$

Here r and t are reflection and transmission matrices for the left lead and r' and t' the same for the right lead. The reflection matrices are square matrices with dimensions $N_1 \times N_1$ and $N_2 \times N_2$ whereas the transmission matrices have the dimensions $N_1 \times N_2$ and $N_2 \times N_1$.

2.2.2 Combining scattering matrices

Sometimes it is convenient to separate the scattering problem into two or more parts containing scatterers with known scattering matrices. Then the question is how the total scattering matrix is related to the individual ones. Consider the example shown in Fig. 2.5. The left part of the system satisfies

$$\begin{pmatrix} b_L \\ b_{iL} \end{pmatrix} = \begin{pmatrix} r_L & t_L \\ t'_L & r'_L \end{pmatrix} \begin{pmatrix} a_L \\ a_{iL} \end{pmatrix} \tag{2.30}$$

and analogously for the right part after replacing the index L by R.

It is evident that the total scattering matrix is not simply a product of the two scattering matrices, but one has to rather employ a *transfer matrix*, connecting the states from the left lead to those in the island. Now assume the amplitude of the incoming state in the left part of the "island" equals the amplitude of the outgoing state in the right part, i.e., we have $a_{iL} = b_{iR}$ and similarly for the incoming state of the right part, $a_{iR} = b_{iL}$. We may now eliminate a_{iR} and b_{iL} and find the total scattering matrix

$$\begin{pmatrix} b_L \\ b_R \end{pmatrix} = \begin{pmatrix} r & t \\ t' & r' \end{pmatrix} \begin{pmatrix} a_L \\ a_R \end{pmatrix} \tag{2.31}$$

where

$$\begin{aligned} t &= t_R[I - r'_L r_R]^{-1} t_L, \quad r = r_L + t'_L r_R [I - r'_L r_R]^{-1} t_L \\ t' &= t'_L[I - r_R r'_L]^{-1} t_R, \quad r' = r'_R + t_R[I - r'_L r'_R]^{-1} r'_L t'_R. \end{aligned} \tag{2.32}$$

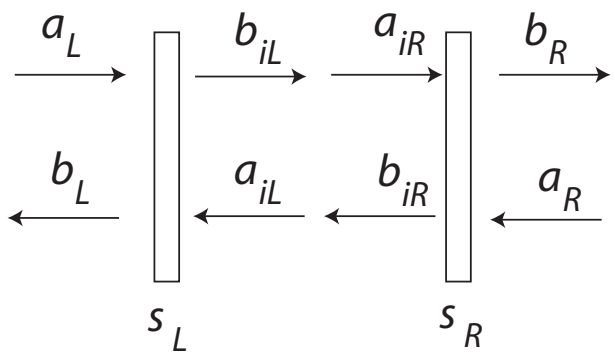


Figure 2.5: System with scatterers in series.

This result can be interpreted as "Feynman paths" the electron takes within the structure by expanding the matrix inverses in geometric series. For example the transmission amplitude is

$$t = t_R t_L + t_L(r'_L r_R)t_L + t_R(r'_L r_R)(r'_L r_R)t_L + \cdots . \tag{2.33}$$

Hence, a transmission of a particle through this double-barrier structure can be seen as a sum of transmitting directly through both scatterers, transmitting from the left, reflecting from both scatterers and then transmitting to the right, and so on.

2.3 Current from scattering

For the calculation of the average current for a general scattering matrix, we could simply include the mode-dependent transmission in Eqs. (2.14,2.15) and obtain for the two-probe conductance in the case of energy-independent scattering

$$G = \frac{2e^2}{h} \sum_{n,m} T_{nm}. \tag{2.34}$$

The generalization of this formula to a many-probe case is straightforward and is done below with the more general formalism.

However, such a treatment would not be easy to generalize for the calculation of the current fluctuations, discussed in Ch. 8. Therefore, we calculate the current by introducing the creation and annihilation operators $\hat{a}^\dagger_{\alpha n}$ and $\hat{a}_{\alpha n}$ for the incoming state n in lead α and similar operators $\hat{b}^\dagger_{\alpha n}$ and $\hat{b}_{\alpha n}$ for the

outgoing state. These operators are related via the scattering matrix,

$$\hat{b}_{\alpha n}(E) = \sum_{\beta} \sum_{m\in\beta} s^{\alpha\beta}_{nm}(E+i\eta)\hat{a}_{\beta m}(E) \tag{2.35a}$$

$$\hat{b}^{\dagger}_{\alpha n}(E) = \sum_{\beta} \sum_{m\in\beta} (s^{\alpha\beta}_{nm})^{*}(E-i\eta)\hat{a}^{\dagger}_{\beta m}(E). \tag{2.35b}$$

In this relation, the operators $\hat{b}(E)$ and $\hat{a}(E)$ are Fourier transforms of the corresponding Heisenberg operators $\hat{b}(t)$ and $\hat{a}(t)$. The fact they are evaluated at the same energy is a consequence of the assumption of only elastic scattering inside the scattering region. The infinitesimal imaginary part $i\eta$ ($\eta > 0$) reflects the causality of this relation (i.e., incoming state must precede the outgoing one), but this fact is not essential for the topics of this course and thus η is dropped from the subsequent equations.

The current operator in lead α at time t is expressed via these operators as[7]

$$\hat{I}_{\alpha} = \frac{2e}{h} \sum_{n\in\alpha} \int dE dE' e^{i(E-E')t/\hbar} \left[\hat{a}^{\dagger}_{\alpha n}(E)\hat{a}_{\alpha n}(E') - \hat{b}^{\dagger}_{\alpha n}(E)\hat{b}_{\alpha n}(E') \right], \tag{2.36}$$

where the notation $n \in \alpha$ refers to the channels n in lead α. This equation can be written in terms of only operators $\hat{a}$ and $\hat{a}^{\dagger}$ by defining a matrix

$$A^{\beta\gamma}_{mn}(\alpha, E, E') = \delta_{mn}\delta_{\beta\alpha}\delta_{\gamma\alpha} - \sum_{j} (s^{\alpha\beta}_{jm})^{*}(E) s^{\alpha\gamma}_{jn}. \tag{2.37}$$

This matrix satisfies

$$\mathrm{Tr}[A^{\beta\beta}] \equiv \sum_{n} A^{\beta\beta}_{nn} = -\mathrm{Tr}[(s^{\alpha\beta})^{\dagger} s^{\alpha\beta}] = -\bar{T}_{\alpha\beta}, \quad \alpha \neq \beta \tag{2.38a}$$

$$\mathrm{Tr}[A^{\alpha\alpha}] = M_{\alpha} - \bar{R}_{\alpha}, \tag{2.38b}$$

where $\bar{T}_{\alpha\beta}$ is the total transmission probability from β to α (summed over the modes), and $\bar{R}_{\alpha} = \mathrm{Tr}[(s^{\alpha\alpha})^{\dagger} s^{\alpha\alpha}]$ is the total reflection probability for electrons in lead α. The current operator can thus be expressed by

$$\hat{I}_{\alpha} = \frac{2e}{h} \sum_{\beta\gamma} \sum_{m\in\beta, n\in\gamma} \int dE dE' e^{i(E-E')t/\hbar} \hat{a}^{\dagger}_{\beta m}(E) A^{\beta\gamma}_{mn}(\alpha, E, E') \hat{a}_{\gamma n}(E'). \tag{2.39}$$

In this chapter, we only concentrate on the average current $\langle \hat{I}_{\alpha} \rangle \equiv \mathrm{Tr}[\rho \hat{I}]$, where ρ is the density matrix for the system. This means that the average contains both the statistical and the quantum-mechanical averaging. For this, we only need to know the statistical average of the number operator,

$$\langle \hat{a}^{\dagger}_{\alpha n}(E)\hat{a}_{\beta m}(E') \rangle = \delta_{mn}\delta_{\alpha\beta}\delta(E-E') f_{\alpha}(E). \tag{2.40}$$

[7] See Y. Blanter and M. Büttiker, Phys. Rep. **336**, 1 (2000) or M. Büttiker, Phys. Rev. B **46**, 12 485 (1992) for the derivation of this formula from the field operators for the lead states.

In fact, this defines the distribution function in reservoir α. Thus, the average current is

$$\langle \hat{I}_\alpha \rangle = \frac{2e}{h} \sum_\beta \int dE \mathrm{Tr}[A^{\beta\beta}] f_\beta(E) = \frac{2e}{h} \int dE [(M_\alpha - \bar{R}_\alpha) f_\alpha(E) - \sum_{\beta \neq \alpha} \bar{T}_{\alpha\beta} f_\beta(E)]$$
$$\approx \frac{2e}{h} (M_\alpha - \bar{R}_\alpha) \mu_\alpha - \sum_{\beta \neq \alpha} \bar{T}_{\alpha\beta} \mu_\beta, \tag{2.41}$$

where the latter approximate equality is valid in the linear response limit where the total scattering probabilities are independent of energy. This is the most general form for the average current. For a two-probe system we have $M_1 - \bar{R}_1 = \bar{T}_{12}$ and thus we obtain

$$\langle \hat{I}_\alpha \rangle = \frac{2e}{h} \int dE \bar{T}_{12} (f_\alpha - f_\beta) \approx \frac{2e^2}{h} \bar{T}_{12} V. \tag{2.42}$$

The latter equality is valid in the linear response regime, i.e., when the total transmission probability is independent of energy within the energy region of width $\max(k_B T, eV)$ around E_F.

2.4 Examples

2.4.1 Voltage probe

A voltage probe is a terminal into which the average current vanishes. Consider a three-terminal system, where the probe 3 is a voltage probe. In the linear response regime, the current I_3 is

$$I_3 = \frac{2e}{h} [(M_3 - \bar{R}_3) \mu_3 - \bar{T}_{31} \mu_1 + \bar{T}_{32} \mu_2]. \tag{2.43}$$

Using unitarity of the total scattering matrix, we have $M_3 - \bar{R}_3 = \bar{T}_{31} + \bar{T}_{32}$. Now requiring that $I_3 = 0$ we get the potential in the probe,

$$\mu_3 = \frac{\bar{T}_{31} \mu_1 + \bar{T}_{32} \mu_2}{\bar{T}_{31} + \bar{T}_{32}}. \tag{2.44}$$

This is a result that could have also obtained from elementary circuit analysis with conductances proportional to $\bar{T}_{\alpha\beta}$. However, this result applies also to ballistic systems!

2.4.2 Two-probe vs. four-probe setup

Originally, Landauer did in fact not get the result of Eq. (2.34). Consider the direction-averaged energy distribution function of the particles inside the lead in the single-channel two-probe case. It is given by

$$f(E) = \frac{1}{2} [(1 + R) f_L(E) + T f_R(E)], \tag{2.45}$$

where $R = 1 - T$ is the reflection probability and T is the transmission probability from the right reservoir. Thus, the distribution of particles follows the distribution in the reservoirs. Now, the chemical potential of the lead can be defined through

$$\mu = \int dE f(E). \tag{2.46}$$

This yields for the left lead $2\mu_{LL} = (1 + R)\mu_L + T\mu_R$ and for the right lead $2\mu_{RL} = (1 + R)\mu_R + T\mu_L$. Hence we get

$$\mu_{LL} - \mu_{RR} = \frac{1}{2}(1 + R - T)(\mu_L - \mu_R) = R(\mu_L - \mu_R). \tag{2.47}$$

Hence, the current divided by the potential difference between the leads is

$$G_{4p} = \frac{2e^2}{h} T \frac{\mu_L - \mu_R}{\mu_{LL} - \mu_{RR}} = \frac{2e^2}{h} \frac{T}{R}. \tag{2.48}$$

This is the original Landauer formula. It has the property that as the reflection probability $R \to 0$, the conductance becomes infinite. Now, the two-probe formula can be obtained by summing the resistances of the contacts and the four-probe resistance.

2.4.3 Resonant tunneling

In the exercises, you will calculate the conductance through a system defined by two scatterers, but where the system has also its own dynamics. This calculation is in fact a model for a noninteracting quantum dot with its own discrete energy levels. Assuming only a single level with energy ϵ_1 would be relevant for transmission, the resulting linear conductance is of the Breit-Wigner form,[8]

$$G = \frac{2e^2}{h} \frac{\Gamma_L \Gamma_R}{(\epsilon_1 - E_F)^2 + (\Gamma_L + \Gamma_R)^2/4}, \tag{2.49}$$

where Γ_L and Γ_R are energy scales characterizing the couplings to the reservoirs. Now, assume the coupling is weak, i.e., Γ_L and Γ_R are small. Then the conductance is generally small, roughly $\Gamma_L \Gamma_R / \epsilon^2$ times the conductance quantum. But if the system is brought into resonance with the dot, $\epsilon_1 \to E_F$, the conductance becomes $4\Gamma_L \Gamma_R / (\Gamma_L + \Gamma_R)^2 2e^2/h$, which for a symmetric system becomes $2e^2/h$, i.e., a relatively large conductance. This is an example of a resonance.

In the system with two scatterers, the general result for the transmission is

$$T_{12} = \frac{T_1 T_2}{1 + R_1 R_2 - 2\sqrt{R_1 R_2}\cos(2\theta)}, \tag{2.50}$$

where $T_i = 1 - R_i$ $(i = 1, 2)$ are the transmission probabilities through the individual scatterers and $\theta = \sqrt{2mE}d/\hbar$. Now one may obtain a generalized

[8] G. Breit and E. Wigner, Phys. Rev. **49**, 519 (1936).

Breit-Wigner equation by defining "quasi-eigenstates" with energies E_n from the condition $\theta = n\pi$, where T_{12} has a maximum. Expanding $\cos(2\theta)$ around one maximum yields

$$\begin{aligned}\cos(2\theta(E)) &= \cos(2(\theta(E) - \theta(E_n))) \approx 1 - 2(\theta(E) - \theta(E_n))^2 \\ &\approx 1 - 2[\theta'(E)(E - E_n)]^2 = 1 - \frac{md^2}{\hbar^2 E_n}(E - E_n)^2.\end{aligned} \tag{2.51}$$

Substituting this to Eq. (2.50) yields a Breit-Wigner-type transmission formula with $\Gamma_i = T_i \hbar \frac{v_n}{2d}$. The factor $\frac{v_n}{2d}$ can be seen as an "attempt frequency", i.e., the frequency with which the particle hits the scatterers once it is inside the island.

2.4.4 Diffusive wire and Drude formula

In the previous chapter, we argued that the resistance of a disordered (diffusive) wire follows the Einstein relation, Eq. (1.3), or the equivalent Drude formula, Eq. (1.5). This can be roughly argued from the Landauer-Büttiker formula, Eq. (2.34), by assuming that the average transmission is

$$\langle T \rangle \sim \frac{\ell_{\rm el}}{L}. \tag{2.52}$$

Combining this with the fact that the number of channels in the wire, $M \sim A^2/(2\lambda_F^2)$, we obtain (roughly) the Einstein relation. Now, this has been also proven rigorously[9] with the *random matrix theory*. This theory does not only yield the average transmission probability, but also their whole probability distribution. For a diffusive wire, this is given by

$$\rho(T) = \frac{\pi\hbar G}{e^2} \frac{1}{T\sqrt{1-T}}, \tag{2.53}$$

where G is the conductance of the wire. This distribution is useful for the description of current fluctuations.

2.5 Calculating the scattering matrix (**)

Above and in the exercises, there are examples of calculating the scattering properties of some model systems. But also the general problem for an arbitrary system can be calculated, at least numerically. The numerical methods typically rely on Green's-function methods discussed for example in S. Datta's book, but let us just give a brief idea of them.

The (retarded) Green's function for the Schrödinger equation describing the whole scattering system obeys the equation

$$(E + i\eta - H(x))G^R(x, x') = \delta(x - x'). \tag{2.54}$$

[9] See, for example, Yu. Nazarov, Phys. Rev. Lett. **73**, 134 (1994), or C. W. J. Beenakker, Rev. Mod. Phys. **69**, 731 (1997).

Here $H(x)$ is the Hamiltonian specified in the real-space coordinates. Now any numerical method would discretize the spatial coordinates, and Eq. (2.54) would be a matrix equation with a formal solution

$$G^R = [(E + i\eta)I - H]^{-1}, \tag{2.55}$$

where I is a unit matrix. But H is an infinite-dimensional matrix, and therefore the problem is not yet computationally well defined. The solution to this is to divide the matrix G into parts,

$$G^R = \begin{pmatrix} G_l & G_{lS} \\ G_{Sl} & G_S \end{pmatrix} \equiv \begin{pmatrix} (E + i\eta)I - H_l & \tau_{lS} \\ \tau_{Sl} & (E + i\eta)I - H_S \end{pmatrix}^{-1}. \tag{2.56}$$

Here $g_l^{-1} = (E + i\eta)I - H_l$ represents the isolated leads with Green's functions g_l^{-1}, τ_{lS} and τ_{Sl} are the connections between the leads and the scattering region, and $g_S^{-1} = (E + i\eta)I - H_S$ describes an isolated scattering region. Now, the problem is reduced to a part with an infinite matrix, but a well-known system (clean, semi-infinite lead), and to a non-trivial but finite-dimensional scattering region. The previous can be calculated exactly[10], whereas the latter can be calculated using numerical methods. The coupling matrix to the leads is also typically finite-dimensional as only the ends of the leads need to be coupled to the ends of the scattering region.

Once the total Green's function is obtained, the scattering matrix can be calculated from[11]

$$s_{nm}^{\alpha\beta} = -\delta_{nm}^{\alpha\beta} + i\hbar\sqrt{v_n^\alpha v_m^\beta}\langle\chi_{n,\alpha}|G^R|\chi_{m,\beta}\rangle. \tag{2.57}$$

With this method, one can calculate the transmission through any given transmission region. This type of a method has for example been used in the study of the critical exponents in the Anderson localization transition[12] discussed briefly in the next chapter, but it has been applied to model essentially to all types of mesoscopic systems.

2.6 Problems

1. Prove the sum rules for the total transmission probabilities

$$\sum_\alpha \bar{T}_{\alpha\beta} = M_\beta \tag{2.58a}$$

$$\sum_\beta \bar{T}_{\alpha\beta} = M_\alpha. \tag{2.58b}$$

Using these, show that the total current into the scattering region vanishes.

[10] See S. Datta's book for details.

[11] For derivation, see Datta's book or D. S. Fisher and P. A. Lee, Phys. Rev. B **23**, 6851 (1981).

[12] See B. Kramer and A. MacKinnon, Rep. Prog. Phys. **56**, 1469 (1993).

2. Show that for a two-probe system $\bar{T}_{12}$=$\bar{T}_{21}$ even in the absence of time-reversal symmetry.

3. **Resonant tunneling**: assume we have two scatterers connected in series in a single-mode setup. Denote the separation of the scatterers by d. Assume that between the scatterers, the wave amplitudes acquire a dynamical energy-dependent phase $\phi = kd = \sqrt{2mE}d/\hbar$, i.e., $a_{iR} = e^{i\phi}b_{iL}$ and $a_{iL} = e^{i\phi}b_{iR}$. For simplicity, assume that the two scatterers have the same scattering matrix with $t = t' = \sqrt{T}$ and $r = -r' = \sqrt{R}$, where $R = 1 - T$ is the reflection probability.

   Calculate the total transmission probability through this system. Plot this result as a function of ϕ for $T = 0.01$.

4. For a multi-probe system with energy independent scattering, prove the Onsager reciprocity relation

$$R_{\alpha\beta,\gamma\delta}(B) = R_{\gamma\delta,\alpha\beta}(-B), \tag{2.59}$$

   where $R_{\alpha\beta,\gamma\delta} \equiv ed(\mu_\gamma - \mu_\delta)/dI_\alpha$, and current flows only between terminals α and β. Hint: write the resistances in terms of the conductances and use their symmetry.

Chapter 3

Interference effects

One of the important quantum-mechanical effects present in nanoelectronic systems are the interference effects. An example of such an effect was shown in the previous chapter, in the case when two scatterers were connected to each other and the modes obtained a dynamical phase between the scatterers. The phase coherence of the waves lead in this case to the resonant tunneling effect. The overall transmission probabilities of this system was not just a sum of transmission probabilities of independent paths that the particle may take, but a squared sum of transmission amplitudes.

Apart from resonant tunneling, there are three main types of interference effects present in nanoelectronic systems: Aharonov-Bohm effect that arises due to a magnetic field in structures containing a ring, (Anderson) localization due to the enhanced backscattering in a disordered wire, and universal conductance fluctuations in disordered wires. All of these are very dependent on phase coherence. In real-world systems, there are many factors that break this phase coherence, essentially as a result of inelastic scattering: depending on the temperature, the major contributors are electron-electron (important at low temperatures) and electron-phonon (important at high temperatures) scattering. Phase-breaking or *dephasing* is typically described with a dephasing time τ_φ or a dephasing length ℓ_φ, such that conductors much shorter than ℓ_φ exhibit the strongest interference effects, and conductors much longer than ℓ_φ behave in a "classical" way, i.e., exhibiting no interference.

3.1 Aharonov-Bohm effect

Consider a structure shown in Fig. 3.1. Experimentally[1] the conductance through such a device has been found to oscillate as a function of the magnetic field B (or the flux Φ) through the ring with a period corresponding to a flux equal to "flux quantum" $\Phi_0 \equiv h/e$ through the ring.

[1] S. Washburn and R. A. Webb, Adv. Phys. **35**, 375 (1986).

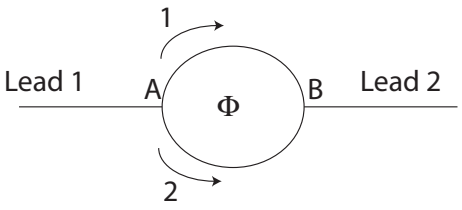


Figure 3.1: Aharonov-Bohm ring.

Now consider a similar "Feynman path" idea as was found for the resonant tunneling effect. Now the transmission amplitude t between given modes in the left and right leads is separated into two parts, $t = t_1 + t_2$, where t_1 describes transmission via the "upper" path and t_2 via the lower path (see Fig. 3.1). The total transmission probability for this mode is

$$T_{AB} = |t_1 + t_2|^2 = T_1 + T_2 + 2\mathrm{Re}(t_1 t_2) = T_1 + T_2 + 2\sqrt{T_1 T_2}\cos(\phi), \qquad (3.1)$$

where $T_i = |t_i|^2$ and ϕ is the phase difference between t_1 and t_2. The oscillations in the conductance occur due to the last, "interference" term in this equation. In the classical case, only the probabilities for the two paths would be added, i.e.,

$$T_{AB}^{\mathrm{cl}} = T_1 + T_2. \qquad (3.2)$$

This would be the result from a classical circuit theory, i.e., the total conductance of two conductors in parallel is the sum conductance.

The effect of a magnetic field can be explained in terms of the vector potential $\mathbf{A}$: The dynamical phase obtained within the transmission through the upper arm is

$$\phi_1 = \frac{1}{\hbar}\int_{\text{upper arm}} \vec{p}\cdot\vec{dl} \approx \frac{\sqrt{2mE}\pi r}{\hbar} - \frac{e}{\hbar}\int_{\text{upper arm}} \mathbf{A}\cdot\vec{dl}. \qquad (3.3)$$

Here r is the radius of the ring. The quadratic term in $\mathbf{A}$ was neglected as it would yield the same phase for both paths (it could change the dynamical phase, though). Thus, the transmission amplitude for the upper arm is $t_1 = t_1^0 \exp(i\phi_1)$, where t_1^0 is the amplitude for $E = 0$, $\mathbf{A} = 0$. For the lower arm, we have the phase ϕ_2 which is the same as ϕ_1 but where the integration is carried out over the lower arm. This all assumes that the ring is very thin compared to its diameter, so that the areas enclosed by the outer and inner circumferences are essentially the same. Now the difference in the phases ϕ_1 and ϕ_2 is proportional to the magnetic flux enclosed by the ring:

$$\phi_2 - \phi_1 = \frac{e}{\hbar}\oint \mathbf{A}\cdot\vec{dl} = \frac{e}{\hbar}\int_S \mathbf{B}\cdot\vec{dS} = \frac{eBS}{\hbar} = \frac{2\pi\Phi}{\Phi_0}. \qquad (3.4)$$

Hence,

$$T_{12} = T_1 + T_2 + 2\sqrt{T_1 T_2}\cos(\frac{2\pi\Phi}{\Phi_0} + \varphi), \qquad (3.5)$$

where φ is the phase difference in the absence of a field. The relative oscillations are greatest when $T_1 = T_2$, in which case the total transmission oscillates between 0 and $4T_1$.

Note that this calculation did not assume anything about the position dependence of the magnetic field. The Aharonov-Bohm effect could be realized also if the actual field within the wires would vanish, as long as it is finite inside the ring. The transmission phase only depends on the total magnetic flux penetrating the ring.

The above calculation takes in fact into account only a subset of the electron paths. There are also paths that traverse around the ring multiple times. Associated with these paths, there then arise additional oscillations in the conductance vs. flux with periods h/Ne, $N = 1, 2, \ldots$ being an integer.

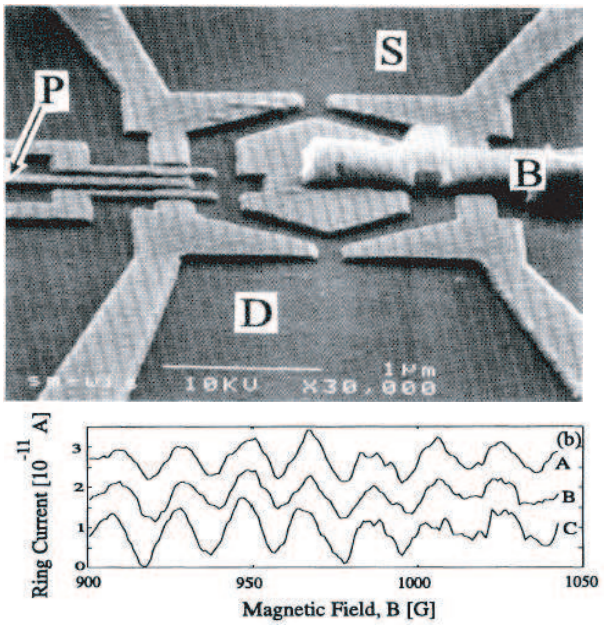


Figure 3.2: Up: SEM Micrograph of a setup for measuring the transmission phase shift of the island (on the left). Down: Aharonov-Bohm oscillations in the current as a function of the magnetic field for three different gate voltages controlling the island. Reproduced from A. Yacoby, *et al.*, Phys. Rev. Lett. **74**, 4047 (1995).

Recently this effect has been applied to measure phase shifts produced by other mesoscopic systems (see Fig. 3.2), such as quantum dots, placed in one of the ring arms.

3.2 Localization

In the previous section, we considered some examples where the conductance could be calculated from the scattering theory. Now consider a more complicated object, a wire with many scatterers in series. Let us concentrate first in the single-channel case. The transmission through two such scatterers was found to be (see Eq. 2.50)

$$T = \frac{T_1 T_2}{1 + R_1 R_2 - 2\sqrt{R_1 R_2}\cos(2\theta)}. \tag{3.6}$$

The dynamical phase θ depends strongly on the distance between successive scatterers. Therefore, we may assume that an ensemble of such scatterers averages the effects related with θ. To get an estimate of a resistance of a wire without the contact effects, let us consider the ensemble average of a four-probe resistance of two scatterers in series (for this discussion, the prefactor $2e^2/h$ is unessential).

$$\begin{aligned} \rho_{12} &= \left\langle \frac{1-T}{T} \right\rangle = \int_0^{2\pi} \frac{d\theta}{2\pi} \frac{1 + R_1 R_2 - T_1 T_2 - 2\sqrt{R_1 R_2}\cos(2\theta)}{T_1 T_2} \\ &= \frac{1 + R_1 R_2 - T_1 T_2}{T_1 T_2} = \frac{R_1 + R_2}{T_1 T_2}. \end{aligned} \tag{3.7}$$

Here the averaging sign $\langle\cdot\rangle$ simply refers to averaging over θ. Defining the individual resistances of the scatterers via

$$\rho_1 = \frac{R_1}{T_1}, \quad \rho_2 = \frac{R_2}{T_2}, \tag{3.8}$$

we get

$$\rho_{12} = \rho_1 + \rho_2 + 2\rho_1\rho_2. \tag{3.9}$$

From this discussion, it is not totally clear which quantity we should average over. For comparison, let us consider the four-probe resistance with averaged T's,

$$\rho_{\rm cl} = \frac{1 - \langle T\rangle}{\langle T\rangle} = \frac{1-T_1}{T_1} + \frac{1-T_2}{T_2} = \rho_1 + \rho_2. \tag{3.10}$$

That is, we get a classical scaling law (see also the first exercise problem for another way of obtaining the classical result). It turns out that for a proper description of phase-coherent elements, we have to use the law given in Eq. (3.8).

Now let us try to obtain a resistance as a function of length. Assume that $\rho_1 = \rho(L)$ describes a conductor of length L, and we add a small piece of length ΔL and resistance $\rho_2(\lambda) = \Delta L/\lambda$ to it. Now from Eq. (3.8) we obtain

$$\rho(L + \Delta L) = \rho(L) + (1 + 2\rho(L))\frac{\Delta L}{\lambda}. \tag{3.11}$$

In the limit $\Delta L \to 0$ it yields a differential equation

$$\frac{d\rho}{dL} = \frac{\rho(L+\Delta L) - \rho(L)}{\Delta L} = \frac{1 + 2\rho}{\lambda}. \tag{3.12}$$

The solution to this equation is

$$\rho(L) = \frac{1}{2}\left(e^{2L/\lambda} - 1\right). \tag{3.13}$$

Thus, the resistance of disordered wire grows exponentially with its length. This is the phenomenon of (Anderson) *localization*.[2] If the wire is much shorter than λ, we get an "Ohmic" result,

$$\rho(L) \approx \frac{L}{\lambda}, \quad L \ll \lambda. \tag{3.14}$$

Comparing this scaling to the Einstein relation, Eq. (1.3), we find that λ should be of the order of the mean free path in this single-mode case. It was shown by Thouless[3] that in the general quasi-onedimensional case with many modes, the localization length (the size of a localized wave packet within the scattering system) is

$$\lambda_{\rm loc} = \frac{2Ak_F^2\ell_{\rm el}}{3\pi^2} \sim M\ell_{\rm el}, \tag{3.15}$$

where A is the cross section of the wire and M is the number of modes. Thus, in a 100 nm $\times$ 100 nm metal wire with 10^5 modes (see discussion below Eq. (2.17)) and elastic mean free path some tens of nm, the localization length would be a few millimeters. This is much longer than the typical dephasing lengths in such wires, and thus Anderson localization in metal wires is difficult to observe. However, in semiconductor quantum wires or carbon nanotubes the number of modes is much less, and hence also the localization length can be taken clearly below ℓ_φ.

Figure 3.3 shows the results of a numerical simulation[4] of the ensemble-averaged conductance of a disordered wire as a function of its length. It illustrates how the wire turns from the "quasiballistic" (where G is almost independent of length) via the "diffusive" regime (where G scales inversely with L) to a localized regime (where G decays exponentially with L).

3.2.1 Weak localization

Expanding the resistance of Eq. (3.13) into second order yields

$$\rho(L) \approx \frac{L}{\lambda} + \left(\frac{L}{\lambda}\right)^2 = \rho_{\rm cl} + \Delta\rho. \tag{3.16}$$

The quantum correction $\Delta\rho$ to the classical resistance is quadratic in the length L. The corresponding correction to the classical conductance is independent of

[2] The name "Anderson" localization is typically used to make a difference to "Mott" localization, which takes place in interacting systems as the electron density grows.

[3] D. J. Thouless, Phys. Rev. Lett. **39**, 1167 (1977).

[4] Such a simulation typically considers a tight-binding lattice describing the scattering region, and the disorder is simulated by a random on-site potential. The width of the random distribution then determines the "magnitude" of disorder, i.e., $\ell_{\rm el}$.

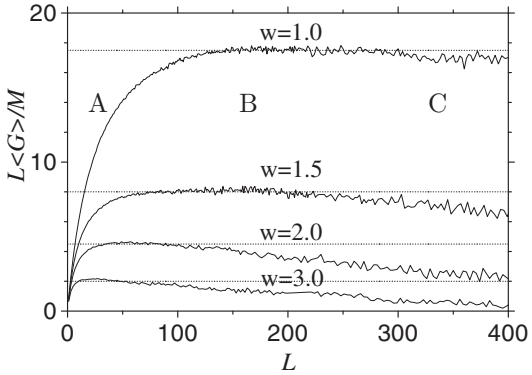


Figure 3.3: Numerically evaluated conductance for a disordered sample as a function of its length L. The four curves correspond to four magnitudes of disorder (different mean free paths). The three labels are for quasiballistic (A, $L \lesssim \ell_{\rm el}$), diffusive (B, $\ell_{\rm el} \ll L \ll \lambda_{\rm loc}$) and localized (C, $L > \lambda_{\rm loc}$) regions. Reproduced from T. T. Heikkilä, *et al.* Phys. Rev. B **60**, 9293 (1999).

L:

$$\frac{G}{2e^2/h} \approx \frac{1}{\rho_{\rm cl} + \Delta\rho} = \frac{1}{\rho_{\rm cl}} - \frac{\Delta\rho}{\rho^2} = \frac{1}{\rho_{\rm cl}} - 1. \tag{3.17}$$

The conductance of a disordered phase-coherent sample is thus reduced by $\Delta G = 2e^2/h$ compared to the classical conductance — provided the above expansion is realistic, i.e., the classical conductance far exceeds the correction. Although the derivation of this result was somewhat crude, this turns out to be the correct result.

3.2.2 Weak localization from enhanced backscattering

This *weak localization* effect can be understood by considering the different paths a particle takes in a disordered medium. Consider a process that scatters an incoming mode n (with wave vector $\vec{k}_n$ into an outgoing mode m (with wave vector $\vec{k}_m$) via multiple scattering events from impurities. The individual scattering effects can be written as scattering between states ("paths") with wave vectors $\vec{k}'_j$ and $\vec{k}'_k$. Thus, the total process (the "path" between states n and m) is of the form

$$\vec{k}_n \to \vec{k}'_1 \to \vec{k}'_2 \to \cdots \to \vec{k}'_N \to \vec{k}_m. \tag{3.18}$$

For a given number N of scattering events, there are $N!$ paths which go via the same wave vectors in different orders. Now denote these paths between states n and m with index p and the probability amplitude of each path with A_p^{mn}. Then the total scattering amplitude between states n and m is of the form

$$s_{mn} = \sum_p A_p^{mn}. \tag{3.19}$$

In a conductor much larger than the mean free path, the conductance is (almost) self-averaging:[5] the total conductance averages out the effects of individual scattering events, i.e., the forms or locations of the individual impurity potentials are not relevant, only their average effect. Then it is enough to consider the average transmission/reflection probability between modes n and m. This is

$$\langle |s_{mn}|^2 \rangle = \langle \sum_{pp'} A_p^{mn} (A_{p'}^{mn})^* \rangle = \langle \sum_{pp'} |A_p||A_{p'}| e^{i\varphi_{pp'}} \rangle. \tag{3.20}$$

Here $\varphi_{pp'}$ is the phase difference between the amplitudes A_p and $A_{p'}$. The disorder average of this phase factor vanishes except in two cases: a) $p = p'$ are the same paths. These processes yield the classical conductance of the wire. The second case is b) when $p' = \bar{p}$ is the time-reversed path of p, i.e., it goes through the same wave vectors $\vec{k}'_j$ as the path p, but in a reverse order (see Fig. 3.4). The latter corresponds to a closed path, and therefore contributes only when $m = n$. If the system is symmetric under time reversal, the amplitudes for the paths p and $\bar{p}$ have to equal.

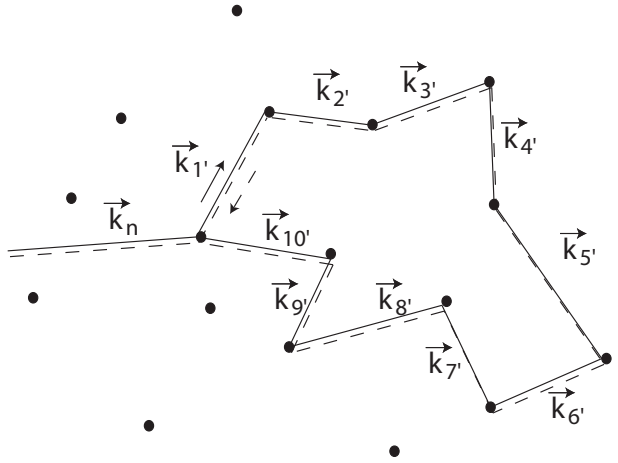


Figure 3.4: Time reversed paths with multiple scattering events.

Thus, for $n \neq m$, we obtain

$$\langle |s_{mn}|^2 = \sum_p \langle |A_p^{mn}|^2 \rangle \equiv T_{nm}^{\rm cl}, \tag{3.21}$$

whereas for $n = m$ we get

$$\langle |s_{nn}|^2 = \sum_p \langle |A_p^{nn}|^2 + A_p^{nn} (A_{\bar{p}}^{nn})^* \rangle = 2R_{nn}^{\rm cl}. \tag{3.22}$$

[5] See the section on universal conductance fluctuations.

As a result, the two-probe conductance is

$$\frac{G}{2e^2/h} = \bar{T}_{12} = M_1 - \bar{R}_1 = M_1 - \sum_m \sum_{k\neq m} |s_{nm}|^2 - \sum_m |s_{mm}|^2 = M_1 - \bar{R}_1^{\rm cl} - \sum_n R_{nn}^{\rm cl}. \tag{3.23}$$

Here the superscripts "cl" denote the "classical" case, i.e., when the enhanced backscattering is suppressed.

We may assume that the scattering is on average isotropic, i.e., mode-independent. Then, the classical backscattering probability between any two modes is

$$R_{nm}^{\rm cl} = \frac{L}{(L+\ell_{\rm el})M}. \tag{3.24}$$

This would yield the classical result for the conductance, $G \sim M_1 \ell_{\rm el}/L$. Now using this in Eq. (3.23) yields

$$\frac{G}{2e^2/h} = M_1 \frac{\ell_{\rm el}}{L} - \frac{L}{L+\ell_{\rm el}} \approx M_1 \frac{\ell_{\rm el}}{L} - 1. \tag{3.25}$$

We thus again find the conductance correction $\Delta G = 2e^2/h$.

3.2.3 Microscopic calculation with Green's-function techniques (**)

The weak-localization correction to the conductance can also be calculated using the equilibrium Green's function theory of condensed-matter physics. This also allows to treat the effect of a magnetic field and of dephasing quantitatively. For linear response, the conductivity can be obtained from the equilibrium Green's function using the Kubo formula[6]

$$\langle\sigma\rangle = \frac{\hbar e^2}{\pi V}\left(\frac{\hbar}{m}\right)^2 \sum_{kk'} k_z k_z' \langle\langle k|G^R|k'\rangle\langle k'|G^A|k\rangle\rangle. \tag{3.26}$$

Here V is the volume of the sample (taken to infinity at the end of the calculation), and G^R and G^A are the retarded and advanced Green's functions of the electron system. The outer brackets refer to disorder averaging.

The classical (Drude or Einstein) conductivity $\sigma = ne^2\tau/m$ can be found by approximating

$$\langle G^R G^A\rangle \approx \langle G^R\rangle\langle G^A\rangle. \tag{3.27}$$

It turns out that the major corrections to this result can be expressed with two types of Feynman diagrams: diffusons and cooperons. The previous simply modifies the diffusion constant to take into account the possible average angular dependence of scattering, but in spherically symmetric case its contribution

[6] See, for example, Chapter 1 in the book T. Dittrich, *et al.*, 'Quantum transport and dissipation' (Wiley-VCH, 1998). Essentially the Kubo formula relates the equilibrium fluctuations to the linear conductance through the fluctuation-dissipation theorem.

vanishes. The latter contains the weak localization correction. The calculation of these diagrams is beyond the scope of this course, and we refer to the book of T. Dittrich *et al.* for the details. However, below we quote some of the results of this theory.

3.2.4 Dephasing

Weak localization is an interference effect. If the time-reversed paths are longer than the dephasing length ℓ_φ, their relative phase gets randomized and as a result, their contribution to the enhanced backscattering is lost. In the exercise problems, you will show that in the one-dimensional case where the width W of the sample is much smaller than ℓ_φ, but its length is much longer, the correction to the conductance is

$$\Delta G(L \gg \ell_\varphi) = \frac{2e^2}{h}\frac{\ell_\varphi}{L}. \tag{3.28}$$

This means that the correction in the conductivity (proportional to the relative correction $\Delta G/G$) scales as

$$\Delta\sigma \sim \frac{\ell_\varphi}{W} \tag{3.29}$$

In the two-dimensional case also the width of the sample exceeds the dephasing length. Then the correction turns out to be (see Exercise 3.4)

$$\Delta\sigma = \frac{2e^2}{\pi h}\ln\left(\frac{\ell_\varphi}{\ell_{\rm el}}\right). \tag{3.30}$$

(Here: add text about the theories for the temperature-dependent ℓ_φ and origins of dephasing).

3.2.5 Magnetic field effect

Assume we lift the time reversal symmetry by applying a perpendicular magnetic field to the sample. Due to this magnetic field, the paths between states n and m will obtain an additional phase,

$$A_p^{mn} \to A_p^{mn}\exp\left(i\frac{e}{\hbar}\int_p \mathbf{A}\cdot\vec{dl}\right). \tag{3.31}$$

This additional phase is irrelevant for the paths whose phase dependence averaged out in disorder averaging. It will also not contribute to the terms $A_p^{mn}(A_p^{mn})^*$ in Eq. (3.20). However, the contribution from a pair of a path p and its time reverse $\bar{p}$ it changes to

$$A_p^{mn}(A_{\bar{p}}^{mn})^* = |A_p^{mn}|^2\exp\left(2i\frac{e}{\hbar}\oint_p \mathbf{A}\cdot\vec{dl}\right) = |A_p^{mn}|^2\exp\left(2i\frac{2\pi\Phi_P}{\Phi_0}\right). \tag{3.32}$$

Here Φ_P is the flux through the area enclosed by the path. Now as the different paths enclose different areas, the sum over the paths will eventually wash out

this additional backscattering term. The critical magnetic field required for this corresponds roughly to a flux Φ_P equal to one flux quantum through the path with the largest loop area. As only the paths whose length are below the dephasing length ℓ_φ contribute, we may estimate this area as roughly ℓ_φ^2. Thus, the critical field is of the order of

$$B_c \sim \frac{h}{e\ell_\varphi^2}. \tag{3.33}$$

For $\ell_\varphi = 1\mu$m, this field equals 40 G or 4 mT.

The microscopic linear-response calculation mentioned above yields also the form of the conductance correction as functions of the magnetic field and the dephasing time $\tau_\varphi = \ell_\varphi^2/D$. In the two-dimensional case this is[7]

$$\Delta\sigma = -\frac{e^2}{\pi h}\left[\Psi\left(\frac{1}{2}+\frac{\hbar}{4eBD\tau_{\rm el}}\right) - \Psi\left(\frac{1}{2}+\frac{\hbar}{4eBD\tau_\varphi}\right)\right]. \tag{3.34}$$

Here $\Psi(x)$ is the digamma function ($\Psi(x) = \Gamma'(x)/\Gamma(x)$ where $\Gamma(x)$ is the Gamma function) and τ is the scattering time.

An example of the measured weak localization correction in a multi-walled carbon nanotube is shown in Fig. 3.5.

3.2.6 Interaction contribution

Weak localization makes rise to a temperature dependence in the conductance of a disordered metal wire, as the dephasing length depends on the temperature. In the case of two-dimensional localization, this temperature dependence is logarithmic. But there is also another effect that gives rise to a temperature dependent conductivity. This effect is due to the scattering of an electron with the rest of the electron system and its calculation requires the use of the (Kubo) Green's-function theory mentioned above. In the one-electron picture, this contribution can be calculated with the self-consistent Hartree approximation, inserting the electron density into the Poisson equation and calculating the potential self-consistently. This potential turns out to be temperature dependent, and therefore also the resulting conductivity variation depends on temperature. The calculation of this correction is outside the scope of this course, so we just state the result. The conductance correction is also logarithmic,

$$\Delta\sigma_{e-e} = -\frac{e^2 g}{\pi\hbar}\ln\left(\frac{\hbar}{k_B T\tau}\right), \tag{3.35}$$

where g is a numerical factor of the order of unity. Thence, the weak-localization correction and the interaction correction both decrease the conductivity in a temperature-dependent fashion. They can only be separated by applying a magnetic field, which kills the WL contribution, but does not essentially affect the interaction contribution.

[7] See, e.g., Eq. (5.5.30) in S. Datta's book.

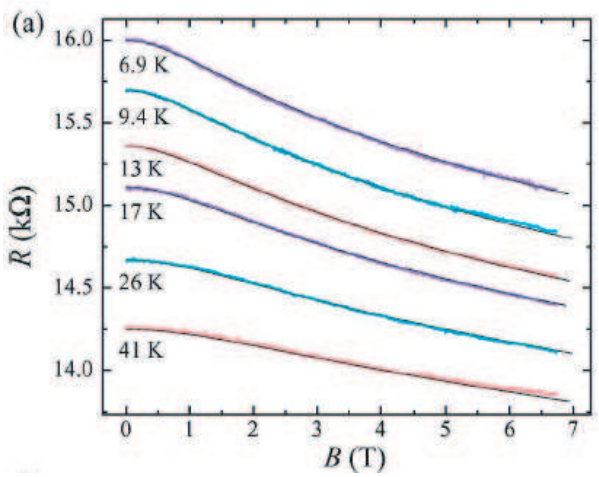


Figure 3.5: Weak localization correction to the conductance as a function of magnetic field in a disordered multiwalled carbon nanotube measured in the Low Temperature laboratory, TKK. The experimental results are fitted with the 2d WL theory (black lines on top of the experimental points), Eq. (3.34) modified to take into account the geometry of the nanotube. The temperature dependence is due to that of the dephasing time. Reproduced with permission from R. Tarkiainen, *et al.*, Phys. Rev. B **69**, 033402 (2004).

Weak antilocalization

In some materials, one has also observed a weak antilocalization effect, which results from the spin-orbit scattering effect, and increases the conductivity compared to the classical case. It is essentially based on a similar picture as the weak localization effect, but the presence of spin-orbit scattering simply changes the direction of the correction.[8]

3.2.7 Strong localization (**)

Above we saw that a one-dimensional conductor is in fact never a conductor in the strict statistical limit, where its size tends to infinity. This is because of the localization of electron wave functions within the localization length scale $\lambda_{\rm loc}$. This rather surprising result has in fact been generalized to two and three dimensions as well: for strong enough disorder, the localization length $\lambda_{\rm loc}$ is finite for a sample of any dimension. But $\lambda_{\rm loc}$ may be very long. For example, in a two-dimensional case it is[9]

$$\lambda_{\rm loc,2d} \approx \ell_{\rm el} \exp\left(\frac{\pi}{2} k_F \ell_{\rm el}\right). \tag{3.36}$$

[8] For more information on this effect, see S. Chakravarty and A. Schmid, Phys. Rep. **140**, 193 (1986).

[9] See Eq. (1.118) in the book by T. Dittrich, *et al.*

For a weakly disordered sample, $\ell_{\rm el} \gg \lambda_F$, this localization length would become indeed very large. However, it always stays finite,[10] whereas in three dimensions, there is a transition from a non-localized state ($\lambda_{\rm loc} = \infty$) to a localized one ($\lambda_{\rm loc}$ is finite). This *Anderson transition* is a quantum phase transition, as it takes place at a zero temperature.

For those interested, the existence of the localization transition can be shown by using the single-parameter scaling theory.[11] The arguments are based on a single-parameter scaling hypothesis: it is assumed that as the size of the system tends to infinity, its properties are described by a single relevant parameter, which in this case is the ensemble-averaged dimensionless conductance $g = \langle G\rangle/(2e^2/h)$.

The scaling arguments go as follows. Consider the scaling of g in a d-dimensional system composed of b^d blocks of size L. Combine these blocks into a single larger system with a scale bL. According to the scaling hypothesis, there is a function f that satisfies

$$g(bL) = f(b, g(L)). \tag{3.37}$$

This hypothesis can also be written in the form

$$\frac{d\ln g(L)}{d\ln L} = \beta(g(L)) \tag{3.38}$$

with some function β. Thus, it is assumed that the scaling of $g(L)$ at some point L only depends on the value of g at that point.

At large and small g we can find the asymptotics of β from general physical arguments. For large g, the conductance follows Ohm's law,

$$G(L) = \sigma L^{d-2}. \tag{3.39}$$

For the β-function, this means that for $g \to \infty$,

$$\frac{d\ln g}{d\ln L} = \frac{\frac{d\sigma\exp((d-2)\ln L)}{d\ln L}}{\sigma L} = d-2. \tag{3.40}$$

For $g \ll 1$, we have seen above that exponential localization is valid and g falls off exponentially:

$$g = g_0 e^{-L/\lambda_{\rm loc}} \tag{3.41}$$

Hence, in this limit

$$\beta = \frac{d(-e^{\ln L}/\lambda_{\rm loc})}{d\ln L} = -L/\lambda = \ln(g/g_0(d)). \tag{3.42}$$

The rough behavior of $\beta(g)$ for different dimensionalities d=1,2,3 is sketched in Fig. 3.6. For $\beta > 0$, the conductance grows with an increasing L, i.e., it behaves in a "metallic" manner, whereas for $\beta < 0$, it decreases when increasing the scale L. Thus, the point $\beta = 0$ indicates a metal-insulator transition, only present in strictly 3-dimensional systems.

[10] This is so for non-interacting electrons. There is evidence of nonlocalized states of weakly disordered two-dimensional conductors.

[11] See Anderson, *et al.*, Phys. Rev. Lett. **42**, 673 (1979).

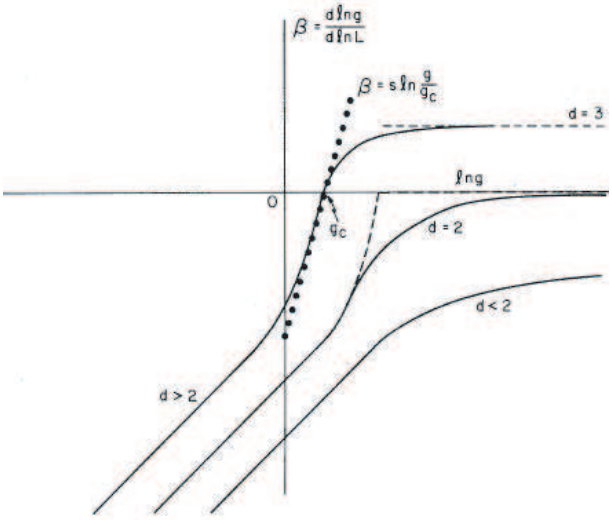


Figure 3.6: Sketch of the behavior of the β-function for different dimensionalities of the conductor.

3.3 Universal conductance fluctuations

Above, we considered the behavior of the conductance averaged over the disorder configuration. In a macroscopic sample (much larger than the dephasing length), the conductance would be self-averaging, i.e., its conductance would equal the disorder-averaged conductance, and changing the disorder configuration but keeping its strength (mean free path) constant would leave the conductance unaltered.

This is different for mesoscopic conductors. The conductance depends greatly on the disorder configuration even in the diffusive limit $L \gg \ell_{\rm el}$. This dependence shows up in the magnetic field or Fermi velocity dependence of the conductance. The latter can be tuned by a gate voltage coupled capacitively to the conductor through the tuning of the Fermi energy. Changing either the field or the gate amounts to changing the phase shifts the electrons acquire between scattering events. This reproducible "magnetofingerprint" is a typical property of mesoscopic samples whose conductance is not far from the conductance quantum $2e^2/h$. It turns out that the variance of the conductance for different disorder configurations is independent of the conductance (in the weak localization regime), and only depends on the global symmetry of the sample, such as the time-reversal symmetry. This is why the variations in the conductance are called *universal conductance fluctuations*.

Let us estimate the variance of the conductance. It is given by

$$\langle \delta G^2 \rangle = \langle G^2 \rangle - \langle G \rangle^2 = \frac{4e^4}{h^2} \langle \delta \bar{R_1}^2 \rangle, \tag{3.43}$$

where $\bar{R}_1 = \sum_{m,n} |s^{11}_{nm}|^2$ is the total reflection probability into lead one. Assuming that the variations in the different $R_{nm} = |s^{11}_{nm}|^2$ are independent, we find

$$\langle \delta \bar{R}_1^2 \rangle \approx M^2 \langle \delta R_{nm}^2 \rangle. \tag{3.44}$$

Here we neglected the effect from the enhanced backscattering. Now the variation of the reflection coefficients for individual channels can be expressed in terms of the different paths as above for weak localization,

$$\begin{aligned} \langle R_{nm}^2 \rangle &= \sum_{pp'p''p'''} \langle A_p A_{p'} A^*_{p''} A^*_{p'''} \rangle = \sum_{pp'p''p'''} |A_p|^2 |A_{p'}|^2 (\delta_{pp''}\delta_{p'p'''} + \delta_{pp'''}\delta_{p'p''}) \\ &= 2\langle R_{nm} \rangle^2. \end{aligned} \tag{3.45}$$

Therefore, we obtain $\langle \delta R_{nm}^2 \rangle = \langle R_{nm} \rangle^2 \approx L^2/(M(L+\ell_{\rm el}))^2$. Substituting to Eqs. (3.44) and (3.43) we get

$$\langle \delta G^2 \rangle = \frac{4e^2}{h} \frac{L^2}{(L+\ell_{\rm el})^2} \approx \frac{4e^2}{h}. \tag{3.46}$$

Thus, the average variation in the conductance is of the order of a single conductance quantum. The exact variance can be obtained with the random matrix theory.[12] The result is that $\sqrt{\langle \delta G^2 \rangle} = ce^2/h$, where $c \approx 0.36$ in the presence and $c \approx 0.26$ in the absence of time-reversal symmetry.

Figure 3.7 shows an example of measured conductance fluctuations in a normal-superconducting contact.

3.3.1 Effect of dephasing

Let us now estimate the effect of dephasing on the magnitude of fluctuations. Assume we have a conductor of length $L \gg \ell_\varphi$. Then we can divide it into $N = L/\ell_\varphi$ coherent subunits, whose resistances we can add incoherently. The variance of a resistance R_i in a single subunit is

$$\mathrm{var}(R_i) = \mathrm{var}(G_i) R_i^4. \tag{3.47}$$

Now the total resistance is $R = NR_i$ and the total resistance variation is $\mathrm{var}R = N\mathrm{var}(R_i)$. The total conductance variation can now be obtained from

$$\mathrm{var}(G) = \frac{\mathrm{var}R}{R^4} = \frac{N\mathrm{var}(G_i)R_i^4}{N^4 R_i^4} = N^{-3}\mathrm{var}G_i = \left(\frac{\ell_\varphi}{L}\right)^3 c^2 (\frac{e^2}{h})^2. \tag{3.48}$$

Thus, the fluctuations decay in a power-law fashion with the decreasing dephasing length and can thus persist to very high temperatures.

[12] See C. W. J. Beenakker, Rev. Mod. Phys. **69**, 731 (1997).

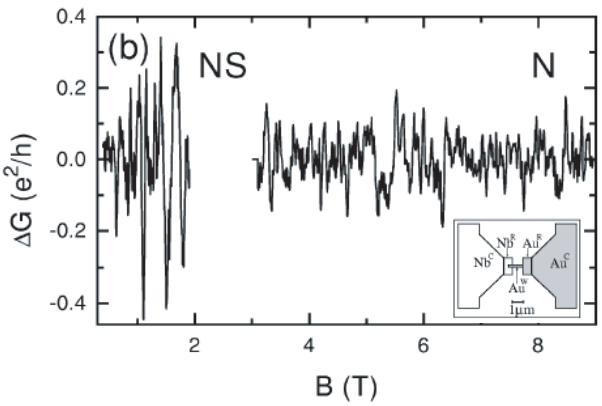


Figure 3.7: Universal conductance fluctuations measured in a normal-superconducting contact sketched in the inset. Superconductivity enhances the magnitude of fluctuations, which is seen when the field is increased above the critical field of the superconductor. Reproduced from K. Hecker, *et al.*, Phys. Rev. Lett. **79**, 1547 (1997).

3.4 Problems

1. Compare the transmission probabilities of a two-scatterer system obtained by summing the transmission probabilities for individual "Feynman paths" of the particle and the quantum-mechanical result obtained in Exercise 2.3. Show that the previous method results into a classical Ohm's law -type summing of the resistances.

2. Consider an Aharonov-Bohm ring, where both arms contain a single propagating mode. Assume that the scattering matrix for the three-way junction at both ends (positions "A" and "B" in Fig. 3.1) are described with the scattering matrix

$$S = \begin{pmatrix} c & \sqrt{\epsilon} & \sqrt{\epsilon} \\ \sqrt{\epsilon} & a & b \\ \sqrt{\epsilon} & b & a \end{pmatrix}, \tag{3.49}$$

where a, b, c, and ϵ are real numbers. (a) Show that in order to ensure unitarity of S, we have to have

$$c = \pm\sqrt{1-2\epsilon}, \quad a = (1-c)/2, \quad b = -(1+c)/2. \tag{3.50}$$

Thus the entire matrix is specified by the function ϵ. (b) Assuming that the mode in the upper arm obtains a phase shift $\theta + \phi/2$ and that in the lower arm a phase shift $\theta - \phi/2$, where $\phi = \Phi/\Phi_0$ and $\theta = \sqrt{2mE}(\pi r/\hbar)$ is the dynamical phase (equal in both arms), calculate the total transmission probability for the ring as functions of ϵ, θ and ϕ. Plot the result for $\theta = 0.1$, $\epsilon = 0.1$ as a function of ϕ.

3. Calculate the weak localization correction ΔG to the conductance of a quasi one-dimensional wire with length $L \gg \ell_\varphi$ and width $W \ll \ell_\varphi$. Hint: divide the wire into L/ℓ_φ coherent sections. Then sum up the resistances (with the WL correction) of these sections. Calculate also the relative correction $\Delta G/G$. This is also the conductivity correction.

4. Show that in the two-dimensional case $W, L \gg \ell_\varphi$, the conductivity is

$$\sigma_Q = \sigma_{CL} - \frac{2e^2}{\pi h} \ln(\ell_\varphi/\ell_{\rm el}). \tag{3.51}$$

   Hint: show that the conductance of a cylindrical conductor is

$$G = \frac{\pi\sigma}{\ln(W_{\rm max}/W_{\rm min})}, \tag{3.52}$$

   where $W_{\rm max}$ and $W_{\rm min}$ are the outer and inner radii of the cylinder. Now combine the conductances of such circular units with $W_{\rm max} = \ell_\varphi$ and $W_{\rm min} \sim \ell_{\rm el}$.

Chapter 4

Semiclassical Boltzmann theory

Above, the transport problem was formulated such that the energy distribution function $f(E)$ of the electrons was specified inside the reservoirs, and the current depended on the details of scattering inside the sample, and on the difference in $f(E)$ between the different reservoirs. In some cases, it is also useful to ask what the energy distribution function is inside the scattering region, and how the current or some other observables can be related to this space-dependent distribution function. It turns out that this is particularly useful in diffusive structures whenever one wants to treat the current fluctuations, the effects or magnitude of relaxation, or certain nonlinear phenomena. It is also useful in estimating the thermal balance of a mesoscopic structure, which determines the static temperature of the electron gas in the reservoirs. For a review of such phenomena, see F. Giazotto, T. T. Heikkilä, A. Luukanen, A. M. Savin, and J. P. Pekola, Rev. Mod. Phys. **78**, 217 (2006).

The aim of this lecture is to understand in which cases the electron energy distribution function can actually be defined, and to find how it behaves in a mesoscopic sample. We also aim to find prescriptions for computing the values of different observables from this behavior.[1]

In macroscopic samples, the electron system under study can be described, even in the presence of an applied voltage, by an equilibrium (Fermi) distribution $f^0(E)$[2] with given (but perhaps varying) (chemical) potential profile $\mu(x) = eV(x)$ and temperature T equal to that of the underlying lattice, T_l (i.e., the temperature of the cryostat in a low-temperature experiment). In mesoscopic

[1] Useful references in this context are at least Chapters 13-16 in N. W. Ashcroft and N. D. Mermin, "Solid State Physics" (Saunders College Publishing), Sect. 10.4 in M. Plischke and B. Bergersen, "Equilibrium Statistical Physics" (2nd edition, World Scientific), K. Nagaev, Phys. Rev. B **52**, 4740 (1995), and E. Sukhorukov and D. Loss, Phys. Rev. B **59**, 13054 (1999).

[2] In this chapter, the Fermi function $f^0(E;\mu,T) = (1+\exp((E-\mu)/kT))^{-1}$ is denoted with the superscript 0.

systems, the temperature T_e of the electrons can become position dependent, and even the distribution function can deviate from the Fermi-function form.

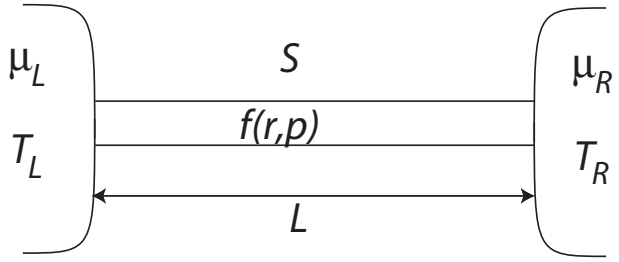


Figure 4.1: System studied in this chapter. Reservoirs at the ends have well-defined potentials μ_i and temperatures T_i.

We again assume that the system under consideration can be separated in the studied, interesting subsystem S and the reservoirs (electrodes) to which it is connected (see Fig. 4.1). Let us rephrase the assumptions about the reservoirs relevant for the study of distribution functions

- There is sufficiently strong relaxation within the reservoirs such that the electron distribution function in them is a Fermi function $f^0(E)$.
- There are no voltage or temperature gradients inside the reservoirs. The previous requires that the resistance of the subsystem S dominates over the resistance of the reservoirs, and the latter is a similar condition for the thermal resistance. The previous condition can be overcome with four-probe measurements. In practice, the latter condition is often not satisfied, and the reservoir temperature has to be calculated for each applied voltage.

On the course of the treatment of relaxation, it will become fairly clear how these assumptions can be met.

The aims of this description can now be stated more clearly:

- What is the distribution function $f(\vec{r},\vec{p})$ for given $\{\mu_L, T_L\}$ and $\{\mu_R, T_R\}$?
- From this $f(\vec{r},\vec{p})$, how do we find the observables such as charge and heat current?

The answer to these questions depends in general on (i) the type of the studied system S and of the reservoirs (for example, are they formed from normal metals, semiconductors or superconductors) and ii) the type and strength of electron scattering within S.[3]

In general, the scattering mechanisms are described by a collision integral which we specify below. But to get some feeling of their strength, one may apply the relaxation-time approximation and attribute a rate or time scale for

[3]It can also be modified due to supercurrents flowing in the system, see for example T. T. Heikkilä, T. Vänskä, and F. K. Wilhelm, Phys. Rev. B **67**, 100502 (2003).

the various scattering events. Moreover, for the given relaxation time scale, one may find a corresponding length scale which is, loosely speaking, the average distance a particle traverses between the scattering events. The relevant scattering mechanisms and the corresponding scales are

- Elastic scattering from impurities, dislocations, etc. The length scale for this process is the elastic mean free path $\ell_{\rm el}$.
- Electron–electron scattering, $\ell_{\rm e-e}$
- Electron–phonon scattering, $\ell_{\rm e-ph}$.

The latter two are inelastic, i.e., energy is transferred in the scattering process. Such a scattering makes rise to a random phase shift to the particle wave function, and thus the dephasing length ℓ_φ discussed in the previous Chapter is related to these scattering events.

At low temperatures in metals the typical order of these length scales is

$$\lambda_F \ll \ell_{\rm el} \ll \ell_{\rm e-e}, \ell_{\rm e-ph}. \tag{4.1}$$

The relative order between $\ell_{\rm e-e}$ and $\ell_{\rm e-ph}$ depends strongly on temperature: typically down to roughly 1 K $\ell_{\rm e-ph} < \ell_{\rm e-e}$, and below this $\ell_{\rm e-e} < \ell_{\rm e-ph}$. This is due to the strong temperature dependence of the electron-phonon scattering rate.

4.1 Semiclassical Boltzmann equation

The electron distribution function $f(\vec{r},\vec{p})$ is defined so that the average number of electrons in the element $\{\vec{dr},\vec{dp}\}$ around the point $\{\vec{r},\vec{p}\}$ in the six-dimensional position-momentum space is

$$f(\vec{r},\vec{p})\frac{\vec{dr}\vec{dp}}{(2\pi)^3}. \tag{4.2}$$

Strictly speaking, Heisenberg uncertainty principle states that we cannot simultaneously talk about the precise momentum and position of an electron, questioning the mere validity of the concept of the distribution function $f(\vec{r},\vec{p})$. However, it can be defined in a coarse-grained manner, averaging the position (and momentum) values over some interval (i.e., the width of the wave packet). In practice, this means that with this approach, we may only describe phenomena occurring at scales much larger than the electron Fermi wavelength λ_F. This implies that with this description, we will not be able to deal with such interference effects as localization or resonant tunneling, and $f(\vec{r},\vec{p})$ essentially describes the classical kinetics of the electrons under the applied fields.[4]

[4]The semiclassical Bolzmann equation can be rigorously derived from the full quantum-mechanical description in the appropriate limit using the quasiclassical Keldysh Green's-function technique, see, for example, J. Rammer and H. Smith, Rev. Mod. Phys. **58**, 323 (1986).

Assume at time t the electrons are described by the distribution $f(\vec{r},\vec{p},t)$. If there is no scattering, after time dt this must equal to the distribution $f(\vec{r}+\vec{v}dt,\vec{p}+\vec{F}dt,t+dt)$, where $\vec{v}=\vec{p}/m$ is the electron velocity, and $\vec{F}$ is a force acting on them (Lorenz force due to electromagnetic fields). The difference between them must therefore result from "collisions", i.e., scattering from impurities, other electrons or phonons. Let us denote this difference by the term $I_{\rm coll}[f]dt$. The term $I_{\rm coll}[f]$ is called a collision integral, and it is generally a functional of $f(\vec{r},\vec{p})$. We thus have

$$f(\vec{r}+\vec{v}dt,\vec{p}+\vec{F}dt,t+dt)-f(\vec{r},\vec{p},t)=I_{\rm coll}[f]dt. \tag{4.3}$$

Expanding to the first order in dt and taking the limit $dt\to 0$ we get the Boltzmann equation

$$\left(\partial_t+\vec{v}\cdot\partial_{\vec{r}}+\vec{F}\cdot\partial_{\vec{p}}\right)f(\vec{r},\vec{p},t)=I_{\rm coll}[f]. \tag{4.4}$$

The terms $\partial_{\vec{x}}\equiv\partial/\partial_{\vec{x}}$ should be understood as partial derivatives: they act only on the respective coordinates of f. In what follows, we neglect the magnetic fields and take only the electric field $\vec{E}$ into account. Then $\vec{F}=-e\vec{E}$.

In metals the dispersion relation $\varepsilon(\vec{p})$ between the kinetic energy and the momentum around the Fermi level $\varepsilon(\vec{p})\approx E_F$ is almost independent of the direction of the momentum.[5] Hence, we may describe the momentum by its direction $\hat{p}$ and by the kinetic energy $\varepsilon(\vec{p})\approx\varepsilon(|\vec{p}|)$. Then, including the effect of the electric field as a potential energy term $\mu(\vec{r})$, $\vec{E}=-\partial_{\vec{r}}\mu(\vec{r})$, so that we get the total energy $E=\varepsilon+\mu(\vec{r})$, we may write

$$\begin{aligned}&\vec{v}\cdot\partial_{\vec{r}}f(\vec{r},\vec{p})-e\vec{E}\cdot\partial_{\vec{p}}f(\vec{r},\vec{p})=\vec{v}\cdot\partial_{\vec{r}}f(\vec{r},\hat{p},E)+(\partial_{\vec{r}}\mu(\vec{r}))\cdot(\partial_{\vec{p}}E)\partial_E f(\vec{r},\hat{p},E)\\&=\vec{v}\cdot[\partial_{\vec{r}}+(\partial_{\vec{r}}E)\partial_E]f(\vec{r},\hat{p},E)=\vec{v}\cdot\nabla f(\vec{r},\hat{p},E).\end{aligned} \tag{4.5}$$

Here ∇ is the total position derivative. In the derivation, we used the fact that $\vec{v}=\partial_{\vec{p}}E$.

Thus, the Boltzmann equation becomes

$$\left(\partial_t+\vec{v}\cdot\nabla\right)f(\vec{r},\hat{p},E)=I_{\rm coll}[f]. \tag{4.6}$$

This equation is solved in the exercises in the ballistic limit $I_{\rm coll}[f]=0$. For the general case, one should specify $I_{\rm coll}[f]$, solve Eq. (4.6), and calculate the observables from the solution.

4.2 Observables

As the number of electrons in the volume element $d\vec{p}$ is $2f(\vec{p})d\vec{p}/(2\pi)^3$,[6] the charge current density at position $\vec{r}$ is

$$\vec{j}_C(\vec{r})=-e\int\frac{d^3p}{4\pi^3}\vec{v}(\vec{p})f(\vec{r},\vec{p})=-e\int dEN(E)\int\frac{d\hat{p}}{4\pi^3}\vec{v}(\hat{p})f(\vec{r},\hat{p},E). \tag{4.7}$$

[5] This property is typically denoted as a "spherical Fermi surface".

[6] The number 2 comes from spin.

Here $N(E) = d|\vec{p}|/dE$ is the density of states. Note that in the absence of scattering ($I_{\rm coll}[f] = 0$), Eq. (4.6) can be considered as the continuity equation for a "spectral current" $\vec{v}f(\vec{r}, \hat{p}, E)$ — this is simply a consequence of the fact that in the absence of scattering, elastic or inelastic, momentum and energy are conserved.

From $f(\vec{r}, \hat{p}, E)$ we may also calculate the heat current $\vec{j}_Q(\vec{r})$. Using the thermodynamic identity $dQ = dU - \mu dN$ between the change dQ in heat, dU in the internal energy, and dN in the number of particles, we obtain

$$\vec{j}_Q(\vec{r}) = \int dE N(E)(E-\mu) \int \frac{d\hat{p}}{4\pi^3} \vec{v}(\hat{p}) f(\vec{r}, \hat{p}, E). \tag{4.8}$$

This quantity is briefly described below in the context of thermoelectric effects.

4.3 Elastic scattering and diffusive limit

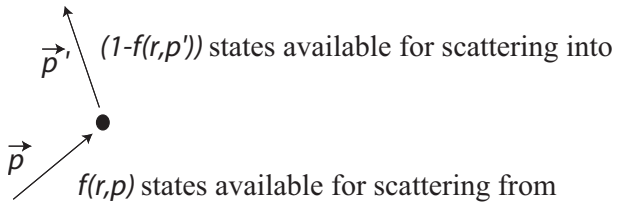
Elastic scattering from impurities can be described by the collision integral of the form

$$I_{\rm el}[f] = \sum_{\vec{p}'} (J_{\vec{p}'\vec{p}} - J_{\vec{p}\vec{p}'}), \tag{4.9}$$

where

$$J_{\vec{p}\vec{p}'} = W_{\vec{p}\vec{p}'}(\vec{r}) f(\vec{r}, \vec{p}, t)(1 - f(\vec{r}, \vec{p}', t)) \tag{4.10}$$

describes scattering from the momentum state $\vec{p}$ to state $\vec{p}'$. For a description of elastic scattering, we have to have $|\vec{p}| = |\vec{p}'|$ and thus it is enough to describe the scattering between directions $\hat{p}$ and $\hat{p}'$. Such a description is based on the idea that the scatterer is "fixed" (i.e., has an infinite mass) and the scattering can be pictorially represented as



The function $W_{\hat{p}\hat{p}'}(\vec{r})$ describes the strength and position of the scattering potential and it may depend in general on both $\hat{p}$ and $\hat{p}'$. The general solution of the Boltzmann equation with elastic scattering would require the knowledge of the position and type of each scatterer, i.e., of the function $W_{\hat{p}\hat{p}'}(\vec{r})$. However, the problem simplifies considerably if we assume that there are many scatterers and that it is enough to know $f(\vec{r}, \vec{p})$ only

- for a region much larger than the typical separation of scatterers, $\ell_{\rm el}$ (that is, we again use a coarse-grained description) and
- ensemble averaged over all positions of the scatterers.

In fact, it can be shown that these two requirements are analogous as the classical $f(\vec{r},\vec{p})$ is a self-averaging quantity (limit of large system corresponds to an effective ensemble average). This would not be the case if we were to include interference effects due to the universal conductance fluctuations. The tendency of the elastic scattering is to uniformly span all the momentum directions such that the distribution function becomes almost independent of the momentum direction $\hat{p}$.

We may then expand the dependence on the direction $\hat{p}$ in spherical harmonics,

$$f(\vec{r},\hat{p},E) \approx f_0(\vec{r},E) + \vec{\delta f}(\vec{r},E)\cdot\hat{p}. \tag{4.11}$$

As this is an expansion, we assume that the second term is much smaller than the first. As $f_0(\vec{r},E)$ is direction independent, it does not contribute to $I_{\rm el}[f]$ (we would get $J_{\hat{p}\hat{p}'} = J_{\hat{p}'\hat{p}}$). We may hence write

$$I_{\rm el}[f] \approx 2\vec{\delta f}\sum_{\hat{p}'} W_{\hat{p}\hat{p}'}(\hat{p}'-\hat{p}). \tag{4.12}$$

Here we assumed that $W_{\hat{p}\hat{p}'} = W_{\hat{p}'\hat{p}}$.[7] For s-wave scattering (assuming that on average the scatterers scatter the particle in all directions with the same probability), $W_{\hat{p}\hat{p}'}$ is independent of the angle between $\hat{p}$ and $\hat{p}'$. Then the first term averages to zero and we get

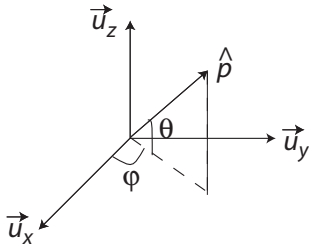
$$I_{\rm el}[f] = -\frac{1}{\tau}\vec{\delta f}\cdot\hat{p}, \quad \frac{1}{\tau} = \sum_{\hat{p}'} W_{\hat{p}\hat{p}'}. \tag{4.13}$$

The quantity τ is called the scattering time.

The Boltzmann equation can now be written as

$$(\partial_t + \vec{v}\cdot\nabla)\left(f_0(\vec{r},E,t) + \vec{\delta f}(\vec{r},E,t)\cdot\hat{p}\right) = -\frac{1}{\tau}\vec{\delta f}(\vec{r},E,t)\cdot\hat{p} + I_{\rm inel}[f_0]. \tag{4.14}$$

Here $I_{\rm inel}[f]$ describes inelastic scattering and we assume it to be much weaker than elastic scattering, such that it essentially only can see that lowest-order term f_0. Remembering that $\vec{v} = \vec{p}/m = |\vec{p}|/m\hat{p} \approx v(p)\hat{p}$, we can now integrate Eq. (4.14) over the directions $\hat{p}$. This is the same as making the integral over the unit sphere, in spherical coordinates described by the angles θ and φ:



[7] This is a consequence of time-reversal symmetry valid in the absence of a magnetic field.

We may choose the reference direction to, e.g., $\vec{u}_z$. Then $\hat{p} \cdot \vec{u}_z = \cos(\theta)$ and

$$A_n \equiv \int d\hat{p}\hat{p}^n = \frac{1}{4\pi}\int_0^{2\pi} d\varphi \int_0^{\pi} d\theta \sin(\theta)\cos^n(\theta). \tag{4.15}$$

We get $A_0 = 1$, $A_1 = 0$, $A_2 = 1/3$, $A_3 = 0$, etc. Thus we get from $\int d\hat{p}$(4.14)

$$\left(\partial_t f_0 + \frac{1}{3} v \nabla \cdot \vec{\delta f}\right) = I_{\rm inel}[f]. \tag{4.16}$$

We can also multiply by $\hat{p}$ and then integrate, i.e., $\int d\hat{p}\hat{p}$(4.14), which yields

$$\frac{1}{3}\left(v\nabla f_0 + \partial_t \vec{\delta f}\right) = -\frac{1}{3\tau}\vec{\delta f} \Leftrightarrow \vec{\delta f} = -v\tau\nabla f_0 - \tau\partial_t \vec{\delta f}. \tag{4.17}$$

In the proper diffusive limit, we have to assume that the temporal change in the distribution functions are much slower than the time τ describing elastic scattering. In this case we may neglect the last term from Eq. (4.17). Combining this result with Eq. (4.16) we get

$$(\partial_t - D\nabla^2) f_0(\vec{r}, E, t) = I_{\rm inel}[f_0], \tag{4.18}$$

where $D = v^2\tau/3$ is the diffusion constant.[8] Note that for $I_{\rm inel}[f] = 0$, Eq. (4.18) has the form of a diffusion equation, hence the term "diffusive limit".

In exercise 4.2, you will calculate the nonequilibrium distribution function in a wire without inelastic scattering. This nonequilibrium form, together with the crossover to the "quasiequilibrium" limit (see below and exercise 4.3), was measured in 1997 by H. Pothier, *et al.* (see Fig. 4.2).

4.3.1 Currents in the diffusive limit

With the same prescription as in the derivation of Eq. (4.18), we may now integrate the current densities, Eqs. (4.7) and (4.8), over $\hat{p}$. Then we get

$$\vec{j}_C(\vec{r}, t) = -eN_0 \int dE \nu(E) D(E) \nabla f_0(\vec{r}, E, t) \tag{4.19}$$

and

$$\vec{j}_Q(\vec{r}, t) = -N_0 \int dE (E - \mu) \nu(E) D(E) \nabla f_0(\vec{r}, E, t). \tag{4.20}$$

The diffusion coefficient $D(E)$ can be (weakly) energy dependent due to the energy dependence of v, or the energy dependence of the scattering time τ. Typically such dependence is small, but it makes rise to the thermoelectric effects described below. For convenience, we also separated the density of states as $N(E) = N_0\nu(E)$, where N_0 is the density of states at the Fermi level and $\nu(E)$ describes the energy dependence of the density of states. In normal metals,

[8] The factor 3 comes from averaging in three dimensions. For a 2-dimensional system we would get the same equation, but with $D = v^2\tau/2$.

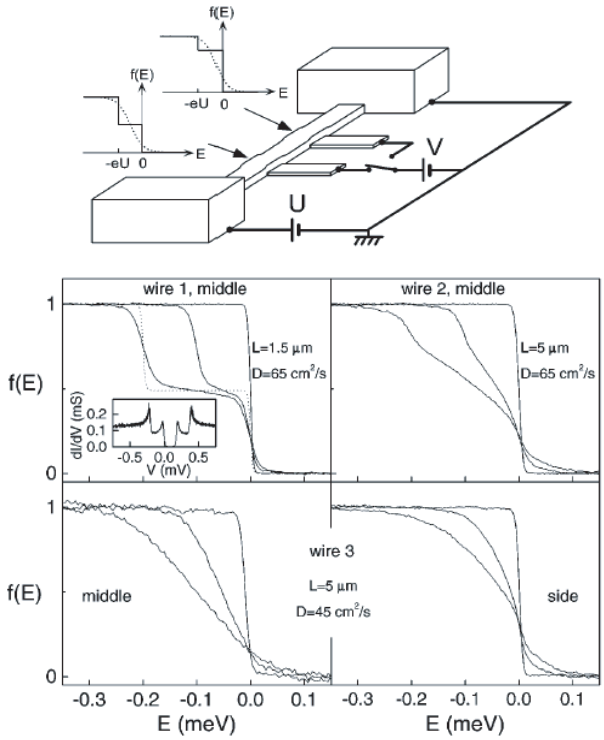


Figure 4.2: Top: Schematics of the measuring setup to probe the nonequilibrium distribution function inside a mesoscopic wire. The measurement is done with two superconducting probe tunnel junctions. In this case, the tunnelling current depends on the detailed form of the distribution function (analogous to Eq. (1.6)). Bottom: the measured distribution functions in wires of different length (i.e., different relative amount of inelastic scattering). Revised from H. Pothier, *et al.*, Phys. Rev. Lett. **79**, 3490 (1997).

$\nu(E) \approx 1$ is a very good approximation as long as ∇f_0 deviates from zero only near E_F. Deviations from this behavior will also contribute to the thermoelectric effects. In a superconductor, $\nu(E) = \mathrm{Re}[|E|/\sqrt{E^2 - \Delta^2}]$ (see Ch. 6), where Δ is the superconducting "energy gap".

Note again that Eq. (4.18) can be viewed as the continuity equation for the "spectral current" $D\nabla f_0(\vec{r}, E, t)$.

As below we concentrate on the diffusive limit where it is enough to follow the position dependence of the "s-wave" term $f_0(\vec{r}, E)$, we drop the superscript 0 unless the $\vec{p}$-dependence is explicitly required.

4.4 Inelastic scattering

The main sources of inelastic scattering are those that scatter two electrons from states with energy E and E' to those with energy $E+\hbar\omega$ and $E'-\hbar\omega$ (such that the total energy is conserved) and those that scatter an electron from state E to state $E+\hbar\omega$ and absorb a phonon with energy $\hbar\omega$ (or from $E \to E-\hbar\omega$ and emit a phonon with energy $\hbar\omega$).[9] These can be pictorially represented as in Fig. 4.3.

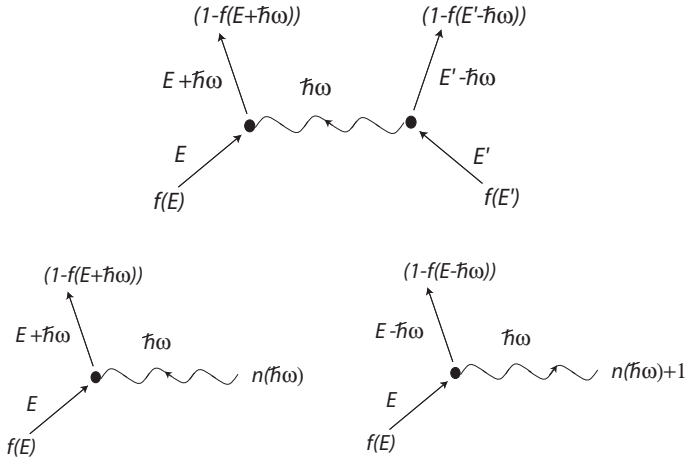


Figure 4.3: Top: electron-electron scattering. Bottom: electron-phonon scattering (left: absorption, right: emission).

The corresponding collision integrals are of the form (for simplicity, the arguments $\vec{r}$ and t are dropped from the distribution function)

$$I_{\mathrm{e-e}}[f] = -\int d\omega dE' K_{ee}(E,E',\omega)[f(E)f(E')(1-f(E+\hbar\omega))(1-f(E'-\hbar\omega)) - (1-f(E))(1-f(E'))f(E+\hbar\omega)f(E'-\hbar\omega)] \tag{4.21}$$

$$I_{\mathrm{e-ph}}[f] = -\int_0^\infty d\omega K_{eph}(E,\omega)\Big\{ [f(E)(1-f(E-\hbar\omega)) - f(E+\hbar\omega)(1-f(E))] \times [n(\hbar\omega)+1] - [f(E-\hbar\omega)(1-f(E)) - f(E)(1-f(E+\hbar\omega))]\, n(\hbar\omega)\Big\}. \tag{4.22}$$

[9] At very low temperatures, a further inelastic scattering effect becomes relevant: that between the electrons and the photons of the electromagnetic environment, see D. R. Schmidt, R. J. Schoelkopf and A. N. Cleland, Phys. Rev. Lett. **93**, 045901 (2004). This effect is empirically well-known and often it can be efficiently reduced by filtering the wires that connect the sample to the measuring equipment. Therefore, the lower one aims in the temperature, the better the filtering has to be and typically this effect can be clearly seen at temperatures below some 100 mK.

Both collision integrals are written in the form "scattering out" minus "scattering in" from/to the state with energy E. The factor $n(|\hbar\omega|) + 1$ in the electron-phonon "scattering out" term comes from the Bose commutation relation. The minus sign in front is due to the fact that scattering out decreases the number of electrons at state E.

The terms $K_{ee}(E, E', \omega)$ and $K_{eph}(E, \omega)$ are called "scattering kernels", and they have to be computed separately for each type of scattering. Such a calculation is usually based on the Fermi golden rule, but its details are outside the scope of this course and only some of the results are presented below.

4.4.1 Relaxation time approach

For small deviations δf from equilibrium (or, as defined below, "quasiequilibrium"), one can describe the collision integral by the relaxation time approach. This is obtained by expanding the collision integral as

$$I_{\rm coll}[f^0 + \delta f] \approx \left(\frac{\partial I_{\rm coll}}{\partial f}\right)_{f=f^0} \delta f \equiv -\frac{1}{\tau}\delta f. \tag{4.23}$$

The equilibrium state f^0 is chosen such that the term $I_{\rm coll}[f^0] = 0$. The functional derivative $\partial I_{\rm coll}/\partial f$ defines a *relaxation time* τ, which in general is time dependent. The relaxation time, when it is well defined, is a useful quantity in characterizing the strength of the scattering effect (for example, the above treatment of elastic scattering was an example of a relaxation time approach, where the relaxation time is well defined).

4.4.2 Electron-electron scattering

There are several sources that contribute to electron-electron scattering.[10] In metallic mesoscopic wires, the most relevant are the direct (screened) Coulomb interaction between the electrons and the second-order effects from collisions between electrons and magnetic impurities.

In most cases, the kernel $K_{ee}(E, E', \omega)$ depends only on ω (see the examples below). In this case, one may show that for a Fermi function $f(\vec{r}, E, t) = f^0(E, \mu(r,t), T(r,t))$ with any potential $\mu(r,t)$ and temperature $T(r,t)$, $I_{\rm e-e}[f]$ vanishes. Thus, if the electron-electron scattering is sufficiently strong (i.e., the prefactor to the collision integral is large), the solution to the Boltzmann equation (4.18) tends towards a Fermi function with (in general) a position dependent electron temperature $T_e(x)$. This limit is can hence be called the "quasiequilibrium" limit (another term used for this limit is the "hot electron limit"). In the exercises, you will derive the "kinetic" equations for the potential $\mu(x)$ and temperature $T_e(x)$.

[10] For a detailed summary, see F. Pierre, PhD thesis, (Paris, 2001), http://www.pa.msu.edu/cmp/birge-group/pierre/ (in French).

Strength of electron-electron interaction (**)

In the case of direct interaction, the kernel is of the form[11]

$$K_{ee}(E, E', \omega) = \kappa_d \omega^{-\alpha}, \tag{4.24}$$

where $\alpha = 2 - d/2$ for a d-dimensional wire.[12] In the one-dimensional case, the prefactor $\kappa_1 = 1/(\pi\sqrt{2D}\hbar^2 N_0 A)$, where A is the cross section of the wire.

For the case of magnetic impurities, one has

$$K_{ee}(E, E', \omega) = \kappa_m \omega^{-2}, \tag{4.25}$$

with $\kappa_m = \frac{\pi}{2}\frac{c_m}{\hbar N_0} S(S+1)\left[\ln\left(\frac{eV}{k_B T_K}\right)\right]^{-4}$. Here c_m is the concentration, S is the spin and T_K is the "Kondo temperature" of the impurities.

Relaxation time and length? (**)

For $d = 3$, the bare electron-electron collision integral can be approximated by $-\delta f/\tau_{e-e}$, where $\tau_{e-e} = \zeta(3/2)(k_B T)^{3/2}/(k_F \ell_{\rm el}\sqrt{\hbar\tau} E_F)$ is the relaxation time.[13] For $d < 3$, a relaxation-time approach does not work as the functional integral $\partial I_{\rm coll}/\partial f$ has an infrared divergence. In this case, however, one can use dimensional analysis.[14] As a result, the electron-electron scattering is described by a length scale

$$\ell_{\rm e-e} = \sqrt{R_Q A \sigma \sqrt{2\hbar D} E^*}, \tag{4.26}$$

where $R_Q = h/(2e^2)$, $\sigma = e^2 N_0 D$ is the Drude conductivity, and E^* is an energy scale describing the deviation from equilibrium. With some typical values, $\sigma = 20(\mu\text{m}\Omega)^{-1}$, $D = 0.02$ m^2/s, $A = 100\text{nm} \times 50\text{nm}$, and $E^* = 100\mu$eV, we would get $\ell_{\rm e-e} \approx 24\mu$m.

Electron-electron scattering in the one-dimensional case has been actively studied in the recent years. At present, it seems that the scattering length is an order of magnitude lower than what the conventional theory depicted above predicts.[15] The reason for this discrepancy is not known.

4.4.3 Electron-phonon scattering

A typical assumption for electron-phonon scattering is that the phonon bath is sufficiently large, such that the electron-phonon scattering does not essentially perturb the phonons. Then we may set $n(E) = (\exp(E/(k_B T_b)) - 1)^{-1}$, where T_b is the temperature of the lattice (typically the temperature of the cryostat).

[11] Sections indicated with stars (**) are for the interested reader — their detailed knowledge is not required in the exams.

[12] In this case, the dimension is determined by comparing the dimensions of the sample to the scale $\xi_D \equiv \sqrt{\hbar D/\max(eV, k_B T)}$.

[13] See Sec. 10.8 in J. Rammer, "Quantum transport theory" (Perseus books, Reading, Massachusetts).

[14] See F. Giazotto, *et al.*, Rev. Mod. Phys. **78**, 217 (2006) for details.

[15] See B. Huard, *et al.*, Solid State Communications **131**, 599 (2004).

The collision integral vanishes when $f(\vec{r}, E, t) = f^0(E; \mu(\vec{r}), T_b)$, i.e., when the electron temperature equals that of the phonons. Thus, the phonons try to thermalize the electrons with the lattice.

Often when driving the samples with a large voltage, the electrons heat the phonons in the film more than the substrate can absorb. In this case, the temperature of the film phonons deviates from that of the substrate (or the phonon distribution function may become a strong nonequilibrium function). The final temperature of the electron gas has in this case to be calculated from a thermal balance, which balances the heat flows between the subsystems (electrons, film phonons and the substrate).

Due to the Bose factor in the collision integral, electron-phonon scattering is strongly dependent on the lattice temperature (see below). At low temperatures, it can often be neglected.

Relaxation time and dissipated power (**)

The kernel of the electron-phonon scattering, obtained from the Fermi golden rule,[16] is related to an "Eliashberg function" $\alpha^2F(\omega)$ by $K_{eph} = 2\pi\alpha^2F(\omega)$, where

$$\alpha^2F(\omega) = \frac{|M_0|^2}{4\pi^2\hbar^2 v_S^3 N_0}\omega^2, \tag{4.27}$$

where v_S is the speed of sound and $|M_0|^2$ is the prefactor of the matrix element for electron-phonon scattering. The latter is dependent on the microscopic details of the lattice, such as the form of the unit cell, and therefore it is often measured by relating it to certain observable quantities, such as the relaxation time or the power dissipated into the phonons at a certain temperature.

The power dissipated into the phonon system (i.e., the heat current times the volume Ω of the sample) equals in the quasiequilibrium limit ($f(\vec{r}, E, t) = f^0(E; \mu(\vec{r}), T_e(\vec{r}))$)

$$P_{eph} = \int dE E I_{\mathrm{e-ph}} = \Sigma\Omega(T_e^5 - T_b^5), \tag{4.28}$$

where $\Sigma = 12\zeta(5)k_B^5|M_0|^2/(\pi\hbar^5 v_S^3)$ and $\zeta(x)$ is a Zeta function, $\zeta(5) \approx 1.0369$. Typical values for Σ in metals are around 10^9 $\mathrm{Wm^{-3}K^{-5}}$.

From Eqs. (4.23) and Eq. (4.27) we obtain that at $E = E_F$

$$\frac{1}{\tau_{e-ph}} = 4\pi\int_0^\infty d\omega \frac{\alpha^2F(\omega)}{\sinh\left(\frac{\hbar\omega}{k_BT}\right)} = \alpha T^3, \tag{4.29}$$

where $\alpha = 7\zeta(3)|M_0|^2k_B^3/(2N_0) = 7\zeta(3)\Sigma/(24\zeta(5)k_B^2N_0)$ and $\zeta(3) \approx 1.2021$. With typical values at Cu, $\Sigma = 2\cdot 10^9\mathrm{WK^{-5}m^{-3}}$, $N_0 = 1.6\cdot 10^{47}\mathrm{J^{-1}m^{-3}}$, $\tau_{e-ph} = 44$ ns at T_e =1K. The corresponding scattering length is obtained from $\ell_{\mathrm{e-ph}} = \sqrt{D\tau_{e-ph}}$. With $D = 0.01$ $\mathrm{m^2/s}$, we would hence get $\ell_{\mathrm{e-ph}} = 21$ μm at

[16] See again Sec. 10.8 in J. Rammer, "Quantum transport theory" (Perseus books, Reading, Massachusetts).

T =1K and $\ell_{\mathrm{e-ph}} = 660$ μm at $T = 100$ mK. This shows why electron-phonon scattering can be neglected at low temperatures.

4.5 Junctions

Tunnel junctions or other types of contacts between materials with different properties (different mean free paths, etc.) are typically very small-scale objects, their width being of the order of the Fermi wavelength. These can hence not be directly treated with the Boltzmann equation. However, one may include them through an effective boundary condition. In the diffusive limit, this is nothing but

- A statement of (spectral) current conservation at the two sides of the junction:

$$D_L N_0^L A_L \nu_L(E) \nabla f(x = x_b^-) = D_R N_0^R A_R \nu_R(E) \nabla f(x = x_b^+), \qquad (4.30)$$

  where $D_{L/R}$, $N_0^{L/R}\nu_{L/R}(E)$, $A_{L/R}$ are the diffusion constant, density of states and cross sections of the wires at the left/right side of the junction.

- This current must equal to that through the junction. In the case of a tunnel junction, this is (see Eq. (1.6)),

$$e^2 D_L N_0^L A_L \nabla f(x = x_b^-) = \frac{1}{R_T} \nu_L(E) \nu_R(E+eV)(f(x = x_b^+) - f(x = x_b^-)). \qquad (4.31)$$

4.6 Thermoelectric effects

One of the direct applications of the Boltzmann-equation approach is the full treatment of thermoelectric effects, i.e., finding a relation between the electric and heat current to the applied voltage and temperature gradient,

$$I = GL(-\nabla\mu/e + \alpha\nabla T) \qquad (4.32a)$$

$$\dot{Q} = L(-\Pi G \nabla\mu/e + G_{\mathrm{th}} \nabla T). \qquad (4.32b)$$

Here the diagonal conductances G and G_{th} are the electrical and thermal conductances of the electron system. The length L of the wire is included in order to get the correct dimension for the coefficients. The off-diagonal components are generally called thermoelectric coefficients.

The coefficient α is called a *Seebeck* coefficient, and it describes a thermoelectric power: assume we have a two-terminal system and we apply a temperature gradient across it. This temperature gradient will make rise to a current $I = G\alpha\nabla T$. But apply now an opposite voltage V that exactly cancels this current. From Eq. (4.32) we see that this voltage has to be $V = \alpha\nabla T$.

The coefficient Π between the heat current $\dot{Q}$ and the voltage V is the *Peltier* coefficient. It indicates the amount of heat one can carry by applying a small voltage. As this effect is linear in voltage, the sign of the heat current follows that of the voltage, and as a result, one can cool one part of the system and heat another one. At larger voltages, however, Joule heating (quadratic in V) starts to heat the sample throughout.

Boltzmann theory allows us to calculate the various thermoelectric coefficients. Let us concentrate on the linear regime, i.e., that we apply only a small voltage or a small temperature gradient. Then the deviation from equilibrium is small and therefore necessarily $f(\vec{r},E) \approx f^0(E;\mu(\vec{r}),T(\vec{r}))$ and the strength of electron-electron scattering is irrelevant. In this case, $\nabla f(\vec{r},E) = \partial_\mu f^0 \nabla\mu + \partial_T f^0 \nabla T$. From Eqs. (4.19) and (4.20) multiplied by the cross section A of the wire we then get

$$G = \frac{A}{L} e^2 N_0 \int dE \nu(E) D(E) \partial_\mu f^0(E,\mu,T) \tag{4.33a}$$

$$G\alpha = \frac{A}{L} e N_0 \int dE \nu(E) D(E) \partial_T f^0(E,\mu,T) \tag{4.33b}$$

$$G\Pi = \frac{A}{L} e N_0 \int dE (E-\mu) \nu(E) D(E) \partial_\mu f^0(E,\mu,T) \tag{4.33c}$$

$$G_{\rm th} = \frac{A}{L} N_0 \int dE (E-\mu) \nu(E) D(E) \partial_T f^0(E,\mu,T). \tag{4.33d}$$

In a normal metal, the density of states and the diffusion constant are typically almost energy independent. We may then express them in the form $\nu(E) \approx 1 + c_N(E-\mu)$ and $D(E) \approx D_0 + c_D(E-\mu)$ and retain only the lowest relevant order in the coefficients c_N and c_D.

For the diagonal coefficients we may neglect the small coefficients c_N and c_D and obtain for the electrical conductance

$$G = e^2 N_0 D A/L, \tag{4.34}$$

i.e., Einstein-Drude conductance of a disordered wire.[17] The thermal conductance is

$$G_{\rm th} = \frac{A}{L}\frac{\pi^2}{3} k_B^2 N_0 D T = \mathcal{L} G T, \tag{4.35}$$

where $\mathcal{L} = \pi^2 k_B^2/(3e^2)$ is the *Lorenz number*. This relation between the thermal and electric conductances is called the *Wiedemann-Franz relation*.

The off-diagonal components rely on the small coefficients c_D and c_N. They are

$$\alpha = e\mathcal{L} G' T/G, \tag{4.36}$$

where $G' \equiv e^2(c_D N_0 + D_0 c_N) A/(L E_F)$. This is the *Mott law*. Finally, the Peltier coefficient satisfies the *Onsager relation* $\Pi = \alpha T$. The appearance of the coefficients c_D and c_N breaks the particle-hole symmetry between the energies

[17] This was the first time we actually derived it in the course.

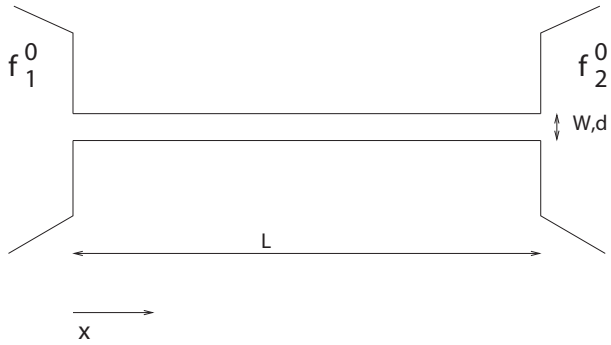
above and below the Fermi level and therefore the thermoelectric effects are sometimes used to justify (or un-justify) theoretical predictions for the electronic band structure in metals. However, from these equations one can see that the thermoelectric coefficients are proportional to the factor k_BT/E_F, and are therefore very small at sub-Kelvin temperatures.

4.7 Problems

In each exercise, you are requested to i) solve for the distribution function in a wire between two reservoirs (where the distribution function has the Fermi function form with chemical potentials μ_1 and μ_2 and temperatures T_1 and T_2) in a stationary limit (time derivative of the distribution function equal to zero) and to ii) compute the current(s) flowing in the sample due to the applied voltage $eV = \mu_2 - \mu_1$ and/or the temperature difference $T_2 - T_1$. This means that we may assume the reservoirs as boundary conditions to the Boltzmann equation. You may assume a constant density of states ($\nu(E) = 1$) and a diffusion constant $D(E) = D = \text{const.}$ throughout.

Note that even if you could not solve the previous problem, you may try to solve the next one.

Each problem considers an essentially quasi-one-dimensional wire (see figure). In the diffusive limit (exercises 2-3), the distribution function changes only along the wire in, say, x-direction, and there are no changes in the perpendicular directions. The ballistic limit has a slightly different ideology and you may try to solve it after the other parts.



1. Consider a ballistic wire (no scattering) of width W, thickness d and length L between two reservoirs (see figure). Find the stationary-state distribution function $f(\vec{r}, \hat{p}; E)$ of electrons inside the wire. Hint: define "left-movers" and "right-movers".

2. Consider a similar situation with a diffusive wire, but no inelastic scattering: $l_{el} \ll L \ll l_{ee}, l_{eph}$. a) Solve the diffusive-limit Boltzmann equation with boundary conditions given as in the figure to find out the stationary (time independent) distribution function $f(x, E)$. Then, b) calculate the charge current flowing in the system for arbitrary potentials μ_1 and μ_2 and temperatures

T_1 and T_2. Finally, c) calculate the heat current flowing in the structure for $\mu_1 = \mu_2 = \mu$ and temperatures T_1 and T_2. Hint: use the formulae given below.

3. Now assume $L \gg l_{ee}$ but $L \ll l_{eph}$ and otherwise the same system. In this case, the space dependence of the distribution function can be described through the form $f(x, E) = f^0(E; \mu(x), T(x))$, where $f^0(E; \mu, T)$ is a Fermi function with potential $\mu(x)$ and temperature $T(x)$. a) Plugging this into the diffusive-limit Boltzmann equation, find the kinetic equations for $\mu(x)$ and $T(x)$. Hint: you can get rid of the energy dependence by integrating over the energies in the same way as one integrated over the momentum directions in the derivation of the diffusive limit - use the formulae given below. b) Once you have the kinetic equations, find the potential/temperature profile $\mu(x)$ and $T(x)$ and the corresponding currents (again for the heat current, it is enough to treat only the case $\mu_1 = \mu_2 = \mu$).

The case of a macroscopic wire, $L \gg l_{eph}$ reduces just to finding the potential profile for a given voltage. In this case, temperature difference would require maintaining a temperature difference also for the phonons and would then go as the quasiequilibrium limit.

Hint: You may have use for the following formulae:

$$\int_{-\infty}^{\infty} (f^0(E; \mu_1, T_1) - f^0(E; \mu_2, T_2))dE = \mu_1 - \mu_2 \tag{4.37}$$

$$\int_{-\infty}^{\infty} E(f^0(E; \mu_1, T_1) - f^0(E; \mu_2, T_2))dE = \frac{\pi^2}{6} k_B^2 (T_1^2 - T_2^2) + \frac{1}{2}(\mu_1^2 - \mu_2^2) \tag{4.38}$$

$$\int_{-\infty}^{\infty} dE \partial_\mu f^0(E; \mu, T) = 1 \tag{4.39}$$

$$\int_{-\infty}^{\infty} dE \partial_T f^0(E; \mu, T) = 0 \tag{4.40}$$

$$\int_{-\infty}^{\infty} dE E \partial_\mu f^0(E; \mu, T) = \mu \tag{4.41}$$

$$\int_{-\infty}^{\infty} dE E \partial_T f^0(E; \mu, T) = \frac{k_B^2 T \pi^2}{3} \tag{4.42}$$

Here $f^0(E; \mu, T)$ are Fermi functions.

Chapter 5

Single-electron tunneling and Coulomb blockade

When a junction between two conductors is made very small, the energy scale required to transfer a single electron through it becomes observable. If two junctions are placed in series such that there is a conducting island between them, or a single junction is placed in high-resistance environment, this *charging energy* E_C limits the current transport through the system, such that at voltages and temperatures lower than E_C/e, no current can flow. This is the regime of *Coulomb blockade*.[1]

Coulomb blockade shows up in various types of small structures. In a *single electron transistor* the island is typically made of a metal with a high density of states (average level spacing much less than E_C, k_BT or eV), and the charge transport is controlled by an extra *gate voltage* coupled capacitively to the junction. For certain values of gate voltages, Coulomb blockade is lifted and current can flow. This is why the system can be used as a transistor and it is in fact the most sensitive detector of charge. In semiconductor quantum dots and molecular structures, the typical level spacing $\Delta\epsilon$ is large (up to the order of 1 eV in molecules). In this case, the transport properties sensitively depend on both energy scales, E_c and $\Delta\epsilon$. These systems are described in this chapter.

In Exercise 1.1, you estimated the charging energy for a typical mesoscopic tunnel barrier. With $C = 0.8910^{-15}$F$= 0.89$fF, the charging energy $E_C = 86\mu$eV$= 1$K$/k_B$.

5.1 Charging energy

Consider a capacitor with capacitance C. Assume we apply a time-dependent voltage of the form $V = V_0(1 - e^{-t/\tau})$ over the capacitor. After time $t \gg \tau$, the

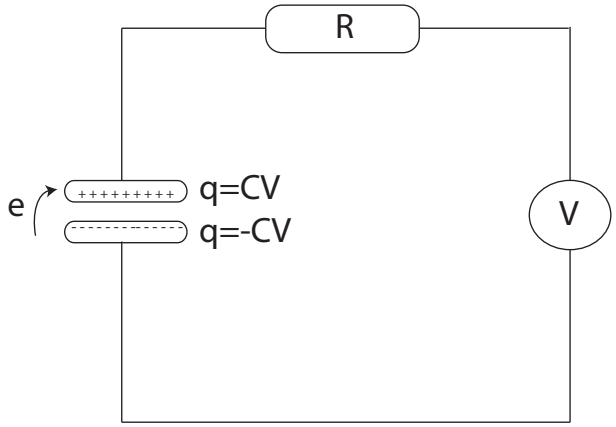
[1] This chapter is partially based on Ch. 3 (by Gerd Schön) in the book T. Dittrich, *et al.*, "Quantum transport and dissipation", (Wiley-VCH, 1998). This chapter can also be downloaded at http://www-tfp.physik.uni-karlsruhe.de/Publications/Pub1998/setrev4.ps.

voltage has become almost constant, $V \approx V_0$, and the charge on the capacitor has changed by $q = \int_0^\infty C\dot{V}dt = V_0 C$. The work done by the voltage is equal to

$$W = \int_0^\infty VI = \int_0^t V\frac{dq}{dt} = \frac{1}{2}V_0^2 C. \tag{5.1}$$

Thus, charging the capacitance by the charge q requires the energy $q^2/(2C)$.

Now let us consider an event where an electron tunnels through a barrier with capacitance C (with a constant voltage bias):



Before tunnelling, the charge across the capacitor is $q = CV$ and the electrostatic energy stored on the capacitor is $E_i = q^2/(2C)$. After an electron has tunnelled through, the charge is $q - e$ and the energy of the final state is

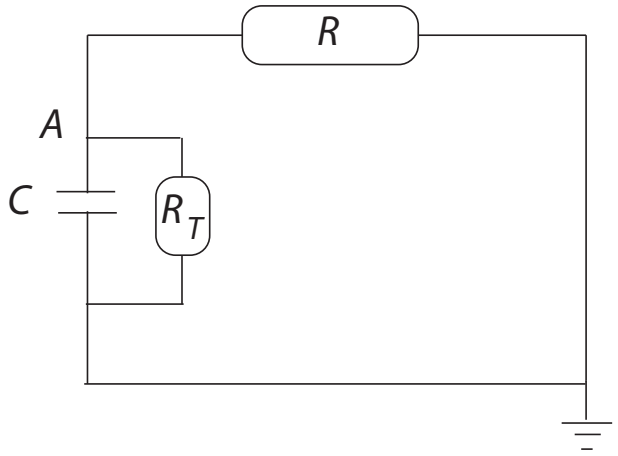
$$E_f = \frac{(q-e)^2}{2C} = \frac{q^2}{2C} - \frac{e}{C}q + \frac{e^2}{2C} = E_i - eV + E_C. \tag{5.2}$$

At a vanishing temperature, tunnelling thus seems to be possible provided that $eV > E_C$.

However, in the case of a single junction, the quantum fluctuations in the environment of the junction may enable the electron tunnelling even at lower voltages. An estimate of this effect is provided by the uncertainty principle: Let us denote the time it takes for the electron charge to relax into the environment of the junction by τ. Within this time, the environment is in a nonequilibrium state, and tunnelling may take place provided that the energy uncertainty $\Delta E = \hbar/(2\tau)$ is greater than E_C, i.e.,

$$E_c\tau \lesssim \hbar/2. \tag{5.3}$$

But what is the time τ? Consider the circuit below:



Here R_T is the resistance of the junction, and R the resistance of the environment. It is easy to derive an equation for the charge q on the capacitor by requiring that the voltage drop across the resistors and the capacitor is the same. One obtains

$$\frac{dq}{dt} = -\frac{q}{R_{||}C}, \tag{5.4}$$

where $R_{||} = (1/R + 1/R_T)^{-1}$. The solution to this is $q(t) = q_0 e^{-t/(R_{||}C)}$, from which we can identify the time $\tau = R_{||}C$. Now substituting this to Eq. (5.3), we obtain the condition for Coulomb blockade (i.e., the inverse of the condition in Eq. (5.3) that stated when tunnelling is made possible by uncertainty):

$$E_c\tau = \frac{e^2}{2C}R_{||}C = e^2 R_{||} > \hbar/2. \tag{5.5}$$

Thus, this imposes two conditions for the observation of Coulomb blockade in a single junction:[2]

- The resistance R_T of the tunnel barrier must exceed the quantum resistance $h/(2e^2) = 12.8\text{k}\Omega$.
- The environment of the junction has to have a resistance that exceeds $h/(2e^2)$.

The current suppression in the case of a single junction is called *dynamical Coulomb blockade* or *environmental Coulomb blockade*, and it is discussed more later in the course, if time allows.

5.1.1 Single-electron box

Now consider the system depicted in Fig. 5.1. It shows a metal island with a tunnel contact (the circuit element left from the island denotes a normal tunnel

[2]Never mind the 2π as the argument is rough anyhow.

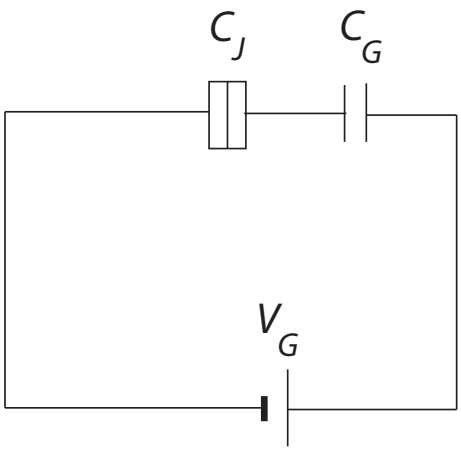


Figure 5.1: Single-electron box.

junction), and a capacitive contact to a gate voltage V_G. For $V_G = 0$, the ground state of the island is neutral, i.e., the number of electrons on the island equals the number of positive ions. The higher-energy states may involve $n = \pm 1, \pm 2, \ldots$ extra (deficit) electrons on the island. If V_G is turned on, the ground state of the system changes in discrete steps from $n = 0$ to another integer value of n. The system reacts to this change of ground state by letting the extra electrons tunnel through the tunnel contact.

The gate voltage induces continuous displacement charges Q_1 and Q_2 on the capacitors. This is due to the spatial distribution of charge on the island - the total charge $Q_1 + Q_2 = -ne$ still remains an integer multiple of the electron charge. The charging energy of the system, $E_{ch} = Q_1^2/(2C_J) + Q_2^2/(2C_G)$ is now obtained in terms of this n and V_G by noting that $V_G = Q_1/C_1 - Q_2/C_G$, and eliminating Q_1 and Q_2 in favor of n and V_G. It turns out from similar considerations as above, that the relevant (grand canonical) free energy has to involve the work $-V_G Q_2 = \mu_G N_2$ done by V_G in polarizing the capacitor C_G. Then this energy can be written as

$$E_{ch}(n, Q_G) = \frac{(ne - Q_G)^2}{2C}, \tag{5.6}$$

plus some terms independent of the index n. Here $C = C_J + C_G$ is the total capacitance of the island. The effect of the gate voltage is to define a continuous offset charge $Q_G = C_G V_G$. This is often denoted as the *gate charge*.

The charging energy for different values of n as a function of Q_G is a series of parabola. The single-electron box has degeneracy points whenever $Q_G/e = n + 1/2$. At these points, the island may change its charge state. At a low temperature, if one now adiabatically changes the gate V_G, the charge n on the island stays constant until the next degeneracy point is reached, at which point an extra electron tunnels in or out of the island, depending on the direction of

the gate charge change.

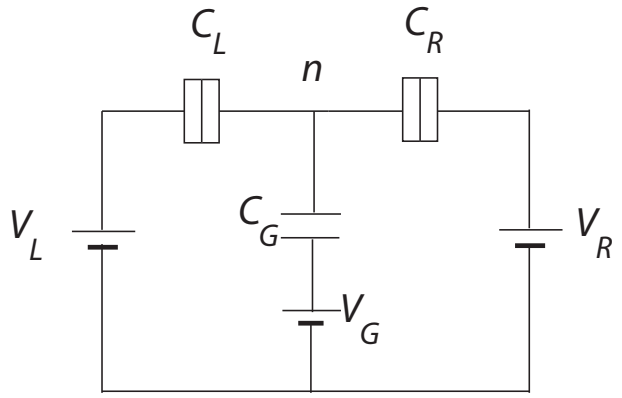


Figure 5.2: Single-electron transistor.

5.1.2 Single-electron transistor (SET)

The most basic and useful example system exhibiting single-electron effects is a single-electron transistor (SET) shown in Fig. 5.2. Here we allow tunnelling through both contacts to the island, such that we may have a net current flowing through it. The current is controlled by the transport or bias voltage $V_L - V_R$, and the state of the island can be controlled by the gate voltage. Similar considerations as for the single-electron box (see also exercises) now show that we again get $E_{ch}(n, Q_G) = (ne - Q_G)^2/C$ with $C = C_L + C_R + C_G$. In this case, the gate charge consists of contributions from all voltage sources. However, in the limit where $C_G \ll C_L, C_R$, we may neglect the effect from the bias voltages $V_{L/R}$, and only get $Q_G = C_G V_G$. As there should be no tunnelling through the capacitance C_G, this limit is the most relevant one in the experiments.

Assume now that we apply a bias voltage such that the potential on the left is higher than on the right. In this case, at $T = 0$ tunnelling from the left electrode (i.e., $n \to n + 1$) to the island is possible if

$$eV_L > E_{ch}(n+1, Q_G) - E_{ch}(n, Q_G) = \left(n + \frac{1}{2} - \frac{Q_G}{e}\right) e^2/C. \tag{5.7}$$

Similarly tunneling from the island to the right lead is possible if

$$eV_R < E_{ch}(n+1, Q_G) - E_{ch}(n, Q_G). \tag{5.8}$$

Both conditions have to be satisfied simultaneously in order for the current to flow.

Let us consider a simple example where we have only the lowest charge state occupied in equilibrium and where we have a symmetric bias, $V_L = -V_R = V/2$. Then for a given Q_G, we have to apply a bias $V > (e - 2Q_G)/C$ in order to

lift the Coulomb blockade. Thus, changing the gate voltage, the current of the island at some finite bias oscillates between almost zero ($Q_G < (e - V/C)/2$) and some finite value depending on the tunnelling resistances. These *Coulomb oscillations* can thus be understood from the stability diagram of the SET, but the quantitative calculation of the current requires the use of *orthodox theory* outlined in the next two sections.

5.2 Tunnel Hamiltonian and tunnelling rates

The orthodox theory of single-electron tunnelling was constructed by D. Averin and K. Likharev in the 1980's. It is based on calculating the tunnelling rates through single junctions that are part of a single-electron transistor or a multi-island system, using the Fermi golden rule, i.e., the lowest-order perturbation theory (this is called a *sequential tunnelling approximation*). With the knowledge of these rates, one constructs a master equation that describes the probabilities of the charge states, and from the solution of this master equation one then obtains the observables, such as the current through the system.

The analysis starts from the Hamiltonian. For a SET it can be written as

$$H = H_L + H_R + H_I + H_{\rm ch} + H_t. \tag{5.9}$$

Here the first three describe the noninteracting electrons in the left and right electrode and the island.[3] They can thus be written in the second-quantized notation as

$$H_L = \sum_{k,\sigma} \epsilon_k c^\dagger_{k,\sigma} c_{k,\sigma}, \tag{5.10}$$

where $c_{k\sigma}$ annihilates an electron in state k, σ in the left electrode. The Hamiltonian H_R for the right electrode is analogous, and that for the island reads

$$H_I = \sum_{q,\sigma} \epsilon_q b^\dagger_{q,\sigma} b_{q,\sigma}, \tag{5.11}$$

$b_{q,\sigma}$ being the annihilation operator for the electrons in the island. The charging energy term is assumed to depend only on the total charge on the island,

$$H_{\rm ch} = \frac{(\hat{n}e - Q_G)^2}{2C}, \tag{5.12}$$

where $\hat{n} = \sum_{q,\sigma} b^\dagger_{q,\sigma} b_{q,\sigma} - N_+$ is the number operator of the island. We subtracted the number N_+ of the positive ions for simplicity. Finally, the tunnelling Hamiltonian $H_t = H_{t,L} + H_{t,R}$ describes a process where an electron is annihilated in one of the electrodes and created in the island, and vice versa,[4]

$$H_{t,L} = \sum_{k,q,\sigma} T_{k,q} b^\dagger_{q,\sigma} c_{k,\sigma} + \text{h.c.} \tag{5.13}$$

[3] In fact, there the interactions are described by a mean field model, such that they simply alter the dispersion relation ϵ_k, which is not necessary of the free-electron form.

[4] The letters h.c. denote the Hermitian conjugate of the previous expression. Here they are for the "vice versa".

A similar Hamiltonian can be written for the contact to the right island.

In Fermi golden rule, the transition rate between an initial state $|i\rangle$ and the final state $|f\rangle$ can be found from

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar}\langle|\langle f|H|i\rangle|^2\delta(E_i - E_f)\rangle. \tag{5.14}$$

The outer brackets denote the statistical averaging. Here the different states are the different charge states on the island. We assume that the electrodes quickly relax into their equilibrium charge state and hence do not consider the differences in the electrode states. Essentially, this means that we assume the system in a low-resistance environment, neglecting the dynamical Coulomb blockade effects.

The only part of the Hamiltonian that changes the charge states (i.e., whose matrix element between different charge states is finite) is H_t. For example, the rate of tunnelling between the state k in the left electrode and the state q in the island, summed over k and q, is

$$\Gamma_{LI}(n) = \frac{1}{e^2R_L}\int_{-\infty}^{\infty} d\epsilon_k \int_{-\infty}^{\infty} d\epsilon_q f_L(\epsilon_k)[1 - f_I(\epsilon_q)]\delta(\delta E_{ch} + \epsilon_q - \epsilon_k). \tag{5.15}$$

Note that the energy conservation includes the change $\delta E_{ch} = \left(n + \frac{1}{2} - \frac{Q_G}{e}\right) - eV_L$ in the charging energy of the island. Above, we also introduced the tunnel conductance $R_L^{-1} = 4\pi e^2 N_I(0)N_L(0)\Omega_I\Omega_L|T|^2/\hbar$, whose expression is analogous to that used in Ch. 1. This assumes that the transmission probability $|T_{k,q}|^2 = |T|^2$ through the junction is independent of the channel indices k, q. In addition, the tunnel conductance depends on the volumes $\Omega_{I/L}$ of the island and the electrode.

Similar considerations would apply for the reverse process $\Gamma_{IL}(n+1)$, and for the other tunnel barrier.

Now assume the electron-phonon relaxation time is short, such that the electrodes and the island can be described by Fermi functions with the lattice temperature for each potential. In this case, we can use the identities

$$\int_{-\infty}^{\infty} dE[f^0(E) - f^0(E+x)] = x \tag{5.16}$$

and

$$f^0(E)(1 - f^0(E+x)) = \frac{f^0(E) - f^0(E+x)}{1 - e^{-x/(k_BT)}}, \tag{5.17}$$

valid for Fermi functions described by temperature T. The resulting single-electron tunnelling rate is

$$\Gamma_{LI}(n) = \frac{1}{e^2R_L}\frac{\delta E_{ch}}{\exp(\delta E_{ch}/(k_BT)) - 1}. \tag{5.18}$$

At $T \to 0$, for $\delta E_{ch} > 0$ the tunnelling would be suppressed. This corresponds to the Coulomb blockade situation depicted in the previous section. For $\delta E_{ch} < 0$,

we would then get

$$\Gamma_{LI}(n) \to \frac{1}{e^2 R_L}|\delta E_{ch}|. \tag{5.19}$$

At finite temperatures all processes are allowed, but their rate depends strongly on the ratio $\delta E_{ch}/k_B T$.

5.3 Master equation

In the lowest-order theory in the tunnel coupling, the charge state of the island can change at most by one electron in a single process. The knowledge of the tunnelling rates allows us to describe the time-dependence of the probability $P(n,t)$ for the system being in a given charge state n at time t. This probability satisfies the master equation

$$\begin{aligned}\frac{dP(n,t)}{dt} &= -[\Gamma_{LI}(n)+\Gamma_{IL}(n)+\Gamma_{RI}(n)+\Gamma_{IR}(n)]P(n,t)\\ &+[\Gamma_{LI}(n-1)+\Gamma_{RI}(n-1)]P(n-1,t)\\ &+[\Gamma_{IL}(n+1)+\Gamma_{IR}(n+1)]P(n+1,t).\end{aligned} \tag{5.20}$$

In other words, the rate of change of the probability $P(n,t)$ is equal to the difference of rates for tunnelling into state n (lower line) and the tunnelling out of state n (upper line).

We can also determine the current from the rates and the probabilities. It is

$$I_L(t) = -e\sum_n [\Gamma_{LI}(n)-\Gamma_{IL}(n)]P(n). \tag{5.21}$$

In the stationary state it is straightforward to show that $I_L = I_R$.

Let us consider an example. To specify the system, assume a symmetric bias, and Q_G between ne and $(n+1)e$. At low temperatures and bias voltages, only the charge states n and $n+1$ are occupied and hence contribute to the current. Then the stationary state occupation probability (i.e., where the left-hand side of Eq. (5.20) vanishes) satisfies

$$[\Gamma_{LI}(n)+\Gamma_{RI}(n)]P(n) = [\Gamma_{IL}(n+1)+\Gamma_{IR}(n+1)]P(n+1). \tag{5.22}$$

Here only the transitions between the charge states n and $n+1$ were taken into account. The energy changes $\pm\delta E^L_{ch}$ determining the rates $\Gamma_{LI}(n)$ and $\Gamma_{IL}(n+1)$ are

$$\pm\delta E^L_{ch} = \pm[(n+\frac{1}{2}-\frac{Q_G}{e})\frac{e^2}{C}-\frac{eV}{2}], \tag{5.23}$$

upper sign denoting the transition from the electrode to the island. The corresponding energy change δE^R_{ch} in the right junction is obtained by replacing eV by $-eV$.

Using the the fact that $\sum_n P(n) = P(n)+P(n+1) = 1$ we then get

$$P(n) = \frac{\Gamma_{IL}(n+1)+\Gamma_{IR}(n+1)}{\Gamma}, \tag{5.24}$$

where $\Gamma = \Gamma_{IL}(n+1) + \Gamma_{LI}(n) + \Gamma_{IR}(n+1) + \Gamma_{RI}(n)$ is the total rate. Now the current is given by

$$\begin{aligned} I &= -e[\Gamma_{LI}(n)P(n) - \Gamma_{IL}(n+1)P(n+1)] \\ &= -e\frac{\Gamma_{LI}(n)\Gamma_{IR}(n-1) - \Gamma_{RI}(n)\Gamma_{IL}(n+1)}{\Gamma}. \end{aligned} \tag{5.25}$$

This can be interpreted as the difference of the rates for going into the island from the left and then out to the right and the opposite process. At low temperatures only one of these directions is possible. This equation can be readily analyzed by inspecting Eq. (5.23). At low temperatures the tunnelling through the left junction is allowed when $Q_G - (n+1/2)e \geq -VC/2$. The rate for tunnelling into the right is finite when $Q_G - (n+1/2)e \leq VC/2$. Thus, both processes are present when $|Q_G - (n+1/2)e| \leq |V|C/2$ with $V > 0$. The opposite processes would be allowed in a similar gate charge regime when $V < 0$. The total current in a symmetric setup $R_L = R_R = R_T$ is then

$$I = \frac{1}{4R_T}\left[V - \frac{4e^2}{C^2V}\left(\frac{Q_G}{e} - n - \frac{1}{2}\right)^2\right]\theta\left(|Q_G - (n+1/2)e| - \frac{1}{2}C|V|\right). \tag{5.26}$$

The Heaviside function $\theta(x) = 1$ for $x < 0$ and 0 otherwise specifies the transport window.

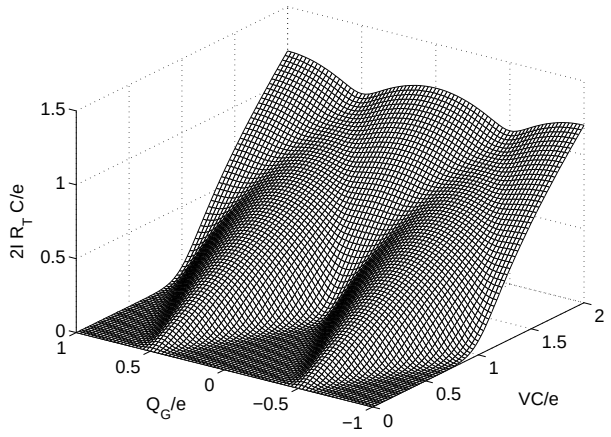


Figure 5.3: Current vs bias voltage and gate charge through a symmetric single-electron transistor in the sequential tunnelling approximation at $T = 0.05E_C/k_B$. Calculated with a code written by Matti Laakso.

Figures 5.3 and 5.4 show the current as a function of the transport voltage and the gate charge. As a function of the latter, the current shows the e-periodic Coulomb oscillations. The points where $Q_G = n + 1/2$ corresponds to a degeneracy point, where even at $V = 0$ two charge states can coexist, and thus a finite current is possible even for a very small bias.

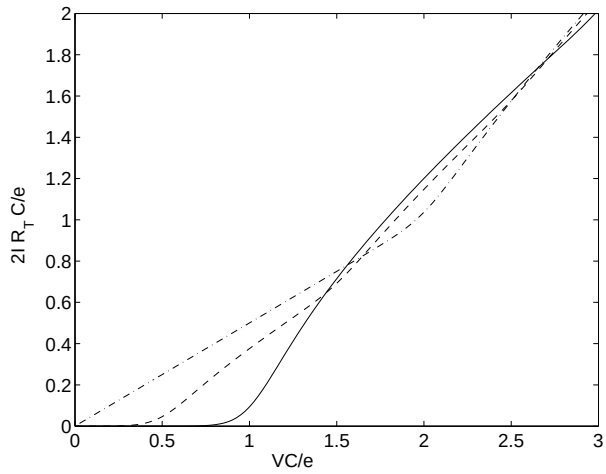


Figure 5.4: Current-voltage characteristic of a symmetric SET for $Q_G = 0$ (solid line, Coulomb blockade), $Q_G = 0.25$ (dashed line, intermediate) and $Q_G = 0.5$ (dash-dotted line, linear $I - V$ due to the degeneracy of $n = 0$ and $n = 1$). Broadening of the edges is due to the finite temperature $T = 0.05E_C/k_B$. Calculated with a code written by Matti Laakso.

Due to the sensitivity of the current on the value of the gate charge, the single-electron transistor can be and has been used used as a very accurate electrometer. For example, the first idea to measure the state of a charge qubit was to employ a SET, but it has also been used for example in measuring charge distributions in semiconductor quantum Hall devices.

5.4 Higher-order effects

The orthodox theory outlined above works mostly in the regime where the tunnelling can be treated with the lowest-order perturbation theory. This is the regime of sequential tunnelling, i.e., an electron tunnels into the island in a single process, and another process tunnels it out. There is no coherence between these processes. The higher-order effects, such as the coherent "cotunnelling" through several junctions, are especially important in the Coulomb-blockade regime, where they make rise to a finite current.

The second-order tunnelling rates can be calculated from the standard second-order theory,

$$\Gamma_{i\to f} = \frac{2\pi}{\hbar}\left|\sum_{\rm int}\frac{\langle i|H_t|{\rm int}\rangle\langle {\rm int}|H_t|f\rangle}{E_{\rm int}-E_i}\right|^2 \delta(E_i - E_f). \qquad (5.27)$$

Here the some goes over all intermediate charge states $|{\rm int}\rangle$. Cotunnelling process thus corresponds to an electron tunnelling from the first electrode into a

virtual charge state (above the Fermi level) and from there to the second electrode. As the rate is inversely proportional to the distance $E_{\rm int} - E_i$ between the virtual state energy and the initial state energy, the process is the weaker the further the closest allowed charge states are. But this is not as strong suppression as in the first-order theory where the rate is exponentially suppressed ($\sim \exp(-(E_{\rm int} - E_i)/(k_B T))$).

In a single-electron transistor, the level spacing is small compared to the charging energy scale. Hence, the electron tunnelling out of the island is not the same as the electron that tunnelled into the island in the same coherent process with a very high probability. Hence, after the process an electron-hole excitation is left on the island. This process is hence often called "inelastic cotunnelling".

Another thing one has to take into account is that two different types of events have to be added coherently: for example, in tunnelling from left to right, either first an electron tunnels into the island from the left electrode, and then another electron tunnels into the right electrode, or first an electron tunnels out of the island, and only then the first electron tunnels in. In the previous case, the energy difference between the initial and the virtual states is $\delta E_L = E_{ch}(n+1, Q_G) - E_c h(n, Q_G) - eV_L$, whereas in the latter case it is $\delta E_R = E_{ch}(n-1, Q_G) + eV_R - E_c h(n, Q_G)$. These processes have to be added coherently, i.e., before taking the square of the matrix element.

The resulting rate for the inelastic cotunnelling in a normal-metal single-electron transistor is

$$\Gamma_{\rm cot} = \frac{1}{\pi^2 h} \frac{R_Q^2}{R_L R_R} \int_{k\in L, q, q' \in I, k' \in R} dE_k dE_q dE_{q'} dE_{k'} f(E_k)[1-f(E_q)]f(E_{q'})[1-f(E_{k'})] \times \left[\frac{1}{E_q + \delta E_L - E_k} + \frac{1}{E_{k'} + \delta E_R - E'_q} \right]^2 \delta(eV + E_k - E_q + E_{q'} - E_{k'}). \tag{5.28}$$

At $T = 0$, $eV \ll \delta E_L, \delta E_R$ the integrals can be performed analytically, with the result

$$\Gamma_{\rm cot} = \frac{e^3}{6\pi^2 h} \frac{R_Q^2}{R_L R_R} \left(\frac{1}{\delta E_L} + \frac{1}{\delta E_R} \right)^2 V^3. \tag{5.29}$$

Hence, we see that the rate for cotunnelling is proportional to $(R_Q/R_L)^2$. Cotunnelling is thus the more important the lower the junction resistance is. In systems with N junctions in series, the relevant higher-order effects can be found from the N'th order perturbation theory and in general they lead to a current-voltage dependence $I \propto V^{2N-1}$. These cotunnelling processes limit the accuracy of the single-electron transistor and other single-electron devices, such as the Coulomb blockade thermometer discussed below.

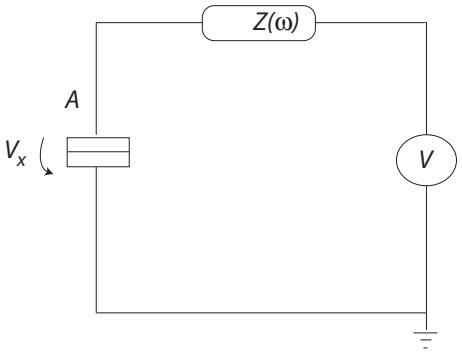


Figure 5.5: Tunnel junction in an environment characterized by the impedance $Z(\omega)$.

5.5 Dynamical Coulomb blockade

In Sec. 5.1, we showed that the speed of charge relaxation is important for the occurrence of Coulomb blockade in a single tunnel junction. This is the phenomenon of *dynamical* or *environmental Coulomb blockade*.[5]

Let us now relax the assumption that the charge on the reservoirs relaxes quickly. This is generally due to the fact that the junction resides in an environment characterized by an impedance $Z(\omega)$ (Fig. 5.5). The presence of $Z(\omega)$ makes rise to two effects: first, the average voltage drop at the junction depends on the current,

$$V_x = V - I(V_x)Z(\omega). \tag{5.30}$$

But since the current may depend on the voltage drop in a nontrivial manner, this problem has to be solved self-consistently. In addition to this, the resistor makes rise to voltage fluctuations, δV_x over the junction. These fluctuations are discussed more in Ch. 8, but for now it is enough to know that their correlation function is in macroscopic systems described by the *fluctuation-dissipation theorem*, and it is written as

$$\langle \delta\hat{V}_x(\omega)\delta\hat{V}_x(-\omega)\rangle = \hbar\omega \mathrm{Re}[Z_t(\omega)]\left(\coth\left(\frac{\hbar\omega}{2k_BT}\right)+1\right). \tag{5.31}$$

Tunnelling is a quantum effect, and therefore we had to treat the voltage as a quantum (i.e., non-commuting) variable. The correlator is then defined through the Heisenberg operators $\delta\hat{V}(t) \equiv \hat{V} - \langle\hat{V}\rangle$ by

$$\langle \delta\hat{V}_x(\omega)\delta\hat{V}_x(-\omega)\rangle \equiv \int_{-\infty}^{\infty} dt e^{i\omega(t_1-t_2)}\langle \delta\hat{V}_x(t_1)\delta\hat{V}_x(-t_2)\rangle. \tag{5.32}$$

[5] For a thorough account, see G.-L. Ingold and Yu. Nazarov, in "Single Charge Tunnelling", edited by H. Grabert and M. H. Devoret, NATO ASI Series B, Vol. 294, pp. 21-107 (Plenum Press, New York, 1992). This article can be found also at [cond-mat/0508728].

The dependence on absolute time $t_1 + t_2$ vanishes in a stationary system.

By Eq. (5.31), we see that the fluctuations only depend on the frequency, on temperature, and on the real part of the impedance $Z_t(\omega)$ between the point A in Fig. 5.5 and the ground. In the case when $Z(\omega) \ll R_T$, we can ignore the tunnel resistance and only use $Z(\omega)$ and the capacitance C of the tunnel junction,

$$Z_t(\omega) = \frac{1}{i\omega C + Z^{-1}(\omega)}. \tag{5.33}$$

The relaxation time of the circuit can now be related to $Z(\omega)$ and C.

The voltage fluctuations lead to time-dependent energies in the electrodes. Essentially this means that there is a k-independent shift of the energy $\epsilon_k \to \epsilon_k - e\delta V(t)$ of the electron dispersion relation in one of the reservoirs, fluctuating in time. Thus, the state of the reduced system (i.e., without the external resistor) described by the Hamiltonian similar to Eq. (5.9), but without the island term, would become time dependent. One can take this time dependence into account with the transformation $H' = UHU^\dagger + i\hbar(\partial_t U)U^\dagger$ where

$$U = \exp(\frac{i}{\hbar}\int^t dt' e(V_x + \delta\hat{V}(t'))\sum_{k,\sigma} c^\dagger_{k,\sigma}c_{k,\sigma}) \equiv \exp(i\phi(t)). \tag{5.34}$$

The Hamiltonian resulting from this transformation is of the form

$$H' = H_L + H_R + H_{\rm bath} + H'_T, \tag{5.35}$$

where H_L and H_R are the electrode Hamiltonian operators defined above, (except that the energies of the quasiparticle states in the left reservoir are shifted by the average potential V_x), $H_{\rm bath}$ describes the resistor and makes rise to the fluctuations $\delta\hat{V}$. The tunnelling Hamiltonian transforms to $H'_T = e^{-i\delta\phi}H_T$, where

$$\delta\phi = \frac{e}{\hbar}\int^t \delta\hat{V}(t')dt'. \tag{5.36}$$

Now the initial and final states in the calculation of the transmission probabilities are $|i\rangle = |i\rangle_e \otimes |i\rangle_R$ and $|f\rangle = |f\rangle_e \otimes |f\rangle_R$, where $|\cdot\rangle_e$ refers to the electronic excitations and $|\cdot\rangle_R$ to the charge state of the electrode(s). For example, the tunnelling rate from the left to the right is given by

$$\Gamma_{LR}(V) = \frac{1}{e^2R_T}\int d\varepsilon_k d\varepsilon_q f(\varepsilon_k)[1-f(\varepsilon_q)]\sum_{f,i} P_i^R |\langle f|_R e^{-i\delta\phi}|i\rangle_R|^2 \delta(\varepsilon_q - \varepsilon_k + E_f^R - E_i^R - eV_x). \tag{5.37}$$

Here P_i^R is the probability of the initial state of the reservoir, and E_f and E_i are the energies of the reservoir final and initial states.

Now let us use the following representation for the δ-function,

$$\delta(\varepsilon_q - \varepsilon_k + E_f^R - E_i^R - eV) = \int \frac{dt}{2\pi\hbar} e^{\frac{i}{\hbar}(eV + \varepsilon_k - \varepsilon_q + E_i^R - E_f^R)t}. \tag{5.38}$$

Inserting this in Eq. (5.37) yields

$$\begin{aligned}\Gamma_{LR}(V) =& \frac{1}{e^2 R_T}\int d\varepsilon_k d\varepsilon_q f(\varepsilon_k)[1-f(\varepsilon_q)]\int \frac{dt}{2\pi\hbar} e^{i(eV+\varepsilon_k-\varepsilon_q)t/\hbar} \\ & \times \sum_{f,i} P_i^R \langle i|_R e^{iE_i t/\hbar} e^{i\delta\phi} e^{-iE_f t/\hbar}|f\rangle_R \langle f|_R e^{-i\delta\phi}|i\rangle_R .\end{aligned} \tag{5.39}$$

Now assume the states $|i\rangle$ and $|f\rangle$ are eigenstates of energy and replace $e^{-iE_f t/\hbar}|f\rangle = e^{-iH_{\rm bath}t/\hbar}|f\rangle$. This amounts to a transformation to the Heisenberg picture, $e^{i\delta\phi(t)} = e^{iH_{\rm bath}t/\hbar}e^{i\delta\phi}e^{-iH_{\rm bath}t/\hbar}$. We may then use the resolution of unity, $\mathbf{1}_R = \sum_f |f\rangle_R\langle f|_R$, and obtain

$$\Gamma_{LR}(V) = \frac{1}{e^2 R_T}\int d\varepsilon_k d\varepsilon_q f(\varepsilon_k)[1-f(\varepsilon_q)]\frac{dt}{2\pi\hbar} e^{i(eV+\varepsilon_k-\varepsilon_q)t/\hbar}\langle e^{i\delta\phi(t)}e^{-i\delta\phi(0)}\rangle, \tag{5.40}$$

where $\langle\cdot\rangle = \sum_i p_i\langle i|\cdot|i\rangle$ denotes the statistical average.

Now let us define the function

$$P(E) = \int \frac{dt}{2\pi\hbar} e^{iEt/\hbar}\langle e^{i\delta\phi(t)}e^{-i\delta\phi(0)}\rangle. \tag{5.41}$$

Using this, we can express the tunnelling rate as

$$\begin{aligned}\Gamma_{LR}(V) =& \frac{1}{e^2 R_T}\int d\varepsilon_k d\varepsilon_q f(\varepsilon_k)[1-f(\varepsilon_q)]P(\varepsilon_k-\varepsilon_q+eV) \\ =& \frac{1}{e^2 R_T}\int dE \frac{E}{1-\exp(-E/(k_B T))}P(eV-E).\end{aligned} \tag{5.42}$$

The corresponding rate for the backward tunnelling (from right to left) satisfies the symmetry $\Gamma_{RL}(V) = \Gamma_{LR}(-V)$. Therefore, the current is given by

$$I(V_x) = e(\Gamma_{LR}-\Gamma_{RL}) = \frac{1}{eR_T}\int dE \frac{E}{1-\exp(-E/(k_B T))}(P(eV_x-E)-P(-eV_x-E)). \tag{5.43}$$

The function $P(E)$ has a probability interpretation (see exercise 5.5): it describes the probability that the environment can absorb the energy E during the tunnelling process.

5.5.1 Phase fluctuations

The function $\langle e^{i\delta\phi(t)}e^{-i\delta\phi(0)}\rangle$ can be interpreted as the characteristic function of the fluctuating quantity $\delta\phi(t)-\delta\phi(0)$, ordered such that $\delta\phi(t)$ precedes $\delta\phi(0)$.[6] A characteristic function can be expanded in the cumulants of the distribution,

$$\chi(t) = \langle e^{itx}\rangle = \exp\left(\sum_{n=0}^{\infty}\kappa_n\frac{(it)^n}{n!}\right). \tag{5.44}$$

[6] Characteristic function is defined as a Fourier transform of the probability density, see http://mathworld.wolfram.com/CharacteristicFunction.html.

Now let us assume that the phase fluctuations are Gaussian.[7] This is the case if the environment of the junction consists of macroscopic resistors much larger than the energy relaxation length. In this case, only the second cumulant of the phase fluctuations is non-vanishing and we may write

$$\langle e^{i\delta\phi(t)}e^{-i\delta\phi(0)}\rangle = e^{\langle[\delta\phi(t)-\delta\phi(0)]\delta\phi(0)\rangle} \equiv e^{J(t)}. \tag{5.45}$$

Here we used the fact that in a stationary situation $\langle\delta\phi(t)^2\rangle = \langle\delta\phi(0)^2\rangle$.

Using the definition (5.36), we may now relate the spectrum of phase fluctuations to the voltage fluctuations over the junction. These are given by

$$S_\phi(\omega) = \langle\delta\phi(\omega)\delta\phi(-\omega)\rangle = \frac{e^2}{\omega^2\hbar^2}\langle\delta\hat{V}(\omega)\delta\hat{V}(-\omega)\rangle. \tag{5.46}$$

5.5.2 Example: resistive environment

Let us now consider the simplest example of a purely resistive environment ($Z = R$ is frequency independent) to the tunnel junction, as in Fig. 5.5. In this case, the voltage fluctuation spectrum is given by Eq. (5.31). Then we have

$$J(t) = \int \frac{d\omega}{\omega}\frac{R}{R_K}\frac{e^{-i\omega t}-1}{1+\omega^2(RC)^2}\left[1+\coth\left(\frac{\hbar\omega}{2k_BT}\right)\right]. \tag{5.47}$$

Here $R_K = h/e^2 = 2R_Q$.

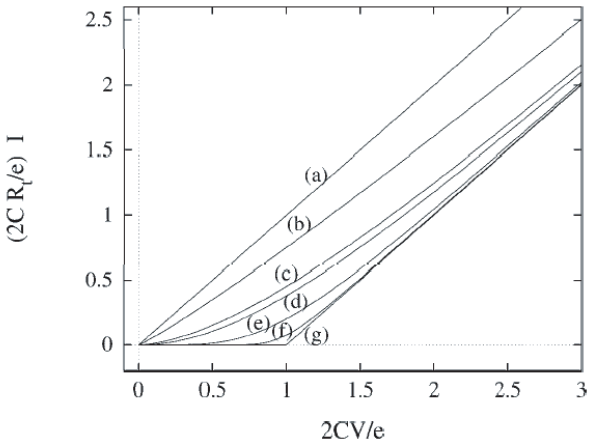


Figure 5.6: Current-voltage curves of a tunnel junction in a resistive environment. From (a) to (g), $R_K/R = \infty, 20, 3.2, 2, 0.4, 0.04, 0$. From Ch. 3 (by Gerd Schön) in the book T. Dittrich, *et al.*, "Quantum transport and dissipation", (Wiley-VCH, 1998).

[7] For an example of a non-Gaussian $P(E)$-theory, see T. T. Heikkilä, P. Virtanen, G. Johansson, and F. K. Wilhelm, Phys. Rev. Lett. **93**, 247 005 (2004).

Let us consider the environment effect in two limits. For $R \ll R_K$, $J(t) \to 0$, and $P(E) \approx \delta(E)$. In this case, we recover Eq. (1.7), and thus a linear $I - V$ curve. For the opposite limit of a very high-impedance environment, $R \gg R_K$, $Z_t \to \pi/C\delta(\omega)$, and we get

$$J(t) \approx -\frac{\pi}{CR_K}\left(it + \frac{k_B T t^2}{\hbar}\right). \tag{5.48}$$

With this $J(t)$, the $P(E)$ function becomes a Gaussian,

$$P(E) = \frac{1}{\sqrt{4\pi E_c k_B T}} \exp\left[-\frac{(E-E_c)^2}{4E_c k_B T}\right]. \tag{5.49}$$

In this case at $k_B T \ll E_C$, $P(E) \to \delta(E - E_c)$. In this limit, the expression for the current simplifies to

$$I(V_x) = \frac{eV_x - E_c}{eR_T}\theta(eV_x - E_c). \tag{5.50}$$

That is, we get a Coulomb blockade, which is washed out by increasing the temperature, or decreasing R/R_K. The $I(V)$ curve for a few environment resistances is shown in Fig. 5.6.

5.6 Other single-electron devices

Besides the single-electron transistor, there are many closely related devices that are based on a similar setting, or improve on the SET concept. We will introduce here a few of the most relevant ones: a Coulomb blockade thermometer (CBT) can be used as a very accurate primary electron thermometer, the radio frequency SET improves the charge measurement resolution of the SET by employing high frequencies, and a single-electron pump is a very accurate current source.

5.6.1 Coulomb blockade thermometer

Coulomb blockade thermometer (CBT) works in the regime $k_B T \gg E_C$. There, we essentially start from the infinite temperature limit of the current-voltage curves and calculate the lowest-order correction from Coulomb blockade. This correction turns out to be a universal function of the temperature with two operating modes: primary (no calibration required) and secondary (easier operation, but one needs to know the junction parameters).[8]

In the limit $k_B T \gg E_C$, the first-order correction from the charging effects to the linear current-voltage curve can be used as a thermometric quantity.

[8] CBT was invented by J. Pekola, *et al.*, see J. P. Pekola, K. P. Hirvi, J. P. Kauppinen, and M. A. Paalanen, Phys. Rev. Lett. **73** 2903 (1994). A brief review of the main CBT properties is given in F. Giazotto, T. T. Heikkilä, A. Luukanen, A. M. Savin, and J. P. Pekola, Rev. Mod. Phys. **78**, 217 (2006).

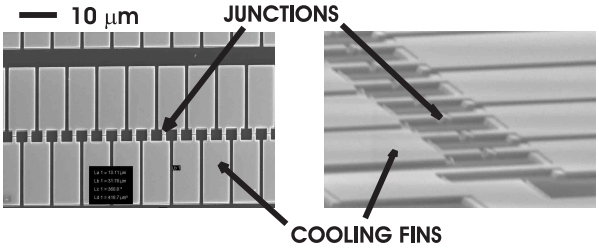


Figure 5.7: A typical CBT sensor for the temperature range 20 mK - 1 K. The structure has been fabricated by electron beam lithography, combined with aluminium and copper vacuum evaporation. Both top view and a view at an oblique angle are shown; the scale indicated refers to the top view. Figure adapted from M. Meschke, *et al.*, J. Low Temp. Phys. **134**, 1119 (2004).

Indeed, in the case of a symmetric SET at $Q_G = 0$, $V_L = -V_R = V/2$, the differential conductance $G = dI/dV$ is of the form (number 2 comes from the fact that we have two junctions - you will derive this in Exercise 5.7)

$$2GR_T = 1 - ug(v), \tag{5.51}$$

where $u = 2E_c/k_BT$, $v = eV/(2k_BT)$ and

$$g(v) = \frac{v\cosh\left(\frac{v}{2}\right) - 2\sinh\left(\frac{v}{2}\right)}{4\sinh^3\left(\frac{v}{2}\right)}. \tag{5.52}$$

Now measuring the conductance correction yields the electronic temperature: the conductance at $v = 0$ is $G = 1/(2R_T) - \Delta G$ with $2\Delta GR_T = E_C/(3k_BT)$. But to find out the precise temperature from here requires calibration, i.e., the knowledge of the tunnel resistance R_T and the charging energy E_C. However, the voltage V for which $g(v) = 1/12$, i.e., the conductance change is half from its maximum change, is a primary thermometric quantity: this position is reached with $v = 2.7196$, i.e., $eV = 5.4392k_BT$, i.e., independent of R_T or E_C. Thus, the measurement of the width of the $G(V)$ curve yields an ideal measurement of the electronic temperature.

In practice, cotunnelling limits the sensitivity of CBT devices. This can be cured by placing several junctions in a row. For N junctions in series, one obtains

$$2GR_T = 1 - u_N g(v_N), \tag{5.53}$$

where $v_N = eV/(Nk_BT)$ and $u_N = 2(N-1)/NE_c/k_BT$ with $E_c = e^2/(2C)$ calculated for a single-junction capacitance C.

Figure 5.7 shows an example of a CBT and Fig. 5.8 its conductance-voltage curve, together with a comparison to another frequently used thermometer.

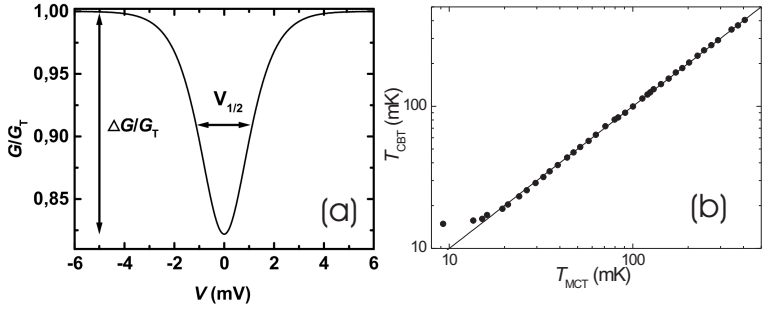


Figure 5.8: (a) Measured conductance-voltage curve of a CBT thermometer. $G(V)/G_{\rm T}$ is the differential conductance scaled by its asymptotic value at large positive and negative voltages, plotted as a function of bias voltage V. $V_{1/2}$ indicates the full width at half minimum of the characteristics. The full depth of the line, $\Delta G/G_{\rm T}$, is another parameter to determine temperature. In (b) temperature deduced by CBT has been compared to that obtained by a ^{3}He melting curve thermometer. Saturation of CBT below 20 mK indicates typical thermal decoupling between electrons and phonons. Figure from Ref. M. Meschke, *et al.*, J. Low Temp. Phys. **134**, 1119 (2004).

5.6.2 Radio frequency SET

The normal-state single-electron transistor is in principle a very sensitive detector of charge. However, it is typically quite slow to operate due to the RC-times of the measuring circuit (C is in such a case given by the capacitive load from cabling, and is typically between 0.1 to 1 nF). Therefore, the speed is limited to below some 1 MHz, i.e., obtaining the charge information requires measuring the state at least for a μs. In this regime, the single-electron devices show large $1/f$ fluctuations (i.e., noise that scales as one over frequency of operation) due to the presence of offset charges in the SET environment.

A radio-frequency SET (RF-SET) cures this problem by placing the SET in a resonant circuit (see Fig. 5.9). Instead measuring the current-voltage curve of the SET, one monitors the reflection of a sinusoidal high-frequency carrier from the resonant circuit. The reflection coefficient is given by

$$\Gamma = \frac{Z - Z_0}{Z + Z_0}, \tag{5.54}$$

where Z is the impedance of the load (the resonant circuit with the SET) and Z_0 is the characteristic impedance of the waveguide (the coaxial cable). These impedances should be measured at the carrier frequency, which can be taken up to a few GHz. At these frequencies, $Z_0 \approx 50\Omega$ and the resonant circuit can be used to transform the SET impedance down. On resonance, the impedance Z depends strongly on the exact value of the SET impedance, which can then be controlled by the gate charge.

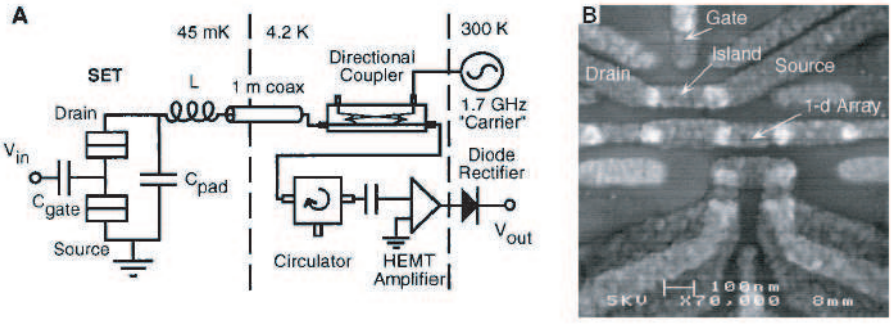


Figure 5.9: A. Scheme of a radio-frequency SET and B. Electron micrograph of a practical device. In A, the resonant circuit is formed from the inductor L and the parasitic capacitance $C_{\rm pad}$ of the contact pad to the SET. From R. Schoelkopf, *et al.*, Science **280**, 1238 (1998).

This scheme allows to use SET for charge detection at frequencies of the order of hundreds of MHz (depends on the bandwidth of the amplifiers, and not so much on the SET itself). In the first demonstration of this device,[9] charge sensitivity of order $10^{-5}e/\sqrt{\rm Hz}$ was obtained (i.e., within 1μs, one could measure the charge with the accuracy of 0.01 e).

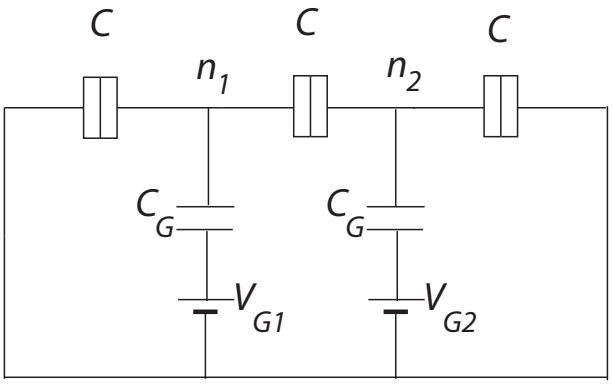


Figure 5.10: A symmetric two-island single-electron device which can be used for pumping.

5.6.3 Single electron pump

Fabricating two or more junctions in series and controlling each of the islands with separate gate voltages allows one to realize a current pump. Let us consider

[9] R. Schoelkopf, *et al.*, Science **280**, 1238 (1998).

the charging energy of the device shown in Fig. 5.10. It turns out that the charging energy in this system is

$$E_{\rm ch} = \frac{2}{3}E_c[(n_1 - n_{G1})^2 + (n_2 - n_{G2})^2 + (n_1 - n_{G1})(n_2 - n_{G2})], \qquad (5.55)$$

where $E_c = e^2/2C$, n_i is the charge on the i'th island ($i = 1, 2$), and $n_{Gi} = V_{Gi}C_{Gi}/e$ is the corresponding gate charge. The charge states (n_1, n_2) corresponding to the ground state of the system are plotted in Fig. 5.11.

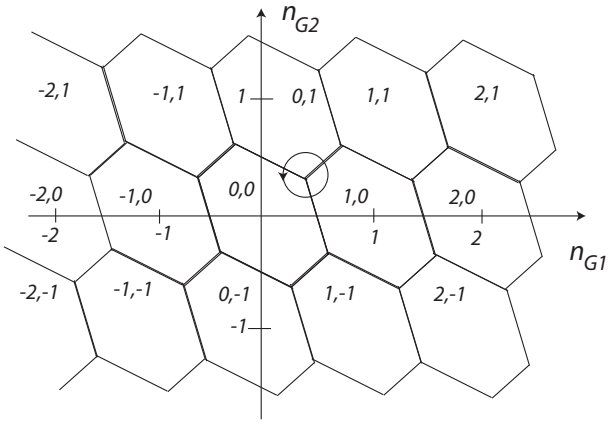


Figure 5.11: Ground state configuration of the two-island system as a function of the two gate charges n_{Gi}.

Consider what happens if the gate voltages are changed in time according to the circle drawn in Fig. 5.11 around the states (0,0), (1,0) and (0,1), such that the system follows the gates adiabatically. When $n_{Gi} < 1/3$, the ground state of the system corresponds to no extra electrons on the island. Increasing n_{G1} but keeping n_{G2} roughly constant changes the ground state to (1,0), i.e., one extra electron tunnels from the left electrode into the left island. Now increasing n_{G2} and keeping n_{G1} between 1/3 and 2/3 changes the ground state to (0,1), and one electron tunnels from the left island to the right island. Now decreasing first n_{G1} and then n_{G2} below 1/3 leads to the charge state (0,0), i.e., one extra electron tunnels from the right island to the right electrode. As a result of this process, one electron tunnelled through the whole system, resulting into a current equal to e/τ, where τ is the time in which the sequence was carried out. Repeating this sequence with frequency f leads to a current

$$I_p = ef \qquad (5.56)$$

pumped through the device. As the frequency f can be controlled very precisely, this is a very precisely defined current.

In practice, cotunnelling limits the accuracy of the current, as it leads to the possibility of processes where electrons are carried back and forth through the

whole device. But fabricating more junctions in series reduces the cotunnelling current, and leads to an increasing accuracy. Figure 5.12 shows a practical device consisting of seven junctions (i.e., six islands) where an accuracy of 10^{-8} was reached. This is the most accurate known way to produce currents. However, the method is not totally suitable for current standards, as the adiabaticity requirement sets the limiting frequency to a few MHz, corresponding to currents of the order of $0.1 \ldots 1$ pA.

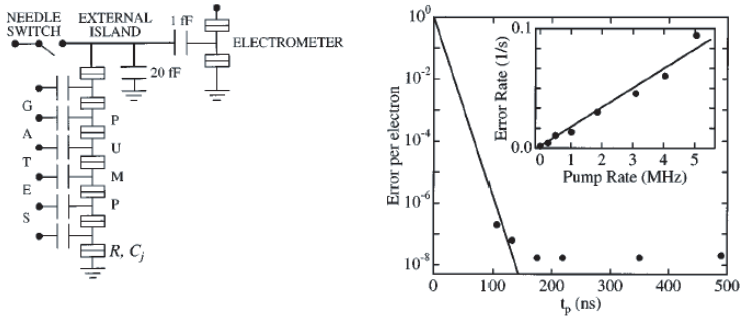


Figure 5.12: Left: Schematics of a charge pump based on a single-electron transistor with many islands in series. Right: Error rate per electron vs. time $t_p \sim 1/f$ to pump a single electron. Revised from M. W. Keller, *et al.*, Appl. Phys. Lett. **69**, 1804 (1996).

5.7 Problems

1. Show that the forward and backward tunnelling rates satisfy the detailed balance criterion,
$$\frac{\Gamma_{LI}(n)}{\Gamma_{IL}(n+1)} = e^{-\delta E_{ch}/(k_B T)}. \tag{5.57}$$

2. **Coulomb diamond:** Construct the stability diagram of the SET at $T = 0$: for given gate charges Q_G and bias voltages $V_L = -V_R = V/2$, find the charge states n whose occupation number is finite, i.e., which can contribute to the transport. Plot this in a 2d plane, varying the gate charge between $-2e$ and $2e$, and the bias voltage between $-2e/C$ and $2e/C$. You do not need to evaluate the tunnelling rates. The figure you obtain is a characteristic conductance plot measured for all single-electron devices. Hint: find the $\{V, Q_G\}$ pairs where the conditions (5.7) and (5.8) apply as an equality. Above and below these lines the allowed charge states are different.

3. Solve the master equation in the finite-temperature equilibrium ($V_L = V_R = 0$) case and calculate the average charge on the island for a given gate charge.

4. Find the charging energy for a single-electron transistor in terms of n, $V_{L/R/G}$ starting from $E_{ch} = Q_L^2/(2C_L)+Q_R^2/(2C_R)+Q_G'^2/(2C_G)-V_LQ_L-V_RQ_R-V_GQ_G'$. Here Q_L and Q_R are the polarization charges on the left and right junctions and Q_G' is the charge on the gate capacitor.

5. Show that the function $P(E)$ satisfies the sum rules

$$\int dEP(E) = 1 \tag{5.58}$$

$$\int dEEP(E) = E_C. \tag{5.59}$$

   The first equation is a requirement for the probability interpretation of $P(E)$. In the latter equation, use the Gaussian result with Eqs. (5.31) and (5.33) with a frequency independent Z.

6. Using the sum rules of the previous exercise, show that at very large voltages $eV_x \gg k_BT, E_c$, the current through a single tunnel junction satisfies

$$I(V_x) = \frac{V_x - E_c/e}{R_T}. \tag{5.60}$$

   Note that the result $I = V_x/R_T$ obtained in the limit of a vanishing environmental impedance does not satisfy the second sum rule. Assuming a small but finite impedance Z satisfies the sum rule and leads to the above offset, but yields an (almost) linear $I - V$ curve.

7. Derive the conductance change ΔG of a CBT in the case of two junctions. *Hint: i) Write the tunnelling rates $\Gamma_{LI}(n)$ and $\Gamma_{IL}(n)$ for a vanishing gate voltage, $Q_G = 0$ and for a symmetric CBT, with $V_L = -V_R = V/2$. ii) Expand the difference $\Gamma_{LI}(n) - \Gamma_{IL}(n)$ in $u \equiv E_c/k_BT$ and find the zeroth- and first-order terms in u. iii) Write the current $I = e\sum_{n=-\infty}^{\infty} P(n)(\Gamma_{LI}(n) - \Gamma_{IL}(n))$ using this expansion, and use the symmetries of $P(n)$ at $Q_G = 0$ to get rid of the sums. iv) Differentiate your results to get the differential conductance.*

Chapter 6

Introduction to superconductivity

Superconductivity is a quantum-mechanical effect, but in contrast to the phenomena discussed in the previous chapters, it does not disappear as the system size is enlarged. Therefore, it is not a mesoscopic effect. However, as most mesoscopic experiments are performed at low temperatures where many materials turn superconducting, it both strongly affects the mesoscopic systems and it can be utilized for creating many different types of phenomena that would not be present in normal (i.e., non-superconducting) metal systems.

A thorough review of superconductivity and superconducting effects is given in the excellent book by Michael Tinkham.[1] In what follows, we will mostly summarize the main phenomena.

6.1 Cooper pairing

Superconductivity of metals was discovered as early as 1911 by H. Kamerlingh Onnes in Leiden, but it took until 1957 before a microscopic theory was laid out by Bardeen, Cooper and Schrieffer (BCS).[2]

The BCS theory is based on the idea that some physical process could make rise to an effectively attractive interaction between the electrons. Due to this interaction, the ground state of the system changes such that it contains correlated pairs of electrons. The amount of these correlations is measured by a correlation function called a *pair amplitude* whose mathematical definition is based on the statistical average of a pair of second-quantized field operators (assumption of a local correlation is valid in the same semiclassical limit described in Ch. 4, i.e., assuming that we are interested only in length scales far

[1] M. Tinkham, "Introduction to superconductivity", McGraw-Hill, Singapore (2nd edition 1996).

[2] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **106**, 162; **108**, 1175 (1957).

exceeding λ_F),

$$F_{\sigma\sigma'}(\vec{r}) = \langle \psi_\sigma(\vec{r})\psi_{\sigma'}(\vec{r})\rangle. \tag{6.1}$$

For conventional superconductors (Al, Nb, Pb, Sn, Ti, etc.), only the spin singlet pair amplitude $\sigma' = -\sigma$ is finite, and $F_{\uparrow\downarrow}(\vec{r}) = F_{\downarrow\uparrow}(\vec{r}) \equiv F(\vec{r})$. In what follows, we concentrate only on this case. Essentially, what Eq. (6.1) means is that in the superconducting state with $F(\vec{r}) > 0$, the positions of pairs of electrons are correlated.

The appearance of superconducting correlations is a phase transition, and it is described by an order parameter $F(\vec{r})$. An alternative choice for the order parameter is the pair potential

$$\Delta(\vec{r}) = g(\vec{r})F(\vec{r}), \tag{6.2}$$

where $g(\vec{r})$ characterizes the strength of the attractive interaction. The pair potential and the pair amplitude are complex functions of $\vec{r}$, and often it is useful to present $\Delta(\vec{r})$ in the form

$$\Delta(\vec{r}) = |\Delta(\vec{r})|e^{i\varphi(\vec{r})}. \tag{6.3}$$

In a bulk superconductor in the presence of a small magnetic field (in the absence of vortices), the absolute value $|\Delta|$ is constant, but the phase $\varphi(\vec{r})$ may be space dependent. This space dependence of $\varphi(\vec{r})$ leads to supercurrent.

The effective attractive interaction turned out to be related to the electron-phonon coupling in the conventional superconducting materials. This can be phenomenologically understood as follows: consider an electron moving in a lattice. As it is negatively charged, its movement exerts a force on the positively charged ions, which then slightly move towards the electron. Now the other electrons in the system may sense this movement, and follow the movement of the ions, i.e., towards the first electron. Thus, an effectively attractive interaction was created and with certain materials parameters and average distance between the electrons, it can even beat the repulsive Coulomb interaction.

6.2 Main physical properties

6.2.1 Current without dissipation

The defining property of a superconductor is that it can carry current without dissipation, i.e., without inducing a voltage drop. Indeed, it turns out that in addition to the usual *quasiparticle current* (similar to the normal-state current), there is a possibility for a *supercurrent*. Whereas the previous is proportional to the potential gradient, the latter may be present due to a gradient in the phase $\varphi(\vec{r})$. The supercurrent density may be written phenomenologically as

$$j_S = 2en_s v_S, \tag{6.4}$$

where n_s is the density of Cooper pairs, related to the superconducting order parameter Δ, and $v_S \propto \nabla\varphi$ is the "superfluid velocity". The previous is generally temperature dependent, such that above a certain temperature, *critical temperature* T_c it vanishes.

If a current is driven through a superconductor, it is converted into a supercurrent, and as a result, no voltage builds up across the superconductor. This means that the resistance of the superconductor vanishes as long as $T < T_c$. This behavior is illustrated in Fig. 6.1, taken from the Nobel lectures of H. Kamerlingh Onnes.

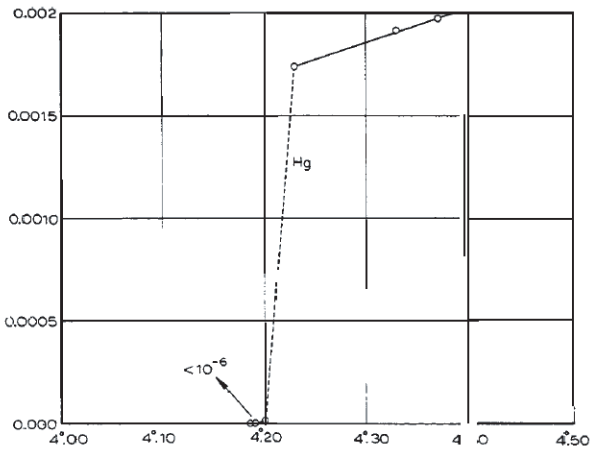


Figure 6.1: Resistance of mercury around its critical temperature $T_c = 4.2K$. From the Nobel lectures of H. Kamerlingh Onnes (1913).

6.2.2 Meissner effect

One of the hallmarks of superconductivity in bulk metals is the Meissner effect: a magnetic field makes rise to circulating supercurrents on the superconductor surface. These supercurrents make rise to an opposing field, such that the total field decays into the bulk superconductor.[3] This is why a superconductor is a *perfect diamagnet*. The corresponding decay length of the magnetic field is called a *penetration depth* λ.

Typically the penetration depth λ is of the same order or longer than the thickness of the films used for mesoscopic expermiments. Therefore, . As the thickness of the mesoscopic samples is generally of the same order, magnetic field may often be considered homogeneous in mesoscopic films. However, the

[3]This is true for magnetic fields below a certain critical field H_c. Above this field, superconductivity is destroyed. In certain types of superconductors, above a given field value $H > H_{c1}$, there appears a topological excitation, a vortex, through which the magnetic field may penetrate the sample, and only above a much higher field H_{c2}, superconductivity is totally destroyed.

presence of circulating supercurrents due to a magnetic field is relevant especially in superconducting loops or SQUIDs, which are discussed in the next chapter.

6.2.3 Energy gap and BCS divergence

It turns out from the BCS theory that the density of states for the quasiparticle excitations from the superconducting ground state is of the form

$$N_S(E) = N_0 \frac{E}{\sqrt{E^2 - |\Delta|^2}} \theta(|E| - |\Delta|). \tag{6.5}$$

Here $N(0)$ is the corresponding density of states at the Fermi level when the sample is in the normal state. This function is plotted in Fig. 6.2.

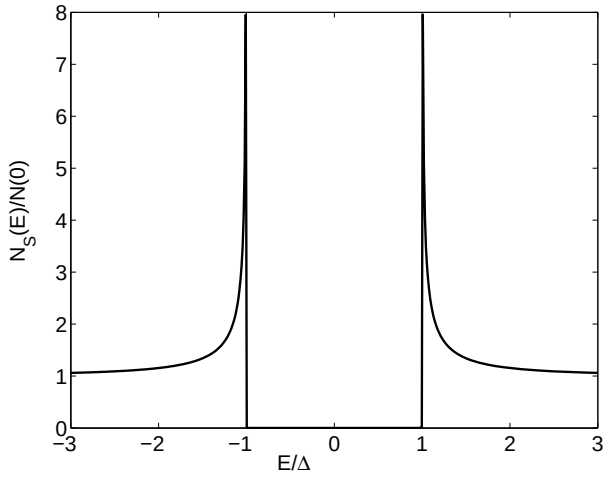


Figure 6.2: BCS bulk superconducting density of states.

There are two main features in this expression that show up in the experiments on superconductors: First, there is the *energy gap* of size $|\Delta|$, below which there are no available states for the quasiparticles. Therefore, the pair potential is often called the gap function. Thus, if the temperature is far below $|\Delta|$, quasiparticles do not affect the behavior of the superconductors. The second notable feature is the *BCS divergence* of the density of states at $E = \Delta$. This shows up for example in the tunnelling current as discussed in the next chapter.

The first experimental evidence of the energy gap was related to the electronic specific heat of the superconductors. In contrast to the linear dependence $C_{en} = \gamma T$ valid in the normal state, well below T_c it was found to decay exponentially with the temperature, $C_{es} \propto T_c e^{-|\Delta|/(k_B T)}$. This is a direct consequence of Eq. (6.5).

The energy gap is generally temperature dependent, and it vanishes at the critical temperature (anyway, it serves as the order parameter of the superconducting state). In the BCS model the energy gap is related with the critical temperature by the relation

$$\Delta(T=0) = 1.764 k_B T_c. \tag{6.6}$$

The critical temperature of Al, frequently used in mesoscopic experiments, is $T_c \approx 1.2$ K.[4] This means that the zero-temperature energy gap for Al is of the order of 180 μeV.

6.2.4 Coherence length

Apart from the penetration depth, there is another important characteristic length scale which shows up in phenomena related to superconductivity. This is the *coherence length* ξ_0, which is a characteristic scale that describes deviations of the superconducting order parameter. For example, at an interface between a superconductor and a normal metal where superconductivity is suppressed because of the presence of the normal metal, this is the length scale in which the order parameter regains its bulk value. The coherence length in a pure superconductor (when $\xi_0 \ll \ell_{\mathrm{el}}$) is given by

$$\xi_0^{\mathrm{clean}} = \frac{\hbar v_F}{\pi |\Delta|}, \tag{6.7}$$

and in the dirty case $\xi_0 \gg \ell_{\mathrm{el}}$ it is

$$\xi_0^{\mathrm{dirty}} = \sqrt{\frac{\hbar D}{2\Delta}}. \tag{6.8}$$

A typical value for the elastic mean free path in mesoscopic films is $\ell_{\mathrm{el}} \sim 30$ nm (note that this is just an order of magnitude estimate!). With this mean free path and with the Fermi velocity of Al, $v_F = 2.03 \times 10^6$ m/s, D would be roughly 0.02 m^2/s. With the above-mentioned energy gap this yields for aluminum $\xi_0 = 190$ nm.

The coherence length and the penetration depth indicate one possible "crossover" length for mesoscopic phenomena: indeed, one of the interesting questions is how superconducting effects behave as the structure sizes are shorter than these lengths.

6.3 Josephson effect

The most important superconducting effect for mesoscopic systems is the Josephson effect, which arises if one places two superconductors in a weak contact, such that the two systems are almost isolated, and the contact only allows the tunnelling of Cooper pairs across it.

[4] Although in Al films, the critical temperature may be higher than this, even up to 2K.

In practice, one talks about three types of Josephson weak links: the most usually encountered weak link is the one where the two superconductors are connected through an insulating barrier (for example, an oxide layer formed between the two). In this case, the transmission probability of all the eigenmodes of the contact (eigenmodes are the eigenvectors of the scattering matrix describing the contact, and the corresponding transmission probabilities are the absolute squares of the eigenvalues). Such a junction is called a SIS junction (Superconductor–Insulator–Superconductor). Other types of weak links are the Superconductor–Normal metal–Superconductor (SNS) weak link, and the Superconductor-constriction-Superconductor (ScS) weak link. In the latter, the weak link is also from the same superconducting material, but its transverse dimension is low.

A phenomenological description of the Josephson effect can be made following the treatment of R. Feynman. Assume the amplitude to find a Cooper pair on one side of the junction is $\psi_1 = \sqrt{n_1}e^{i\varphi_1}$ and on the other side it is $\psi_2 = \sqrt{n_2}e^{i\varphi_2}$. The two are related by

$$i\hbar\frac{\partial\psi_1}{\partial t} = U_1\psi_1 - K\psi_2 \tag{6.9a}$$

$$i\hbar\frac{\partial\psi_2}{\partial t} = U_2\psi_2 - K\psi_1. \tag{6.9b}$$

The constant K depends on the properties of the junction. For an isolated case, $K = 0$, and the remaining equations would simply describe the lowest-energy state of the superconductors. Now let us assume that both sides are connected to two sides of a voltage source, such that there is a potential difference qV between the two, i.e., $U_1 - U_2 = qV$. Here q is the charge of the current carriers ($q = 2e$ for a Cooper pair). Defining the zero of energy in halfway between U_1 and U_2 we get

$$i\hbar\frac{\partial\psi_1}{\partial t} = \frac{qV}{2}\psi_1 - K\psi_2 \tag{6.10a}$$

$$i\hbar\frac{\partial\psi_2}{\partial t} = -\frac{qV}{2}\psi_2 - K\psi_1. \tag{6.10b}$$

Now let us use the definitions of ψ_i in terms of n_i and φ_i and rewrite the above equations as

$$\frac{\partial\psi_1}{\partial t} = \frac{1}{2\sqrt{n_1}}e^{i\varphi_1}\frac{dn_1}{dt} + i\sqrt{n_1}e^{i\varphi_1}\frac{d\varphi_1}{dt} = \frac{qV}{2i\hbar}\sqrt{n_1}e^{i\varphi_1} - \frac{K}{i}\sqrt{n_2}e^{i\varphi_2} \tag{6.11a}$$

$$\frac{\partial\psi_2}{\partial t} = \frac{1}{2\sqrt{n_2}}e^{i\varphi_2}\frac{dn_2}{dt} + i\sqrt{n_2}e^{i\varphi_2}\frac{d\varphi_2}{dt} = -\frac{qV}{2i\hbar}\sqrt{n_2}e^{i\varphi_2} - \frac{K}{i}\sqrt{n_1}e^{i\varphi_1}. \tag{6.11b}$$

Now multiply the above equation by $e^{-i\varphi_1}$ and the below equation by $e^{-i\varphi_2}$, and then separate the real and imaginary parts of the resulting expressions,

remembering that both n_i and φ_i are real. The real part yields

$$\frac{dn_1}{dt} = -2K\sqrt{n_1 n_2}\sin(\varphi) \tag{6.12a}$$

$$\frac{dn_2}{dt} = 2K\sqrt{n_1 n_2}\sin(\varphi), \tag{6.12b}$$

where $\varphi = \varphi_2 - \varphi_1$. The imaginary part can be written as

$$\frac{d\varphi_1}{dt} = -\frac{qV}{2\hbar} + K\sqrt{\frac{n_2}{n_1}}\cos(\varphi) \tag{6.13a}$$

$$\frac{d\varphi_2}{dt} = \frac{qV}{2\hbar} + K\sqrt{\frac{n_1}{n_2}}\cos(\varphi). \tag{6.13b}$$

From the real part, we obtain an expression for the current from the left superconductor, $j = -q\frac{dn_1}{dt}$,

$$j = j_c \sin(\varphi_1 - \varphi_2), \tag{6.14}$$

where $j_c = 2qK\sqrt{n_1 n_2}$ is called the critical current density for this junction. Multiplying this by the area of the junction yields

$$I = I_C \sin(\varphi_1 - \varphi_2), \tag{6.15}$$

with $I_C = Aj_c$. Equation (6.15) is called the *dc Josephson relation*: it shows that even in the absence of a potential drop, one can maintain a current through the junction. Moreover, this current depends on the phase difference across the junction. If the current exceeds the critical current I_C, a voltage starts to appear. A more microscopic approach yields an equation for I_C, relating it to the junction resistance R_T at the normal state and the values of the order parameter in the superconductors. In the case $\Delta_1 = \Delta_2$ this Ambegaokar-Baratoff relation, valid for tunnel contacts, reads

$$I_C = \frac{\pi\Delta}{2eR_T}\tanh\left(\frac{\Delta(T)}{2k_BT}\right). \tag{6.16}$$

For example, for an Al Josephson junction with $R_T = 1\Omega$, we would obtain $I_C = 280\ \mu$A at $T = 0$.

Let us consider the difference of the imaginary parts of the above relations, Eq. (6.13b), with similar superconductors, suitably close to equilibrium. In this case $n_1 \approx n_2$ and we obtain

$$\frac{d\varphi}{dt} = \frac{d}{dt}(\varphi_2 - \varphi_1) = \frac{qV}{\hbar} = \frac{2eV}{\hbar}. \tag{6.17}$$

The latter form is obtained by using the charge $q = 2e$ of a Cooper pair. Solving this for a constant voltage yields $\varphi = \varphi(0) + 2eVt/\hbar$ and

$$I = I_C \sin(\varphi(0) + 2eVt/\hbar). \tag{6.18}$$

Therefore, the current through a Josephson junction with an applied dc voltage oscillates in time. This is the *ac Josephson effect*.

Above, we gave only the simplest model for a Josephson junction, capturing still the two most relevant effects. However, Josephson junctions are one of the most complicated systems studied in physics, and they continue to be studied even now. In the next chapter I will introduce few of the mesoscopic phenomena observed in Josephson junctions.

6.4 Main phenomena characteristic for mesoscopic systems

Apart from the phenomena related to the bulk properties of superconductors, two microscopic effects show up especially in small hybrid structures composed of normal metals and superconductors. These are the Andreev reflection and the proximity effect, which are in many ways related. To illustrate the microscopic theory of superconductivity, we calculate below the Andreev reflection probabilities using the Bogoliubov–de Gennes equation, which is a generalization of the BCS theory to the inhomogeneous case.

6.4.1 Bogoliubov-de Gennes equation

The microscopic state of an inhomogeneous superconductor can be described with an equation derived first by Bogoliubov and a little later by de Gennes, who published it to the western world.[5] This equation is a generalization of the BCS model (that was constructed for a bulk and homogeneous superconductor) to the inhomogeneous case. We will not present its derivation here, it suffices to know that it is a mean-field approximation, and the coupling is assumed to be weak, such that its product with the density of states at the Fermi energy is small, $gN(0) \ll 1$. Most conventional superconductors satisfy this inequality at least roughly, and the deviations from the BCS theory show up mostly in the fact that the critical temperature does not quite scale correctly with the zero-temperature pair potential.

The Bogoliubov-de Gennes equation reads

$$\begin{pmatrix} H_0 & \Delta(\vec{r}) \\ \Delta^*(\vec{r}) & -H_0^\dagger \end{pmatrix} \begin{pmatrix} u(\vec{r}) \\ v(\vec{r}) \end{pmatrix} = E \begin{pmatrix} u(\vec{r}) \\ v(\vec{r}) \end{pmatrix}, \tag{6.19}$$

where

$$H_0 = \frac{1}{2m}\left(\frac{\hbar}{i}\nabla - e\mathbf{A}\right)^2 + U(\vec{r}) - \mu \tag{6.20}$$

and $\Delta(\vec{r})$ is the pair potential. The potential $U(\vec{r})$ describes the impurities and any overall electrostatic potentials.

If $\Delta(\vec{r}) = 0$, the equations decouple into the form

$$H_0 u = Eu \tag{6.21a}$$

$$H_0^\dagger v = -Ev. \tag{6.21b}$$

[5] This was the time of cold war.

The first is just the usual Schrdinger equation for the electrons (with the only difference that the energies are measured with respect to the Fermi energy). The latter equation describes time-reversed excitations, i.e., holes. In this way, the Bogoliubov-de Gennes equation can be understood as the Schrdinger equation for the superconducting state, where the pair potential couples electron- and hole-like excitations.[6]

This equation has to be appended by the self-consistency equation for the order parameter. In terms of the eigenfunctions $u_n(\vec{r})$ $v_n(\vec{r})$ of Eq. (6.19) with the corresponding eigenenergies E_n, it is written as

$$\Delta(\vec{r}) = g(\vec{r})F(\vec{r}) = g(\vec{r})\sum_n v_n^*(\vec{r})u_n(\vec{r})\tanh\left(\frac{E_n}{2k_BT}\right). \tag{6.22}$$

This is a self-consistency equation, as the result $\Delta(\vec{r})$ should be used to solve for the functions $u_n(\vec{r})$ and $v_n(\vec{r})$.

6.4.2 Andreev reflection

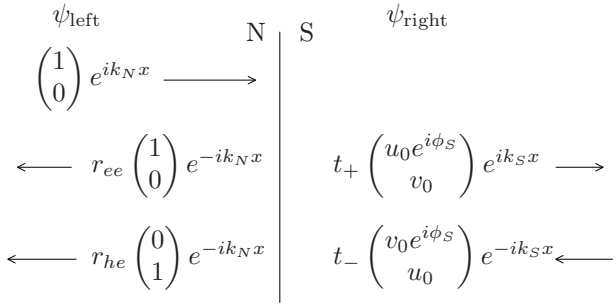


Figure 6.3: Andreev reflection at a NS interface.

An important phenomenon in mesoscopic superconducting systems is the Andreev reflection, where an electron in a normal metal hits a superconductor and is converted into a hole. This is a result of the appearance of the energy gap in the superconductor. Indeed, consider a structure where a normal-metal wire (say, Cu) is brought in good contact with a superconductor (say, Al). Far in the normal metal, we see no effects of superconductivity, and the density of states is flat. On the other hand, far in the superconductor, there are no effects from the normal metal, and the density of states has a gap. Then, if an electron excitation in the normal metal with energy lower than the gap (i.e., $E \in [E_F, E_F + \Delta]$) hits a superconductor, it cannot enter it because there are no states for it below the gap. However, if the contact is clean, there can also be no direct elastic scattering into another electron state traversing out from the contact. What happens is an *Andreev reflection*: the electron reflects as

[6] The electron-hole state space spanned by the vectors of the form $(u_n \quad v_n)^T$ is typically called the Nambu space.

a hole (an excitation below the Fermi sea). As a hole has a negative charge, the net result is an extra Cooper pair (with a double electronic charge) inside the superconductor. This thus means that Andreev reflection carries charge current.[7]

The process can be exemplified with a simple system where a normal metal (with $g(\vec{r}) = 0$) is in contact with a superconductor (with $g(\vec{r}) = g = const.$) (see Fig. 6.3).[8] For simplicity, let us assume that the materials are clean, such that $U(\vec{r}) = 0$, and that there is no magnetic field. Moreover, we can neglect self-consistency by assuming a constant $\Delta(\vec{r}) = \Delta$ inside the superconductor.

The solution to the Bogoliubov–de Gennes equation is of the form $\psi = \psi_{\text{left}} + \psi_{\text{right}}$ (as in Fig. 6.3), where

$$\psi_{\text{left}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{ik_N x} + r_{ee} \begin{pmatrix} 1 \\ 0 \end{pmatrix} e^{-ik_N x} + r_{he} \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{-ik_N x} \tag{6.23}$$

is the wave function on the left-hand side of the interface, composed of an incident electron moving towards the right, a normally reflected electron, and a hole that has been *Andreev reflected* from the electron state. Defining $\epsilon = E - \mu$, the wave number in the normal state is given by

$$k_N = \sqrt{\frac{2m\epsilon}{\hbar^2}}. \tag{6.24}$$

On the superconducting side, the electron can get transmitted into one of the two types of eigenstates of a bulk superconductor (calculating these is left as an exercise). Therefore, the solution in the superconductor is of the form

$$\psi_{\text{right}} = t_+ \begin{pmatrix} u_0 e^{i\phi_S} \\ v_0 \end{pmatrix} e^{ik_S x} + t_- \begin{pmatrix} v_0 e^{i\phi_S} \\ u_0 \end{pmatrix} e^{-ik_S x}. \tag{6.25}$$

Here, the wave number is given by

$$k_S = \sqrt{\frac{2m[\mu \pm \sqrt{E^2 - \Delta^2}]}{\hbar^2}}. \tag{6.26}$$

If the NS interface is clean, the transmission and reflection coefficients are obtained by requiring continuity of the functions and their derivatives at the surface point, $x = 0$. Such a calculation yields $r_{ee} = 0$, $|r_{he}|^2 = 1$ for $|\epsilon| < \Delta$ and $|r_{he}|^2 = (\epsilon - \sqrt{\epsilon^2 - \Delta^2})/(\epsilon + \sqrt{\epsilon^2 - \Delta^2})$ for $|\epsilon| > \Delta$. As $|r_{he}|^2 = 1$ for $\epsilon < \Delta$, Andreev reflection is the only allowed process for this energy range.

If the interface is not clean, also normal reflection takes place. Figure 6.4 illustrates the probabilities $A = |r_{he}|^2$, $B = |r_{ee}|^2$, $C = |t_+|^2$ and $D = |t_-|^2$ of the various processes as functions of energy for a few values of the transmission probability T of the interface (this is the transmission probability one would

[7] However, it does not carry heat as the Cooper pairs do not carry heat.

[8] This example is due to G. E. Blonder, M. Tinkham, and T. M. Klapwijk, Phys. Rev. B **25**, 4515 (1982), frequently abbreviated as BTK.

obtain for $\Delta \to 0$). As the Andreev reflection essentially contains the transmission of two particles through the interface (the initial electron, and the reflected hole), its probability decays as T^2 for $\epsilon \ll \Delta$ and $T \ll 1$. In a tunnel structure, therefore, it can be neglected.

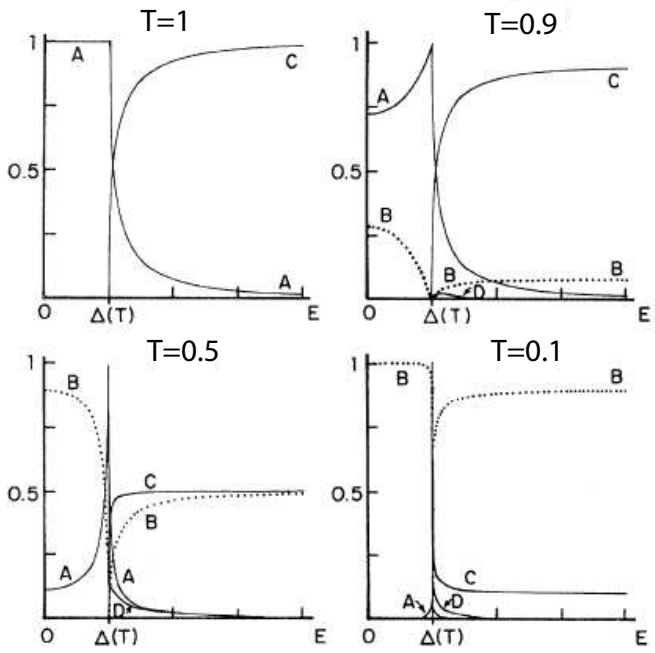


Figure 6.4: Probabilities for Andreev reflection (A), normal reflection (B) and the two types of transmission (C and D) at a NS interface. Adapted from G. E. Blonder, M. Tinkham, and T. M. Klapwijk, Phys. Rev. B **25**, 4515 (1982).

6.4.3 Proximity effect

Due to the Andreev reflection, Cooper pairs can "pop in" the normal metal near the NS interface (see Fig. 6.5). This process induces a finite pairing amplitude $F(\vec{r})$ also inside the normal metal (note that as the coupling constant g is assumed to vanish there, the pair potential stays zero). The inverse effect also exists: because the pairing amplitude vanishes far in the normal metal, that in the superconductor near the interface is diminished. Such a phenomenon is called the superconducting proximity effect.

In most practical cases, the quantitative treatment of the proximity effect has to be carried out with the approach of (quasiclassical) Green's functions.[9]

[9]See, for example, W. Belzig, *et al.*, Superlattices Microstruct. **25**, 1251 (1999).

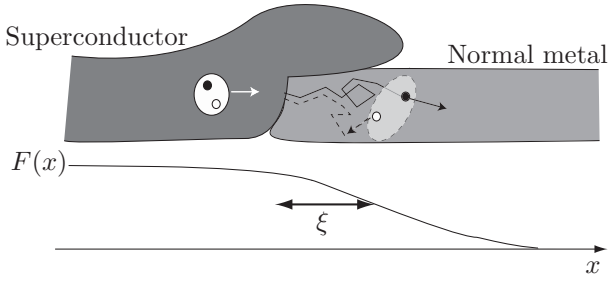


Figure 6.5: Typical normal-superconductor contacts for the study of the proximity effect are deposited as overlap junctions of two materials, for example superconducting aluminum or niobium and normally conducting copper or silver. As a result, Cooper pairs leak from the superconductor to the normal-metal side.

As this is outside the scope of this course, we will not dwell on this subject. However, it is useful to learn at least the basic properties of the proximity effect. It has been shown to slightly increase the local conductivity and decrease the thermal conductivity of the normal metal. It also makes rise to variations in the local density of states near the N-S interface. However, the most important consequence of the proximity effect is the fact that a normal metal sandwiched between two superconductors can carry a supercurrent, and can hence realize a Josephson junction.

6.5 Problems

1. Show that the BCS density of states satisfies

$$\int_{-\epsilon^*}^{\epsilon^*} dE N_S(E) = 2N(0)\epsilon^*, \tag{6.27}$$

   for $\epsilon^* \gg \Delta$, i.e., there are the same total number of states in the superconducting state as in the normal state.

2. Consider a bulk clean superconductor ($\Delta(\vec{r}) = \text{const.}$) at a vanishing magnetic field. Find the eigenenergies and eigenfunctions of Eq. (6.19).

3. Consider a Josephson junction driven with a time-dependent voltage, $V = V_0 + V_1 \cos(\omega_1 t)$. Calculate the corresponding time-dependent phase using the ac Josephson relation, and show that the time-averaged supercurrent,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T dt I(t), \tag{6.28}$$

   is finite for $V_0 = n\hbar\omega_1/2e$.

Chapter 7

Superconducting junctions

The simplest mesoscopic systems with superconducting elements are the tunnel junctions between two superconductors (SIS system) or a superconductor and a normal metal (SIN or NIS). In the previous, there are essentially two regimes of interest: at a voltage-biased junction, the main effect that one can measure is the quasiparticle tunnelling between the superconductors. This process is governed mainly by the superconducting density of states. At a vanishing bias voltage across the junction, or at low voltages in some cases, the Josephson effect is the dominating phenomenon. In the latter system (NIS), only the quasiparticle current can be observed. In the following, we detail these processes.

7.1 Tunnel contacts without Josephson coupling

7.1.1 NIS contact

The current through a NIS contact follows Eq. (1.6) with one of the densities of states assumed constant, $N_R(E) = N_0^R$, and the other one assumed of the form of Eq. (6.5), i.e.,

$$I = \frac{1}{eR_T N_0^S} \int_{-\infty}^{\infty} dE N_S(E)[f(E, T_L) - f(E + eV, T_R)]. \tag{7.1}$$

This cannot be analytically integrated for an arbitrary voltage and temperature, but its numerical evaluation is straightforward, see Fig. 7.1. We can see that in the limit $T \to 0$, the form of the differential conductance dI/dV turns towards the superconducting density of states. Indeed, this is obtained by

$$\frac{dI}{dV} = \frac{1}{R_T} \int_{-\infty}^{\infty} \frac{N_S(E)}{N_0^S} \left[-\frac{\partial f(E + eV)}{\partial (eV)} \right] dE. \tag{7.2}$$

The function $-\partial f(E+eV)/\partial(eV)$ is a Bell-shaped curve peaked at eV and with a width $\sim 4k_BT$. Therefore with $T \to 0$, we get

$$\frac{dI}{dV} \overset{T \to 0}{\to} \frac{N_S(e|V|)}{N_0^S R_T}. \tag{7.3}$$

This allows a fairly easy measurement of the superconducting density of states via the differential conductance.

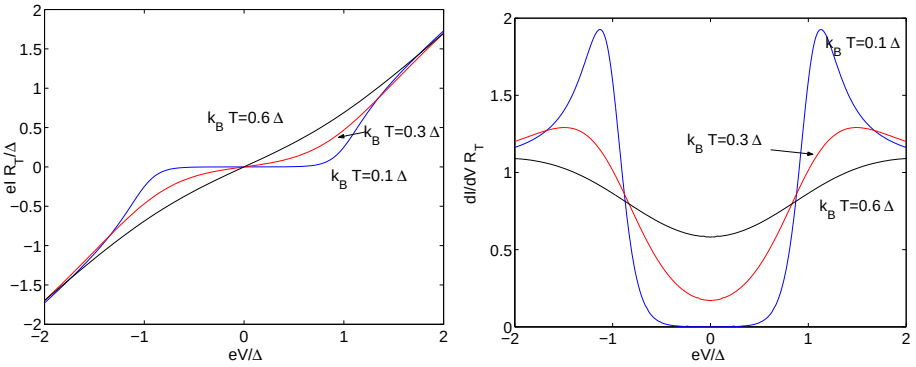


Figure 7.1: Left: Current-voltage curve of a NIS junction at three different temperatures. Right: Corresponding differential conductance.

As the current through the NIS junction at the gap edge is strongly temperature dependent, it can and is used for electron thermometry, or even measuring the shape of the distribution functions.[1]

In the exercises, you will show that the heat current through a NIS junction is a nonmonotonous function of voltage. In particular, the heat current from the normal metal may become positive at a finite bias, i.e., heat can be transferred from the normal metal into the superconductor. Such an effect can be used for electron cooling especially in double NIS systems, i.e., in a SINIS.

7.1.2 SIS contact

In the absence of Coulomb blockade effects and supercurrent, the current through a SIS contact is obtained from

$$I = \frac{1}{eR_T N_0^L N_0^R} \int_{-\infty}^{\infty} dE N_S(E) N_S(E+eV)[f(E,T_L) - f(E+eV,T_R)]. \quad (7.4)$$

Now a current can flow if the voltage exceeds twice the value of the gap. This can be explained using the so-called semiconductor model, see Fig. 7.2: In this picture, the current flows horizontally from left to right, and in order for a finite current to flow, the electron has to leave from an occupied state and enter an unoccupied one. At a low temperature, this is only possible if the difference in the Fermi levels of the two superconductors exceeds 2Δ. The SIS current-voltage curves are shown in Fig. 7.3.

[1] See Sec. III in F. Giazotto, *et al.*, Rev. Mod. Phys. **78**, 217 (2006), and H. Pothier, *et al.*, Phys. Rev. Lett. **79**, 3490 (1997).

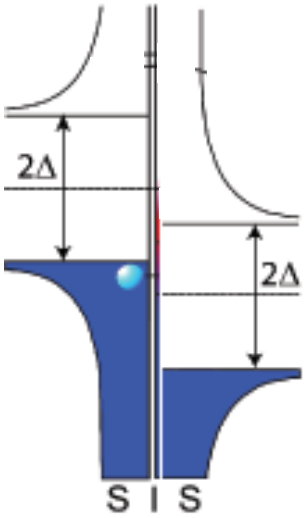


Figure 7.2: Semiconductor model for the SIS junction. At a low temperature, current flow is allowed only if $eV > 2\Delta$.

7.1.3 Superconducting SET

Now let us consider the effects superconductivity induces in single-electron transistors. The lowest-order effects are twofold: Firstly, the presence of the gap in the energy spectrum makes rise to an additional energy scale for the transition rates and for the charging energy corresponding to single-electron (single-quasiparticle) excitations. Secondly, the presence of Josephson coupling makes rise to additional contribution to the current, as the Cooper pairs tunnel through the structure. The latter effect is described in Sec. 7.2.6.

In the second order in tunnel coupling, additional effects (to cotunnelling) arise. These are Andreev reflection in NIS junctions,[2] and Josephson-quasiparticle processes in SISIS-SETs. The latter is schematically explained below.

The modification on the quasiparticle tunnelling rates is fairly obvious. The tunnelling rate for a normal-metal junction, Eq. (5.15), was derived assuming an energy independent density of states. The presence of superconductivity changes the relation to

$$\Gamma_{LI}(n) = \frac{1}{e^2 R_L} \int_{-\infty}^{\infty} d\epsilon_k \int_{-\infty}^{\infty} d\epsilon_q f_L(\epsilon_k) \frac{N_L(E) N_R(E)}{N_L^0 N_R^0} [1 - f_I(\epsilon_q)] \delta(\delta E_{ch} + \epsilon_q - \epsilon_k), \tag{7.5}$$

where $N_{L/R}(E)$ should be replaced by Eq. (6.5) for the superconducting parts of the transistor. This integral can no longer be performed in closed form, but it can be represented using the current $I_t(V)$ of a noninteracting case, Eq. (7.1)

[2] For a description of this process, see Ch. 3 (by Gerd Schön) in the book T. Dittrich, *et al.*, "Quantum transport and dissipation", (Wiley-VCH, 1998).

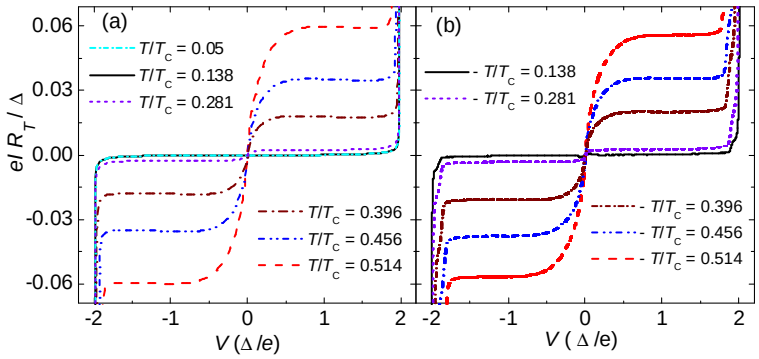


Figure 7.3: (a) Calculated and (b) measured current-voltage curves of a SIS tunnel junction at a few temperatures. Supercurrent has been suppressed. Courtesy of Alexander Savin and Jukka Pekola.

and Fig. 7.1 for the NIS and Eq. (7.4) and Fig. 7.3 for the SIS case:

$$\Gamma_{LI}(n) = \frac{1}{e} I_t \left(\frac{\delta E_{ch}}{e} \right) \frac{1}{\exp[\delta E_{ch}/(k_B T)] - 1}. \tag{7.6}$$

The presence of the gap makes rise to an additional suppression of the quasiparticle tunnelling rates compared to the NIN case. In exercise 7.2, you will find the threshold voltage for the quasiparticle current in different types of superconducting SETs.

For a single junction (NIS or SIS), the environmental Coulomb blockade effect on the quasiparticle current can be found similarly to the NIN case, by replacing the $\delta(E)$-function in Eq. (7.5) by the function $P(E)$ characterizing the environment.

Parity effect

If there is an odd number of conduction electrons in a superconductor, not all of them can be paired to Cooper pairs. The unpaired electron necessarily has an excess kinetic energy of at least Δ (and because of the BCS divergence, most of the excess electron states are near this). Therefore, the charging energy obtains an extra contribution depending on the parity of the electron number,

$$E_{\rm ch}(n, Q_G) = \frac{(ne - Q_G)^2}{2C} + \Delta_n, \tag{7.7}$$

where $\Delta_n = 0$ for even n and $\Delta_n = \Delta$ for an odd n. This function and the corresponding ground state charge are plotted in Fig. 7.4. As a result, the charge, and correspondingly the current is no longer an e-periodic but $2e$-periodic function.

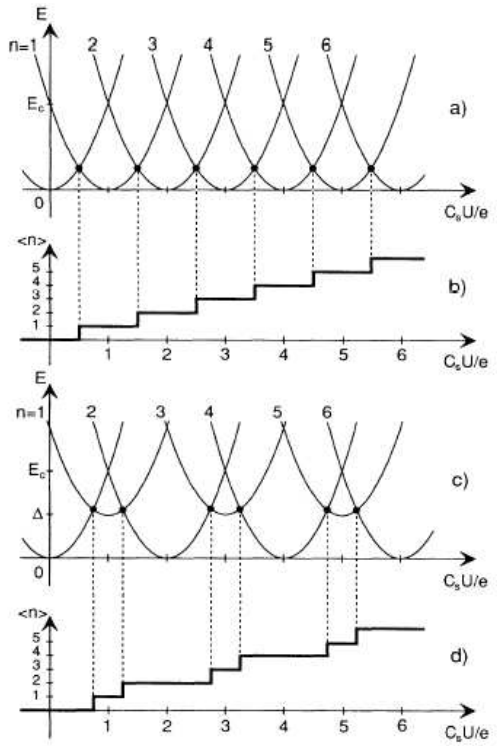


Figure 7.4: Parity effect in the charging energy. From P. Lafarge, *et al.*, Phys. Rev. Lett. **70**, 994 (1993).

The parity effect in the current was first measured by Mark Tuominen, *et al.*[3] and in the charge by P. Lafarge, *et al.* (see Fig. 7.5). This effect can be observed at low temperatures, $T < T_{cr}$, with $T_{cr} \ll \Delta/k_B$. Above T_{cr}, the usual e-periodic conductance is obtained. This can be roughly explained by following what happens to the odd, "unpaired", quasiparticle.[4] This single quasiparticle can tunnel out with a rate which is smaller by a factor $1/N_{\rm eff}$ than the rate Γ of the other $N_{\rm eff}$ electrons, for typical mesoscopic islands $N_{\rm eff} \sim 10^4$. But the rate γ may not be exponentially suppressed compared to the rates for the other electrons, since the excitation energy $\approx \Delta$ is released in this tunnelling process. Hence, $\gamma \approx \Gamma e^{\Delta/(k_B T)}/N_{\rm eff}$. A more precise analysis shows that the number of other electrons is temperature dependent,

$$N_{\rm eff} = N_I^0 \Omega_I \sqrt{2\pi \Delta k_B T}, \tag{7.8}$$

[3] M. Tuominen, Phys. Rev. Lett. **69**, 1997 (1992).

[4] This process is related to the "charge imbalance" in a superconductor, a difference between the chemical potential of the quasiparticles and Cooper pairs inside a single superconductor. See a more detailed explanation in Gerd Schön's text, Ch. 3 in the book T. Dittrich, *et al.*, "Quantum transport and dissipation", (Wiley-VCH, 1998).

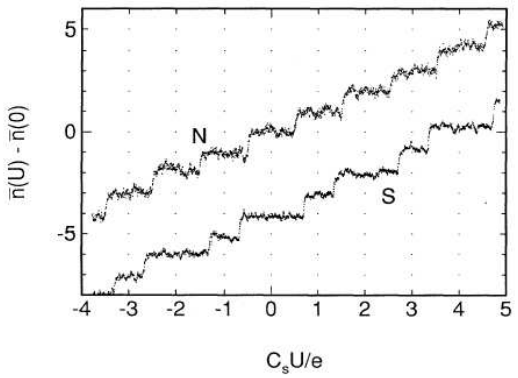


Figure 7.5: Parity effect measured in the charge of a superconducting single-electron transistor island. From P. Lafarge, *et al.*, Phys. Rev. Lett. **70**, 994 (1993).

where N_I^0 is the normal-metal density of states of the island at the Fermi energy, and Ω_I is the volume of the island.

The parity effects survive as long as this single-electron rate is observable, i.e., $\gamma \geq \Gamma$. The crossover temperature can thus be found from $N_{\rm eff}(T_{cr}) = e^{\Delta/(k_B T_{cr})}$. As T_{cr} depends on the volume of the island only logarithmically, for Al islands, this crossover temperature is almost always between 200 and 300 mK.

Important processes which combine quasiparticle and Cooper pair tunnelling in SSS transistors are the Josephson-quasiparticle (JQP) and double Josephson-quasiparticle (DJQP) cycles. In these cycles, a Cooper pair tunnels through one of the junctions, and two quasiparticles tunnel through the other junction (see Fig. 7.6). In the DJQP cycle, this happens twice (Fig. 7.7). These processes are resonant, and thus they make rise to peaks in the current-voltage characteristics. They are especially relevant for the use of superconducting SETs in the measurement of charge qubits.

7.2 Josephson junctions

Josephson junctions are very versatile objects. They can be used to realize highly accurate voltmeters, magnetometers, thermometers, or for radiation detection. Including the charging effects taking place in small Josephson junctions, they have been suggested to use in realization of quantum bits. The above list is definitely not exhaustive, and more examples are found in the book of Tinkham. The versatility of the Josephson junctions lies in the fact that they are nonlinear circuit elements, but often in a very controlled manner (the nonlinearity stems from the dc Josephson relation, Eq. (6.15)).

In what follows, we detail only some of the basic properties of Josephson junctions.

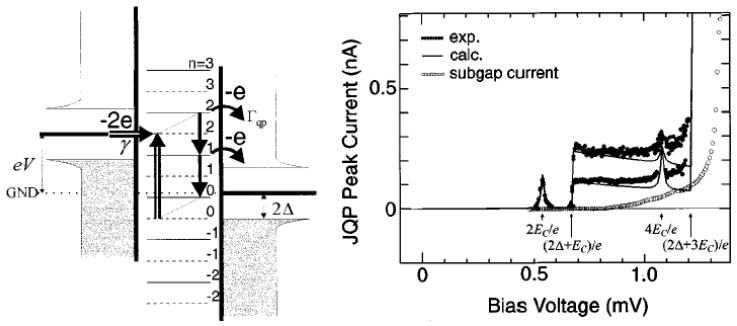


Figure 7.6: Illustration of the Josephson-quasiparticle process. Left: Energy band diagram from which estimates for the occurrence of the JQP peak can be found — when the potential of the left island is resonant with the solid lines, indicating the charging energy difference for an addition of a Cooper pair, and when the two electrons can tunnel into the right island, a JQP process can take place. Right: measured JQP cycle. Revised from Y. Nakamura, C. D. Chen, and J. S. Tsai, Phys. Rev. B **53**, 8234 (1996).

7.2.1 SQUIDs

The discussion on the Josephson effect in the previous chapter was based on the concept of the phase difference φ across the junction, without paying attention how such a phase difference could be created. From the microscopic theory, one finds that the phase is connected to the momentum of the superconducting condensate, i.e., of the Cooper pairs. Hence, a gauge-invariant way to treat the phase difference is to include the effect of the vector potential, such that the total phase difference is

$$\phi = \varphi - \frac{2\pi}{\Phi_S} \int \mathbf{A} \cdot \vec{dl}, \tag{7.9}$$

where the integral is taken from one electrode to the other. Here it is also customary to use the flux quantum for Cooper pairs,

$$\Phi_S \equiv \frac{h}{2e} = \Phi_0/2. \tag{7.10}$$

If the magnetic field is sufficiently high to make rise to a flux exceeding a flux quantum through the junction area, one can observe the *Fraunhofer pattern* in the current-field relation. However, in the following we assume that the fields are much lower.

Consider now a superconducting loop containing two Josephson junctions and connected to an external circuit as shown in Fig. 7.8. This system is called a dc-SQUID, where the acronym SQUID comes from Superconducting QUantum Interference Device. The single-junction version is called an rf-SQUID, but we will not detail its behavior here. In the SQUID, the presence of the magnetic flux makes rise to a phase gradient within the ring. This in turn induces a circulating

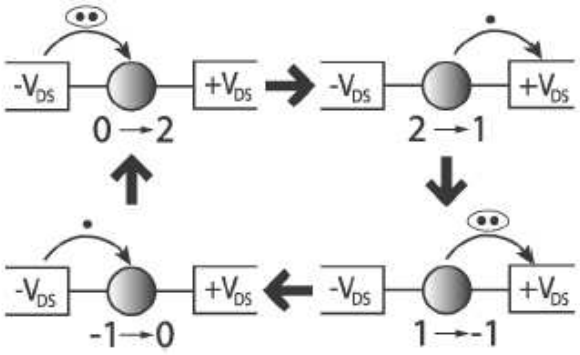


Figure 7.7: Schematic illustration of the DJQP process. Numbers indicate the charge on the superconducting island. From A. A. Clerk, Phys. Rev. Lett. **89**, 176804 (2002).

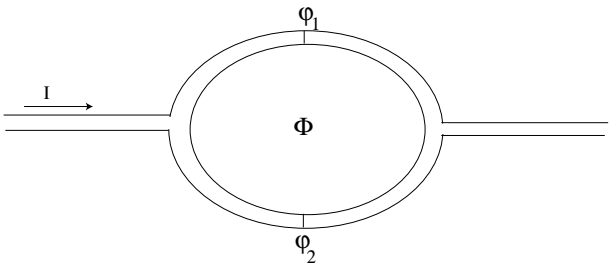


Figure 7.8: DC SQUID.

supercurrent, which screens the flux. The total flux Φ across the ring is then a sum of the applied flux and that produced by the screening current. Now, starting from the upper junction, the total phase difference gathered within one round around the ring is

$$\phi_{\rm tot} = -\varphi_1 + \frac{2\pi}{\Phi_S}\int_1^2 \mathbf{A}\cdot\vec{dl} + \varphi_2 + \frac{2\pi}{\Phi_S}\int_2^1 \mathbf{A}\cdot\vec{dl}. \tag{7.11}$$

Here the first integral is taken around the right side of the ring, and the second around the left. Neglecting the flux exactly on the junctions, we obtain

$$\phi_{\rm tot} = \varphi_2 + \varphi_1 + \frac{2\pi}{\Phi_S}\oint \mathbf{A}\cdot\vec{dl} = \varphi_2 - \varphi_1 + \frac{2\pi\Phi}{\Phi_S}. \tag{7.12}$$

But the phase at a given position should be single-valued (up to 2π), i.e., the total phase difference around the ring should be $\phi_{\rm tot} = 2n\pi$. This means that

$$\varphi_1 - \varphi_2 = \frac{2\pi\Phi}{\Phi_S} \pmod{2\pi}. \tag{7.13}$$

Now the total current through this ring is given by

$$I = I_{c1}\sin(\varphi_1) + I_{c2}\sin\left(\varphi_1 - \frac{2\pi\Phi}{\Phi_S}\right), \tag{7.14}$$

where I_{c1} and I_{c2} are the critical currents of the two junctions. In the symmetric case $I_{c1} = I_{c2} \equiv I_c$, for a given flux Φ, the maximum external current that can flow as a supercurrent through this system is

$$I = 2I_c|\cos\left(\frac{\pi\Phi}{\Phi_S}\right)|. \tag{7.15}$$

Hence, the critical current of the ring can be modulated with a flux. Because of this property, the SQUIDs are very sensitive magnetometers.

7.2.2 Resistively and capacitively shunted junction model

Let us now consider a Josephson junction as a part of an electric circuit. A frequently used model for describing how a finite resistance appears in the junction is called the resistively and capacitively shunted junction model, or RCSJ-model in short. Consider the circuit of Fig. 7.9. The Josephson junction is denoted with the cross inside a box (this is a symbol for a Josephson junction where the capacitance can be neglected - including the capacitance would make rise to an additional line in the middle). This junction is then in parallel with a "shunt" resistance R and its own capacitance C. The resistance R may be due to the quasiparticle current (in which case it would be dependent on the driving current), but it may also be just some additional shunt resistance.

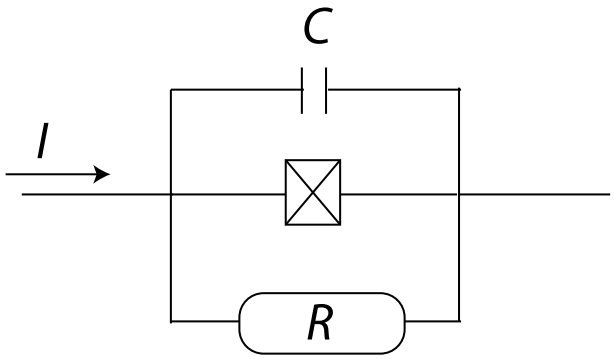


Figure 7.9: Equivalent circuit for the RCSJ model.

Now the driving current I divides in three paths, $I = I_c\sin(\varphi) + I_C + I_R$. In the capacitive path, the current is given by $I_C = \dot{q} = C\dot{V}$, and in the resistive path, it is $I_R = V/R$, where V is the voltage over the circuit elements. Using the ac Josephson relation, Eq. (6.17), this equality can also be written as

$$I = I_c\sin(\varphi) + \frac{\hbar}{2e}\dot{\varphi}/R + C\frac{\hbar}{2e}\ddot{\varphi}. \tag{7.16}$$

Now let us make this equation dimensionless, by introducing a dimensionless time $\tau = \omega_p t$, the "plasma frequency"

$$\omega_p = \sqrt{2eI_c/\hbar C} \tag{7.17}$$

and the "quality factor"

$$Q = \omega_p RC. \tag{7.18}$$

With these definitions, Eq. (7.16) turns into

$$\ddot{\varphi} + Q^{-1}\dot{\varphi} + \sin(\varphi) = I/I_c. \tag{7.19}$$

There is a mechanical analog for this model. This equation can be seen as the equation of motion for a particle with mass $(\hbar/(2e))^2 C$ moving along the φ-axis in an effective potential

$$U(\varphi) = -E_J \cos(\varphi) - (\hbar I/(2e))\varphi, \tag{7.20}$$

where $E_J = \hbar I_c/(2e)$ is the *Josephson energy*. Because of its shape, this potential is called the "tilted washboard potential" (see Fig. 7.10). One can also notice that the second term in Eq. (7.19) describes "friction", which is consistent with an idea of the resistor as a dissipative element.[5]

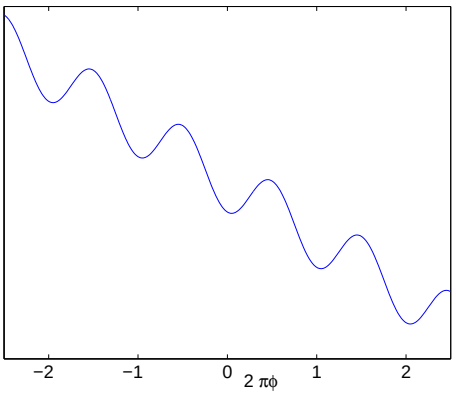


Figure 7.10: Tilted washboard potential.

Generally, the RCSJ model can be separated into "overdamped" ($Q < 1/2$) and "underdamped" ($Q > 1/2$) regimes, based on the behavior of the phase within a single well of the potential (7.20). Assume we drive the phase to some initial value φ_0 with the external current, and then turn the external current suddenly off. If φ_0 is not $2n\pi$, corresponding to a minimum of $-E_J \cos(\varphi)$, the phase starts to do damped oscillations in this potential, approaching the potential minimum. In the underdamped regime, it oscillates many times before decaying into the minimum, and in the overdamped regime, the decay time is less than the oscillation period.

[5] Although, note that the friction is the smaller the higher the resistance is.

7.2.3 Overdamped regime

In the overdamped regime, the behavior of the Josephson junction is regular, and one can derive an equation for the average voltage across the junction (see exercise 7.3),

$$\langle V \rangle = R\sqrt{I^2 - I_c^2}\,\theta(|I| - I_c)\mathrm{sgn}(I). \tag{7.21}$$

This current-voltage curve is plotted in Fig. 7.11 and it is valid at a vanishing temperature. At a finite temperature, there is a finite probability for the junction to "escape" into the finite-voltage state at some lower current than I_c (see below). This process makes the resulting $I - V$ curve smoother, and the measured critical current (i.e., maximum current with which no voltage appears) lower than I_c.

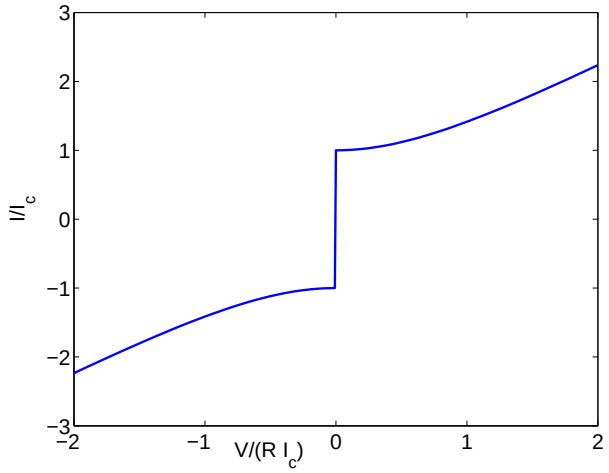


Figure 7.11: $I - V$ curve of an overdamped Josephson junction.

7.2.4 Underdamped regime

When the quality factor is much larger than $1/2$, the $I - V$ curves become hysteretic. Upon increasing the current I from zero, the junction remains in the supercurrent-carrying state until some critical current close to I_c, at which point it jumps discontinuously (at $T = 0$) to some finite value, which turns out to be $2\Delta/e$. If I is now reduced, the junction does not enter the zero-voltage state before the current reaches a "retrapping current" $I_r \approx 4I_c/(\pi Q)$. This hysteretic behavior can be roughly explained with the mechanical analogue: Assume a particle is located in one of the potential wells. Tilting the potential, the particle escapes from the well only once the potential minima no longer are minima. The "running state" of the particle now corresponds to the finite voltage. Now leveling the potential back towards the stable position does not immediately retrap the particle in one of the wells. The physical reason for the

dependence $\propto 1/Q$ is that for retrapping, the energy fed by the driving current as the phase advances by 2π has to be dissipated at the same time.

7.2.5 Escape process

The above description of the Josephson junction was effectively a zero-temperature description. At a finite temperature, the thermal fluctuations of the current will cause an extra driving term in Eq. (7.19),

$$\ddot{\varphi} + Q^{-1}\dot{\varphi} + \sin(\varphi) = I/I_c + \delta I(t)/I_c. \tag{7.22}$$

This is a Langevin-type description, where the fluctuations are treated with the quantity $\delta I(t)$. This satisfies

$$\langle \delta I(t) \rangle = 0 \tag{7.23}$$

and

$$\langle \delta I(t)\delta I(0) \rangle = \tilde{S}_I(t), \tag{7.24}$$

where the average is taken over the realizations of the fluctuations, or in practice, over time. The term $\tilde{S}_I(t)$ characterizes the fluctuations, and by the fluctuation-dissipation theorem (derived in the following chapter), in thermal equilibrium its Fourier transform satisfies

$$S_I(\omega) \equiv \int dt e^{i\omega t} \tilde{S}_I(t) = \frac{2k_B T}{R}. \tag{7.25}$$

This is valid provided that $\omega \ll k_B T/\hbar$.

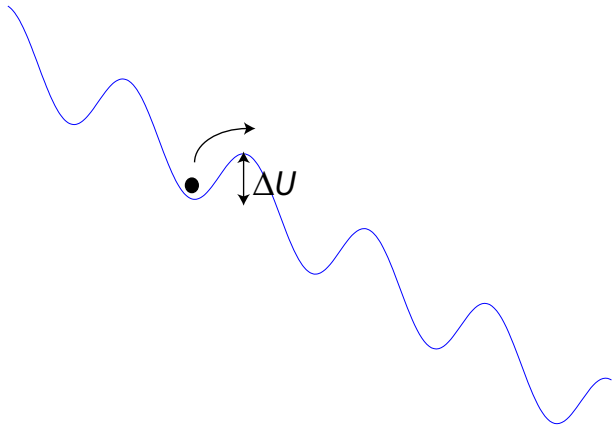


Figure 7.12: Thermal escape of the "phase particle" from a potential well.

The thermal fluctuations make it possible for the phase to thermally escape from the potential well, as schematically illustrated in Fig. 7.12. The rate for this escape process can be written as

$$\Gamma(I) = \frac{\omega_A(I)}{2\pi} \exp\left(-\frac{\Delta U(I)}{k_B T}\right). \tag{7.26}$$

Here $\omega_A \sim \omega_p$ is the frequency that characterizes the dynamics of the "particle" inside a single potential well. The frequency $\omega_A/(2\pi)$ is hence an "attempt frequency" for the escape. On each attempt, there is a probability $\exp(-\Delta U/(k_B T))$ of success, i.e., this is the probability that the instantaneous value of the current (including the fluctuations) exceeds the critical current. Now the attempt frequency has to be taken for the tilted potential, and it reads

$$\omega_A = \omega_p[1 - (I/I_c)^2]^{1/4}. \tag{7.27}$$

The height of the potential well depends on the driving current, vanishing at I_c. Its expression is

$$\Delta U = \frac{8\sqrt{2}}{3} E_J \left(1 - \frac{I}{I_c}\right)^{3/2}. \tag{7.28}$$

This escape probability can be directly measured, as each escape of the junction produces an observable voltage pulse, of the order of $2\Delta/e$. Thus, measuring the state of the junction for a time τ, the probability for escape is

$$P_e(\tau; I) = \exp(-\Gamma(I)\tau). \tag{7.29}$$

A typical escape curve is an S-shaped curve, where P increases from zero to unity within a width $\sim k_B T$ slightly below I_c.

As the escape probability is strongly dependent on the strength of current fluctuations, it can be also used to measure them. The first such measurement was reported very recently, see J. Pekola, *et al.*, Phys. Rev. Lett. **95**, 197004 (2005). The escape process has also been used to measure quantum bits, where the two states of the bit can be translated into two different bias currents over the Josephson junction.[6]

As a result of the finite escape probability below the critical current, the average current-voltage curves will be smeared. The actual details of the smearing depend on the value of the Q-factor, and we will not detail them here.

7.2.6 Quantum effects in small junctions

In small Josephson junctions, the charging energy for the Cooper pairs tunnelling through the junctions will become essential. In this case, a fully quantum-mechanical description of the junction dynamics is necessary. The first question in this case is on the form of the Hamiltonian. The potential energy of the Josephson junction was above shown to be of the form of Eq. (7.20). The kinetic energy requires the identification of a canonical conjugate variable for the phase. This turns out to be the polarization charge Q on the junction, and the corresponding kinetic energy is the charging energy. Thus, the total Hamiltonian is of the form

$$H = \frac{Q^2}{2C} - E_J \cos(\varphi) - \frac{\hbar}{2e} I_b \varphi, \tag{7.30}$$

[6] See, for example, D. Vion, *et al.*, Science **296**, 886 (2002).

where I_b is the bias current. This conjugate pair of variables can be justified through the Hamilton equations for the Josephson junction (see exercise 7.5), which produce the ac and dc Josephson relations. The theory for the quantized Josephson junction is now defined by assuming that φ and Q are operators that satisfy the commutation relation,

$$[\varphi, Q] = 2ei. \tag{7.31}$$

The charge on this equation is naturally the charge of a Cooper pair.

Now the mechanical analogue of the tilted washboard potential turns into an analogue of a quantum-mechanical description of a particle in a periodic potential, following the "dictionary" given in Table 7.1.

Particle	Josephson junction
$H = \frac{p^2}{2m} - U\cos\left(\frac{x}{a}\right) - Fx$	$H = \frac{Q^2}{2C} - E_J\cos(\varphi) - \frac{\hbar}{2e}I_b\varphi$
coordinate x	phase φ
momentum $p = -\frac{\hbar}{i}\partial_x$	$\propto$ charge $\frac{\hbar Q}{2e} = -2ei\partial_\varphi$
velocity $v = \frac{dx}{dt} = \frac{p}{m}$	$\propto$ voltage $\frac{2eV}{\hbar} = \frac{\partial\varphi}{\partial t} = \left(\frac{2e}{\hbar}\right)^2 \frac{1}{C}\frac{\hbar Q}{2e}$
mass m	$\propto$ capacitance $\left(\frac{\hbar}{2e}\right)^2 C$
force F	$\propto$ bias current $\frac{\hbar}{2e}I_b$.

Table 7.1: Translation between the quantum theory of a particle in a periodic potential and the quantum theory of a Josephson junction.

The state of the junction is characterized by a macroscopic wave function Ψ. In the time-independent case it satisfies the Schrdinger equation

$$-\frac{(2e)^2}{C}\frac{\partial^2\Psi_n}{\partial\varphi^2} - E_J\cos(\varphi)\Psi_n = E_n\Psi_n. \tag{7.32}$$

This equation is in mathematics called the Mathieu differential equation and its eigenfunctions Ψ_n are Mathieu functions.[7] Below, we only consider the eigenenergies in two opposite limits, $E_J \gg E_c$ and $E_J \ll E_c$.

7.2.7 "Tight-binding limit"

In the tight-binding limit $E_J \gg E_c$, the lowest eigenstates are peaked around the minima of the potential well, i.e., $\varphi \approx 2n\pi$. In this case, we may find the eigenenergies by expanding the potential around one of the minima, $\cos(\varphi) \approx 1 - \varphi^2/2$. As a result, we get the Hamiltonian of a Harmonic oscillator,

$$H = -4E_c\partial_\varphi^2\Psi_n + \frac{1}{2}E_J\varphi^2\Psi_n = E_n\Psi_n. \tag{7.33}$$

[7] See http://mathworld.wolfram.com/MathieuDifferentialEquation.html and http://mathworld.wolfram.com/MathieuFunction.html.

The solutions can be written in terms of the Hermite polynomials as explained in the basic courses on quantum mechanics. The eigenenergies are

$$E_n = (n + 1/2)\hbar\omega_p, \tag{7.34}$$

where ω_p is the plasma frequency (see Exercise 7.4 and Fig. 7.13).

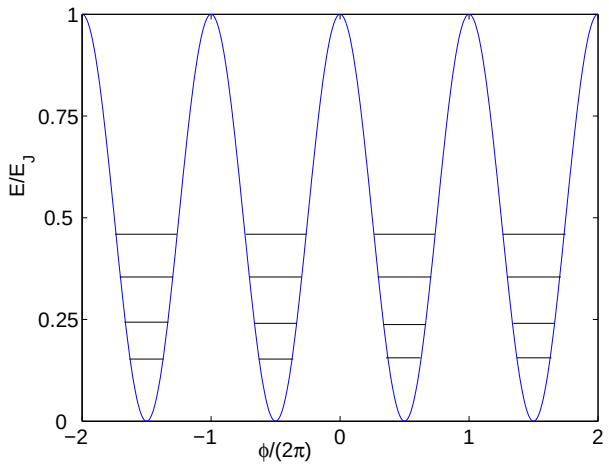


Figure 7.13: Energy levels in the tight-binding limit $E_J \gg E_C$.

7.2.8 "Nearly free-electron limit"

In the weak-coupling case $E_J \ll E_c$, it is useful to start the description from the case of a completely isolated charge, $E_J = 0$. In this case, the solutions to Eq. (7.32) are plane waves corresponding to different fixed charges Q,

$$\Psi_Q(\varphi) = A_Q e^{i\frac{Q}{2e}\varphi}. \tag{7.35}$$

This is the analog of the electron plane wave $\sim e^{i\vec{p}\vec{x}}$. The charge Q is generally of the form $Q = 2en + q$, where q stands for the polarization charge. For the different n we again get the Coulomb parabola, depicted in Fig. 7.14.

The polarization charge can be controlled in a geometry of a Cooper pair box, i.e., the single-electron box (see Fig. 5.1) replaced by a superconductors, with a finite Josephson coupling between the island and the left lead. Therefore, we concentrate on this case below.

It turns out that the introduction of a small $E_J \ll E_c$ makes rise to an "avoided crossing" of the energy levels corresponding to the different n. In this limit, it is useful to write the Hamiltonian in terms of the basis functions corresponding to the different charge states $|n\rangle$. This Hamiltonian is obtained by calculating the matrix elements of the original Hamiltonian, Eq. (7.30), in

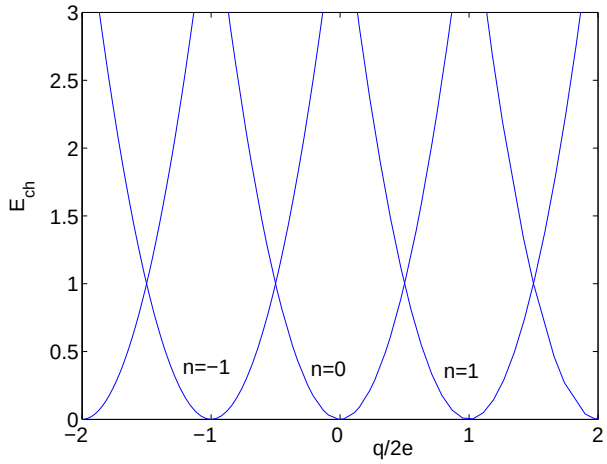


Figure 7.14: Coulomb parabola for Cooper pairs in the case $E_J = 0$.

the basis defined by the charge states. With this scheme, we get

$$\langle n|\frac{Q^2}{2C}|m\rangle = 4E_C(n-n_g)^2\delta_{nm}, \tag{7.36}$$

directly from the definition of Q. The Josephson term can be written as $E_J(e^{i\varphi}+e^{-i\varphi})/2$. The matrix element of these operators can be found in the "real-space" representation of the wave functions,

$$\begin{aligned}\langle n|e^{\pm i\varphi}|m\rangle &= A_nA_m\int e^{-i(n-n_g)\varphi}e^{\pm i\varphi}e^{i(m-n_g)\varphi}d\varphi \\ &= A_nA_m\int e^{i(m\pm1-n)\varphi} = 2\pi A_nA_m\delta_{n,m\pm1}.\end{aligned} \tag{7.37}$$

As plane waves are not normalizable, the correct normalization would require introducing some distant boundary conditions, and then including the normalization. This question is discussed in the basic quantum-mechanics courses, and we will not dwell on it here. From such an analysis one would get that the prefactor in the previous expression equals unity. With these matrix elements, we can use the resolution of unity and get the desired form for the Hamiltonian,

$$H = \sum_{n,m}|n\rangle\langle n|H|m\rangle\langle m| = \sum_n 4E_C(n-n_g)^2|n\rangle\langle n| - \frac{1}{2}E_J(|n\rangle\langle n+1|+|n+1\rangle\langle n|). \tag{7.38}$$

This can now be reduced further to analyze the spectrum at $E_J \ll E_C$.

Near the crossing point of the energies of two charge states, $n_g \approx 1/2$, the Hamiltonian can be described by the two-state model, involving only the two crossing charge states, say $|0\rangle$ and $|1\rangle$. In this basis, the Hamiltonian is a 2×2

matrix, specified by the Pauli matrices,

$$H = -\frac{1}{2}B_z\sigma_z - \frac{1}{2}B_x\sigma_x. \tag{7.39}$$

Here the "magnetic fields" are $B_z = 4E_C(1-2n_g)$ and $B_x = E_J$. The eigenfunctions of this Hamiltonian are

$$\begin{aligned} |e1\rangle &= \cos(\eta/2)|0\rangle + \sin(\eta/2)|1\rangle \\ |e2\rangle &= -\sin(\eta/2)(\eta/2)|0\rangle + \cos(\eta/2)|1\rangle, \end{aligned} \tag{7.40}$$

where $\eta = \arctan(B_x/B_z)$. The corresponding eigenenergies are $\pm\sqrt{B_x^2+B_z^2}/2 = E_J/(2\sin(\eta))$.

It is straightforward to see that in the limit $n_g \ll 1/2$, the ground state goes towards the charge state $|0\rangle$ and the excited state towards $|1\rangle$, and vice versa for $n_g \gg 1/2$. At the degeneracy point $n_g = 1/2$, the difference between the eigenenergies is E_J, which is then also the splitting of the two energy bands.

This type of two-state description allows the use of superconducting SETs as quantum bits.[8]

7.2.9 Dynamic capacitance

The band model of a Josephson junction can be also used to define a "dynamic capacitance", analogous to the effective mass m_e in the theory of electrons in a lattice. This dynamic capacitance is defined as

$$\frac{1}{C^*} = \frac{\partial^2 E_0}{\partial q^2}, \tag{7.41}$$

where q is the polarization charge, and E_0 is the ground-state energy. As this dynamic capacitance depends on the gate charge, it can be used as a tunable capacitance, for example in LC-circuits.[9]

7.3 Problems

1. Show that the heat current from a normal metal through a NIS junction is at finite temperatures ($0 < k_BT \ll \Delta$) a non-monotonous function of voltage (for $V > 0$). Hint: You can evaluate it numerically, but try to show it also analytically by considering the limit of a small voltage, $eV \ll \Delta$ (and expanding to the second order), and on the other hand a very large voltage, $eV \gg \Delta$.

2. At $T = 0$, find the threshold voltage where quasiparticle current starts to flow (in the sequential tunnelling regime) for a) NSN, b) SNN, c) SSN, d) SNS, and e) SSS superconducting SETs (for example, NSN SET refers to a SET composed of normal-metal electrodes and a superconducting island).

[8] See Yu. Makhlin, A. Shnirman, and G. Schön, Rev. Mod. Phys. **73**, 357 (2001).

[9] For its observation, see a very recent paper by M. Sillanpää, *et al.*, Phys. Rev. Lett. **95**, 206806 (2005).

3. Consider the RCSJ model in the strongly overdamped limit, $Q \ll 1$. In this case, you can neglect the first term in Eq. (7.19). From the resulting equation, find the average voltage across the junction as a function of the driving current, both for $I < I_c$ and for $I > I_c$. Hint: Separate the coordinates φ and τ and integrate over one period $[0, 2\pi]$ in the phase. Use this for the averaged phase.

4. a) In the linear regime, show that a Josephson junction represents an inductance. Find this inductance. b) Combine this inductance with the capacitance C of the junction, and find the resonant frequency ω_0 of the resulting LC-oscillator. c) Express ω_0 using the Josephson energy E_J and the charging energy E_c. Use this to justify Eq. (7.34).

5. Show that the Hamilton equations (the equations describing the dynamics of the canonical coordinates) for the Josephson junction in the non-driven case (Eq. (7.30)) produce the dc and ac Josephson relations.

Chapter 8

Fluctuations and correlations

In the above chapters, we already mentioned a few times the fluctuations of currents or voltages in the systems we describe. These fluctuations are in fact an everyday phenomenon for any electronic engineer, or for example someone trying to analyze the resolution of electronic measuring equipment. In those cases the fluctuations act as "noise", i.e., trying to hinder the actual operation based on the average quantities. In this chapter we describe this noise, tell how it can be quantitatively characterized and show how it can actually be used to obtain further information about the studied objects.[1]

8.1 Definition and main characteristics of noise

Consider a time-resolved current measurement with a constant bias voltage source (Fig. 8.1)

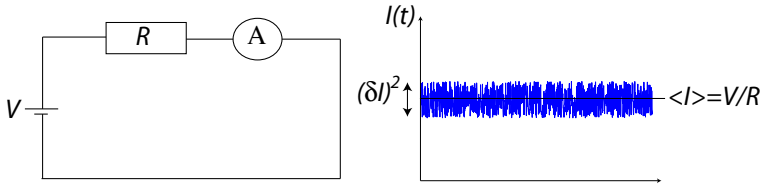


Figure 8.1: Time-resolved current measurement with a constant voltage bias.

[1] Fluctuations in different types of systems are treated in the book by Kogan: "Electronic noise and fluctuations in solids" (Cambridge, 1996). The standard reference for shot noise is the review by Yaroslav Blanter and Marcus Büttiker, "Shot noise in mesoscopic conductors", Phys. Rep. **336**, 1 (2000). Moreover, the basic phenomena, such as equilibrium fluctuations described by fluctuation-dissipation theorem, are described in most books on advanced-level statistical mechanics.

The magnitude of the current fluctuations ("noise") around the average may be described through the correlation function

$$\begin{aligned} S_I(t,t') &= 2\langle \delta I(t)\delta I(t')\rangle, \quad \text{"classical current"} \\ &= \langle \{\delta\hat{I}(t), \delta\hat{I}(t')\}\rangle, \quad \text{"quantum current"}, \end{aligned} \tag{8.1}$$

where $\delta\hat{I} = \hat{I}(t) - \langle\hat{I}(t)\rangle$. Here the brackets $\langle\cdot\rangle$ denote an ensemble average over all the possible (microscopic) states of the system. Usually this is equivalent to an average over repeated measurements, or a time average (this is the ergodic hypothesis).

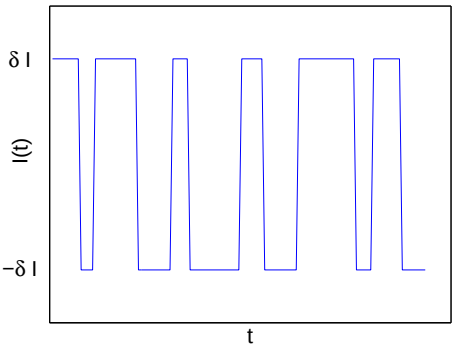
In a stationary system (where the ensemble-averaged current $\langle I\rangle$ does not vary in time), $S(t,t')$ depends only ont he time difference between the events, $S(t,t') = S(t'-t)$. In this case it is customary to describe the Fourier transform of $S(t'-t)$:

$$S(\omega) \equiv \int_{-\infty}^{\infty} d(t'-t) e^{i\omega(t'-t)} S(t'-t). \tag{8.2}$$

In the following, we will concentrate on the frequency domain, and thus there should be no confusion between the different functions $S(\omega)$ and $S(t'-t)$. The function $S(\omega)$ is often called the noise power spectral density, and according to the Wiener-Khintchine theorem, it describes the squared deviation $\delta I(\omega)^2$ of the signal around the frequency ω (see exercise 8.1).

Example: "two-level current"

To clarify the notion of the correlation function and its ensemble average, let us consider a fictitious system where, for some reason, the current may have only either the value $I_+ = +\delta I$ or $I_- = -\delta I$. Which value the current has on a given moment depends on the microscopic features of the system and the value of the current may change at any time, independent of what the states were in a prior time. Thus, the fluctuations are of the form



This type of noise is called telegraph noise. Let us now calculate $S(t,t')$. For $t' \neq t$ we have

$$\begin{aligned} S_I &\equiv 2\langle \delta I(t)\delta I(t')\rangle = 2\sum_{n,n'=\pm 1} p_n p_{n'}(n\delta I)(n'\delta I) \\ &= 2(\delta I)^2\left(\frac{1}{2}\right)^2 (1\cdot 1+(-1)\cdot 1+1\cdot(-1)+(-1)\cdot(-1)) = 0. \end{aligned} \tag{8.3}$$

Here we denoted the probability for state n by p_n, and assumed equal probabilities for both states, $p_{+1} = p_{-1} = 1/2$. The sum over the n and n' is due to the ensemble average, which goes over all possible initial and final states at $t = t$, $t = t'$.

For $t = t'$ we have to have $n = n'$ such that

$$S_I(t,t') = 2\sum_{n=\pm 1} p_n(n\delta I)(n\delta I) = 2(\delta I)^2\frac{1}{2}(1\cdot 1+(-1)\cdot(-1)) = 2(\delta I)^2. \tag{8.4}$$

Thus, we obtained

$$S_I(t,t') = 2(\delta I)^2\delta(t-t'). \tag{8.5}$$

Fourier transforming, we would get

$$S_I(\omega) = 2(\delta I)^2, \tag{8.6}$$

independent of frequency ω. This kind of noise is called "white" and it reflects the fact that we assumed that the correlations between the currents $I(t)$ and $I(t')$ disappear for any, however small time difference $t'-t$.[2] In practice, there always is a (small) time scale τ during which such transitions between different current-carrying states may take place. This makes rise to a finite width $\sim \tau$ for $S_I(t'-t)$, and correspondingly, it makes $S_I(\omega)$ frequency dependent at large frequencies $\omega \sim 1/\tau$.

8.1.1 Motivations for the study of noise

Typically, noise is a feature that limits the accuracy of the detected average signal ($\langle I\rangle$), and hence it can be an *undesired property* that one tries to minimize. Attempts for such minimization are already of course a fair motivation for its study (for example, one can show that for any finite-frequency measurements, devices show an unavoidable noise source, "quantum" or "vacuum fluctuations", that does not vanish even at low temperatures — conversely, using low-frequency measurements also does not help due to the so-called $1/f$-noise which is a consequence of the fact that all the microscopic features of the measurement setup are never fully known and controlled). However, the study of *shot noise* has shown that even the noise itself may *contain information on the measured system*. Thirdly, the *descriptions of the interaction between small quantum systems and their environment*, for example an atom in a photon field,

[2] Such a process without memory is called "Markovian".

employ the concept of noise (although then, one often talks about correlations rather than noise). Such a description is especially important in quantum computing as noise provides the dephasing of the quantum systems, but also is the studied feature in the descriptions of quantum measurements.

The main types of (current) fluctuations are[3]

- Thermal noise $S(\omega) = 4\frac{k_B T}{R}$
- Shot noise $S(\omega) = 2eFI$, where F is the "Fano factor"
- Quantum fluctuations $S(\omega \gg \max(k_B T, eV)/\hbar) = 2\frac{\hbar\omega}{R}$ and
- Correlations at some given frequency ω_0, reflecting some energy or time scale relevant for the studied system.

8.1.2 Fluctuation-dissipation theorem

Fluctuations of equilibrium linear systems can be quantified through the fluctuation-dissipation relation. Derivations of this relation are based on the theory of linear systems, (Kubo) linear response theory. This describes a system characterized by the (for simplicity) time independent Hamiltonian H_0. Observables of this system are denoted by A_i^0 and the corresponding operators by $\hat{A}_i^0$, such that

$$A_i^0 = \langle \hat{A}_i^0 \rangle = \mathrm{Tr}[\rho(t)\hat{A}_i^0] = \mathrm{Tr}[\rho(0)\hat{A}_i^0(t)], \tag{8.7}$$

where in the last form, the operator $\hat{A}_i^0(t)$ is expressed in the Heisenberg picture and in equilibrium, $\rho(0) = \exp(-\beta H_0)$ with $\beta = 1/(k_B T)$.

Assume now we apply a "force" $f(t)$ that acts on the observable A_j^0. The resulting Hamiltonian becomes

$$H = H_0 - f(t)\hat{A}_j^0. \tag{8.8}$$

The force $f(t)$ is assumed to act in this system as a scalar function (however, it can still be an operator of some other system, and hence one should be careful when trying to commute $f(t)$ at different times), such that it commutes with the operators $\hat{A}_i^0$. Then, the dynamics of the observable $A_i(t)$ changes in time according to Eq. (1.26),

$$\frac{d}{dt}\hat{A}_i(t) = \frac{i}{\hbar}[H_0 - f(t)\hat{A}_j, \hat{A}_i(t)]. \tag{8.9}$$

Now the solution to Eq. (8.9) can be expanded in powers of $f(t)$. In the zeroth order, we have nothing but $\hat{A}_i^0(t) = \exp(-iH_0 t/\hbar)\hat{A}_i(0)\exp(iH_0 t/\hbar)$. In the first order, this may be replaced instead of $\hat{A}_i(t)$ in Eq. (8.9), and similarly we may replace $\hat{A}_j \to \hat{A}_j^0$ to get

$$\frac{d}{dt}\hat{A}_i(t) = \frac{i}{\hbar}[H_0, \hat{A}_i^0(t)] - f(t)\frac{i}{\hbar}[\hat{A}_j^0(t), \hat{A}_i^0(t)] + o(f(t)^2). \tag{8.10}$$

[3] In this text, we concentrate on the current fluctuations. However, the concept is very general, and we could use the description for the fluctuations of any observable.

A similar procedure for higher orders in $f(t)$ would yield commutators of the form $[\hat{A}_j^0(t),[\hat{A}_j^0(t),\hat{A}_i^0(t)]]$, and so on. Now, a linear quantum system is characterized by the property that the operators satisfy $[\hat{A}_j^0(t),\hat{A}_i^0(t)]$ is a complex number (function of time), and its further commutators with $\hat{A}_j^0$ vanish. Thus, for such systems we may terminate the series to first order and Eq. (8.10) is exact. Otherwise, one has to require that $f(t)$ is small such that higher orders do not contribute much.

Solving Eq. (8.10) formally yields

$$A_i(t) \equiv \langle \hat{A}_i(t)\rangle = \langle A_i^0(t)\rangle + \frac{i}{\hbar}\int_{-\infty}^{t} dt'[\hat{A}_i^0(t),\hat{A}_j^0(t')]f(t'). \tag{8.11}$$

Now defining a response coefficient (or "susceptibility")

$$\chi_{ij}(t,t') \equiv \frac{i}{\hbar}\langle[\hat{A}_i^0(t),\hat{A}_j^0(t')]\rangle\theta(t-t') \tag{8.12}$$

allows us to represent the time dependence of the observable A_i by

$$A_i(t) = A_i^0(t) + \int_{-\infty}^{\infty} dt'\chi_{ij}(t,t')f(t'). \tag{8.13}$$

In a stationary system, $\chi(t,t') = \chi(t-t')$. Then Eq. (8.13) is a convolution. Hence, in the Fourier transformed space it is

$$A_i(\omega) = A_i^0(\omega) + \chi_{ij}(\omega)f(\omega), \tag{8.14}$$

where $A_i(\omega)$, $\chi_{ij}(\omega)$ and $f(\omega)$ are the Fourier transforms of $A_i(t)$, $\chi_{ij}(t-t')$ and $f(t)$. Thus, the change in the observable $A_i(\omega)$ due to the applied force $f(\omega)$ is described through the susceptibility $\chi(\omega)$. To be precise, only the imaginary part of $\chi(\omega)$ describes a dissipative process and the real part is called the "reactive" part.

Now let us consider the noise correlation function,

$$S_{ij}(t) \equiv \langle \hat{A}_i(t)\hat{A}_j(0)\rangle. \tag{8.15}$$

This is not yet the symmetrized version we introduced above, but we may symmetrize it at the end of the calculation. For simplicity, we prove the FD relation in the case $i=j$, as this is what we mostly need and then the proof is simpler. We will then drop the indices from S_{ij} and χ_{ij} altogether.

Let us first note that in the stationary case $S(t)$ follows the symmetry $S(t)^* = \langle \hat{A}_i(0)\hat{A}_i(t)\rangle = \langle \hat{A}_i(-t)\hat{A}_i(0)\rangle = S(-t)$. Then separate the real and imaginary parts of the correlator,

$$S(t) = B(t) + iC(t) \tag{8.16a}$$

$$B(t) = B(-t) = \frac{1}{2}\left(S(t)+S(-t)\right) \tag{8.16b}$$

$$C(t) = -C(-t) = \frac{1}{2i}(S(t)-S(-t)) = \frac{1}{2i}\langle[A_i(t),A_i(0)]\rangle. \tag{8.16c}$$

From the last equality we get $\chi(t) = -\frac{2}{\hbar}\theta(t)C(t)$. Now the Fourier transform of χ satisfies

$$\chi(\omega) = -\frac{2}{\hbar}\int_0^\infty dt e^{i\omega t}C(t) = -\frac{2}{\hbar}\int_{-\infty}^0 dt e^{-i\omega t}C(-t) = \frac{2}{\hbar}\int_{-\infty}^0 dt e^{-i\omega t}C(t). \tag{8.17}$$

But we also have

$$C(\omega) = \int_{-\infty}^\infty dt e^{i\omega t}C(t) = -\frac{\hbar}{2}\left[\chi(\omega) - \chi(-\omega)\right] = -i\hbar \mathrm{Im}[\chi(\omega)]. \tag{8.18}$$

The last relation follows from the fact that $\chi(t)$ is real (which in turn comes from the reality of $C(t)$).

For an equilibrium state, the correlator $S(t)$ satisfies

$$\begin{aligned} S(t) &= \mathrm{Tr}[e^{-\beta H_0}e^{iH_0 t/\hbar}\hat{A}_i(0)e^{-iH_0 t/\hbar}\hat{A}_i(0)] \\ &= \mathrm{Tr}[\hat{A}_i(0)e^{-\beta H_0}e^{iH_0(t-i\hbar\beta)/\hbar}\hat{A}_i(0)e^{-iH_0(t-i\hbar\beta)/\hbar}] \\ &= \langle A_i(-t-i\hbar\beta)A_i(0)\rangle = S(-t-i\hbar\beta). \end{aligned} \tag{8.19}$$

Here we used the cyclic property of the trace and a simple reordering of the exponents. Correspondingly, the Fourier transform yields the detailed balance relation

$$S(\omega) = \int dt e^{i\omega t}S(-t-i\hbar\beta) = \int ds e^{-i\omega s}e^{\beta\hbar\omega}S(s) = e^{\beta\hbar\omega}S(-\omega). \tag{8.20}$$

Finally, we get for the Fourier transform of $C(\omega)$

$$C(\omega) = \frac{1}{2i}(S(\omega) - S(-\omega)) = \frac{1-e^{-\beta\hbar\omega}}{2i}S(\omega) = \frac{-iS(\omega)}{\coth(\beta\hbar\omega/2)+1}. \tag{8.21}$$

Combining Eqs. (8.18) and (8.21) yields the fluctuation-dissipation theorem,

$$S(\omega) = \hbar \mathrm{Im}[\chi(\omega)][\coth(\beta\hbar\omega/2)+1]. \tag{8.22}$$

The proof is slightly longer for the case $i \neq j$, but a similar theorem holds also in that case.

In an electrical system, the admittance $Y(\omega) = Z^{-1}(\omega)$ is defined through the response of the time derivative of A_i ($\sim$ charge) to the applied "force" (potential). In this case, we have $Y(\omega) = -i\omega\chi(\omega)$ and the fluctuation-dissipation theorem for current noise is

$$S_I(\omega) = 2\hbar\omega \mathrm{Re}[Y(\omega)][\coth(\beta\hbar\omega/2)+1]. \tag{8.23}$$

This consists of two parts, symmetric and antisymmetric in the frequency. As $\mathrm{Re}[Y(\omega)] = \mathrm{Re}[Y(-\omega)]$, the symmetric part is simply the one with the $\coth(\cdot)$-term, and the antisymmetric part is the other term. The symmetrized correlator of Eq. (8.1) is the one usually accessed in experiments — this captures only the $\coth(\cdot)$-part of the above expression. However, one has recently noticed that the negative-frequency noise corresponds to the "emitted" and the positive-frequency noise to the "absorbed" fluctuations. This becomes clearer when we treat the effect of noise on small quantum systems.

8.1.3 Thermal and quantum fluctuations

Consider now the limits of Eq. (8.23). In the limit $\hbar|\omega| \ll k_B T$ (for $T = 1K$, this would correspond to the angular frequency $|\omega| \ll 120$ GHz, or the frequency $f \ll 21$ GHz). In this case, we may take the limit $\omega \to 0$ and get

$$S(|\omega| \ll k_B T/\hbar) = 4\mathrm{Re}[Y(\omega)]k_B T. \tag{8.24}$$

Thus, in equilibrium systems at low frequencies, the current through a resistance R fluctuates by the amount $\sim 4k_B T/R$. This is called the thermal noise.[4]

In the opposite limit $\hbar\omega \gg k_B T$, we get a temperature independent noise,

$$S(\omega \gg k_B T/\hbar) = 2\mathrm{Re}[Y(\omega)]\hbar\omega. \tag{8.25}$$

These fluctuations are present even at $T = 0$, and they are called "vacuum" or "quantum" fluctuations. They may become the major noise "source" in the present-day high-frequency, low-temperature experiments.

Finally, in the limit $-\omega \gg k_B T/\hbar$, the nonsymmetrized noise in Eq. (8.23) vanishes. This shows that the quantum fluctuations can only absorb energy, and hence they cannot be used for example as a power source.

8.1.4 Shot noise

In driven (nonequilibrium) systems, one gets an additional source of noise, shot noise. This was first discovered by Schottky (1918), who studied the noise in vacuum tubes. The shot noise can be schematically described as follows.

Consider a conductor which transfers electrons in an uncorrelated fashion, such that the current may be described through a sequence of δ-peaks:

$$I(t) = e\sum_i \delta(t - t_i). \tag{8.26}$$

Here t_i are assumed random and uncorrelated.

The average current in a stationary system can then be expressed by

$$\langle I\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T I(t)dt = \lim\frac{e}{T}\int_0^T\sum_i \delta(t - t_i)dt = \lim\frac{eN_T}{T} \equiv \frac{e}{\tau}, \tag{8.27}$$

where N_T is the number of electrons transmitted in time T, and $\tau = T/N_T$ is the average time between the transfer of the electrons..

[4] Often one also uses the terms Johnson-Nyquist noise for the thermal current fluctuations, as the two first described such noise processes.

The current noise is then

$$\begin{aligned}\frac{1}{2}S_I(t,t') &= \langle I(t)I(t')\rangle - \langle I(t)\rangle\langle I(t')\rangle \\ &= e^2\langle\sum_{ij}\delta(t-t_i)\delta(t'-t_j)\rangle - \frac{e^2}{\tau^2} \\ &= e^2\langle\sum_{i}\delta(t-t_i)\delta(t'-t_i)\rangle + e^2\langle\sum_{i\neq j}\delta(t-t_i)\delta(t'-t_j)\rangle - \frac{e^2}{\tau^2} \\ &= e^2\langle\sum_{i}\delta(t-t_i)\delta(t'-t_i)\rangle + e^2\langle\sum_{i}\delta(t-t_i)\rangle\langle\sum_{j}\delta(t'-t_j)\rangle - \frac{e^2}{\tau^2} \\ &= e^2\langle\sum_{i}\delta(t-t_i)\delta(t'-t_i)\rangle = \frac{e^2}{\tau}\delta(t-t') = e\langle I\rangle\delta(t-t').\end{aligned} \tag{8.28}$$

In the second last equality, we assumed that t_i and t_j are uncorrelated.

Thus, the noise equals $S_I(t,t') = 2e\langle I\rangle\delta(t-t')$ and its Fourier transform is

$$S_I(\omega) = 2e\langle I\rangle. \tag{8.29}$$

Thus, the current noise is proportional to the average current. The origin of the noise is in the discreteness of electron charge, and it is therefore called the "shot noise".

8.2 Scattering approach to noise

Now let us make a more serious evaluation of the shot noise using the scattering approach. In Ch. 2, we derived an expression for the current operator in lead α in terms of the matrix $A^{\beta\gamma}_{mn}$, see Eq. (2.39). Now we may use this to evaluate the current correlator in the general nonequilibrium (but stationary) state of the system. For simplicity, let us concentrate on the symmetrized correlator defined in Eq. (8.1), but let us generalize this to describing also cross-correlations, i.e.,

$$S_{\alpha\beta}(t) \equiv \langle\delta\hat{I}_\alpha(t)\delta\hat{I}_\beta(t') + \delta\hat{I}_\beta(t')\delta\hat{I}_\alpha(t)\rangle. \tag{8.30}$$

Its Fourier transform satisfies in a stationary system (Exercise 8.2)

$$2\pi\delta(\omega+\omega')S_{\alpha\beta}(\omega) = \langle\delta\hat{I}_\alpha(\omega)\delta\hat{I}_\beta(\omega') + \delta\hat{I}_\beta(\omega')\delta\hat{I}_\alpha(\omega)\rangle. \tag{8.31}$$

To find an expression for this correlator, we need to calculate the statistical expectation value of products of four operators $\hat{a}$. For a noninteracting Fermi (Bose) system at equilibrium this is

$$\begin{aligned}&\langle a^\dagger_{\alpha k}(E_1)\hat{a}_{\beta l}(E_2)\hat{a}^\dagger_{\gamma m}(E_3)\hat{a}_{\delta n}(E_4)\rangle - \langle\hat{a}^\dagger_{\alpha k}(E_1)\hat{a}_{\beta l}(E_2)\rangle\langle\hat{a}^\dagger_{\gamma m}(E_3)\hat{a}_{\delta n}(E_4)\rangle \\ &= \delta_{\alpha\delta}\delta_{\beta\gamma}\delta_{kn}\delta_{ml}\delta(E_1-E_4)\delta(E_2-E_3)f_\alpha(E_1)[1\mp f_\beta(E_2)].\end{aligned} \tag{8.32}$$

The proof of this expressions requires the use of the Wick theorem, valid for a macroscopic equilibrium quantum system (see Exercise 84). The upper sign in the above expression corresponds to Fermi and the lower to the Bose statistics, and f_α is the corresponding equilibrium distribution function in lead α.[5] Now using this equation and the definition, Eq. (2.39), for the current operator, we obtain

$$\begin{aligned} S_{\alpha\beta}(\omega) = \frac{e^2}{h}\sum_{\gamma\delta}\sum_{mn}\int dE A^{\gamma\delta}_{mn}(\alpha;E,E+\hbar\omega)A^{\delta\gamma}_{nm}(\beta;E+\hbar\omega,E) \\ \times\{f_\gamma(E)[1\mp f_\delta(E+\hbar\omega)]+[1\mp f_\gamma(E)]f_\delta(E+\hbar\omega)\}. \end{aligned} \tag{8.33}$$

Below, we only concentrate on the zero-frequency noise for fermions. This has a slightly simpler expression,

$$\begin{aligned} S_{\alpha\beta}(0) =& \frac{e^2}{h}\sum_{\gamma\delta}\sum_{mn}\int dE A^{\gamma\delta}_{mn}(\alpha;E,E)A^{\delta\gamma}_{nm}(\beta;E,E) \\ &\times\{f_\gamma(E)[1-f_\delta(E)]+[1-f_\gamma(E)]f_\delta(E)\}. \end{aligned} \tag{8.34}$$

8.2.1 Two-terminal noise

In the two-terminal case, current conservation implies $S_{LL} = S_{RR} = -S_{LR} = -S_{RL}$. The expression (8.34) can be written down for example for the autocorrelation function where $\alpha = \beta = L$,

$$\begin{aligned} S_{LL} =& \frac{2e^2}{h}\int dE\Big\{ \mathrm{Tr}[A^{LL}(L)A^{LL}(L)]f_L(E)(1-f_L(E)) \\ &+ \mathrm{Tr}[A^{RR}(L)A^{RR}(L)]f_R(E)(1-f_R(E)) \\ &+ \mathrm{Tr}[A^{LR}(L)A^{RL}(L)][f_L(E)(1-f_R(E))+f_R(E)(1-f_L(E))]\Big\}. \end{aligned} \tag{8.35}$$

Utilizing the representation Eq. (2.29), and using the definition (2.37), we will have traces of the form $\mathrm{Tr}[r^\dagger r t^\dagger t]$ in the expression for the noise. These can be written in a form of a sum of transmission probabilities, provided we define an eigenbasis of the different parts of the scattering matrix. The corresponding eigenvalues of $t^\dagger t$ are then the transmission probabilities T_n and the eigenvalues of $r^\dagger r$ are the reflection probabilities R_n, defined such that $R_n = 1 - T_n$. In this basis, the above trace goes to a form

$$\mathrm{Tr}[r^\dagger r t^\dagger t] = \sum_n T_n(1-T_n). \tag{8.36}$$

[5] One can use an analogous theory to derive an expression for photon shot noise, for example.

Now after some algebra (see exercises), we are ready to write the two-terminal noise power at zero frequency in the form

$$S = \frac{4e^2}{h} \sum_n \int dE\{T_n(E)[f_L(1-f_L)+f_R(1-f_R)]+T_n(E)[1-T_n(E)](f_L-f_R)^2\}. \tag{8.37}$$

If the transmission eigenvalues are energy independent, we may integrate this to find

$$S = \frac{4e^2}{h}\left[2k_BT\sum_n T_n^2 + eV\coth\left(\frac{eV}{2k_BT}\right)\sum_n T_n(1-T_n)\right], \tag{8.38}$$

where V is the voltage applied between the two reservoirs.

In the limit $eV \to 0$, this produces the thermal noise, $S = 4k_BTG$, where $G = 2e^2/h\sum_n T_n$. In the opposite, "shot noise" limit, we get

$$S = 2eFG|V| = 2eF|\langle I\rangle|, \tag{8.39}$$

where

$$F = \frac{\sum_n T_n(1-T_n)}{\sum_n T_n} \tag{8.40}$$

is called the "Fano factor". In the tunnelling limit, $T_n \ll 1 \quad \forall n$, we obtain

$$S_I^{\text{tunn}} = 2e|V|G = 2e|\langle I\rangle|, \tag{8.41}$$

which is the Schottky result obtained above. One can also see that for such a noninteracting normal-metal system, the Fano factor always lies between zero and one. For example, a ballistic conductor has $T_n = 1$ for all transmitting channels. In this case $F = 0$, i.e., the shot noise vanishes. Hence, through the shot noise, one may obtain such information about the studied system that may not be present in the average current.

The calculation of the Fano factor for a given system requires in general the knowledge of, if not all the values T_n (in fact, these have really been measured individually in mechanical break junctions), at least their probability distribution $\rho(T)$. As mentioned in Ch. 2, this can be found for some systems by applying the random matrix theory. Fano factors have been calculated in this way for

- Diffusive wires for which $F = \frac{1}{3}$ and for
- Chaotic cavities for which $F = \frac{1}{4}$. Such a chaotic cavity is a ballistic system with an irregular form and only surface scattering, connected to the leads via small point contacts. If the form of the cavity is irregular enough, the electrons dynamics inside it is chaotic, and can at long times be described by the random matrix theory.

In superconducting systems, Andreev reflection modifies the formula for the shot noise, such that at energies below the superconducting energy gap Δ, the zero-temperature shot noise is given by

$$S_{N-S} = 4e|V|\frac{e^2}{h}\sum_n R_{A_n}(1 - R_{A,n}), \tag{8.42}$$

where $R_{A,n}$ is the probability for Andreev reflection. In this way, one obtains for an NIS tunnel contact $F = 2$, and for a diffusive normal-metal wire between normal and superconducting reservoirs in the absence of the proximity effect, $F = 2/3$.

For some systems, such as the tunnel contact to a superconductor, to a fractional quantum Hall sample, or a Luttinger liquid (a strongly interacting one-dimensional electron system), the Fano factor characterizes the effective charge responsible for the charge transport. This is illustrated by the above "charge" $2e$ of Cooper pairs in NIS contacts.

8.3 Boltzmann-Langevin approach

Besides the scattering approach, there are two other formulations from which one can calculate the nonequilibrium current noise in mesoscopic circuits. One relies on the nonequilibrium Keldysh Green's-function theory, introducing the fluctuation effects through a "counting field" defined in the terminals. This formulation is outside the scope of this course, but in the incoherent limit in a diffusive wire, this theory, applied for the second-order correlator, produces the same noise formula as an incoherent semiclassical Boltzmann-Langevin approach.[6] The latter is an extension of the Boltzmann theory introduced in Ch. 4. This assumes that due to the stochastic nature of the scattering inside the sample, the scattering currents, Eq. (4.10), fluctuate in time, i.e., they are of the form

$$J_{\vec{p},\vec{p}'} = \bar{J}_{\vec{p},\vec{p}'} + \delta J_{\vec{p},\vec{p}'}. \tag{8.43}$$

Here $\bar{J}_{\vec{p},\vec{p}'}$ is the average scattering current, and $\delta J_{\vec{p},\vec{p}'}$ is the fluctuation. As a result, also the energy distribution function fluctuates, and we may denote

$$f(\vec{r},\hat{p},t) = \bar{f}(\vec{r},\hat{p},t) + \delta f(\vec{r},\hat{p},t). \tag{8.44}$$

In the stationary case we can neglect the time dependence of the average distribution, but not of its fluctuation part. Now Eq. 4.6 should be replaced by[7]

$$(\partial_t + \vec{v}\cdot\nabla)f(\vec{r},\hat{p},E,t) = I_{\rm coll}[f] + \xi(\vec{r},\vec{p},t), \tag{8.45}$$

[6] This theory was first derived by Kogan and Schul'man, see Sov. Phys. JETP **29**, 467 (1969). This derivation follows that in the review by M. Büttiker and Ya. Blanter.

[7] Here we neglect the fluctuation of the electric field, which would add a term in the left hand side. This is valid if we can ignore the Coulomb-blockade -type effects, and concentrate on the limit of low frequencies.

where the Langevin forces ξ are

$$\xi(\vec{r},\vec{p},t) = \Omega \int \frac{d^3\vec{p}'}{(2\pi\hbar)^3} \left[\delta J_{\vec{p}',\vec{p}} - \delta J_{\vec{p},\vec{p}'}\right]. \tag{8.46}$$

Here Ω is the volume of the system, which cancels out at the end of the calculation. The latter term comes directly from the fluctuating part of $I_{\rm coll}[f]$. Now we should make an assumption about the correlator of the fluctuations. We assume that the currents $J_{\vec{p},\vec{p}'}$ are independent Poisson processes. This means that they are correlated only when the initial and final states and times are the same, and if they are evaluated at the same point, i.e.,

$$\begin{aligned}&\langle \delta J_{\vec{p}_1,\vec{p}_1'}(\vec{r}_1,t_1)\delta J_{\vec{p}_2,\vec{p}_2'}(\vec{r}_2,t_2)\rangle \\ &= \frac{(2\pi\hbar)^6}{\Omega}\delta(\vec{p}_1-\vec{p}_2)\delta(\vec{p}_1'-\vec{p}_2')\delta(\vec{r}_1-\vec{r}_2)\delta(t_1-t_2)\bar{J}_{\vec{p},\vec{p}'}(\vec{r}_1,t_1).\end{aligned} \tag{8.47}$$

As in a Poisson process, the second-order correlator ($\sim$ variance) is thus assumed to be directly proportional to the average. From this we get a relation for the correlator of the Langevin forces,

$$\langle \xi(\vec{r},\vec{p},t)\xi(\vec{r}',\vec{p}',t')\rangle = \delta(\vec{r}-\vec{r}')\delta(t-t')G(\vec{p},\vec{p}',\vec{r},t). \tag{8.48}$$

Here

$$G(\vec{p},\vec{p}') = \Omega\left\{\delta(\vec{p}-\vec{p}')\int d\vec{p}''\left[\bar{J}_{\vec{p}'',\vec{p}}+\bar{J}_{\vec{p},\vec{p}''}\right] - \left[\hat{J}_{\vec{p},\vec{p}'}+\hat{J}_{\vec{p}',\vec{p}}\right]\right\}, \tag{8.49}$$

where all functions are evaluated at the same position $\vec{r}$ and time t.

As in Ch. 4, we may again separate the momentum integrals to integrals over the direction $\hat{p}$ and the magnitude $|p|$, and relate the latter to the energy in the case of a spherical Fermi surface. Approximating the density of states by that evaluated at the Fermi surface, N_0, we then get

$$\langle \xi(\vec{r},\hat{p},E,t)\xi(\vec{r},\hat{p}',E',t')\rangle = \frac{1}{N_0}\delta(\vec{r}-\vec{r}')\delta(t-t')\delta(E-E')G(\hat{p},\hat{p}',\vec{r},E). \tag{8.50}$$

Using Eq. (4.10), we can relate this back to the average distribution function by noting that for purely elastic scattering,

$$\begin{aligned}G(\hat{p},\hat{p}') = \int d\vec{p}''[\delta(\hat{p}-\hat{p}')-\delta(\hat{p}'-\hat{p}'')][W(\hat{p},\hat{p}')\bar{f}(\hat{p})(1-\bar{f}(\hat{p}'')) \\ +W(\hat{p}'',\hat{p})\bar{f}(\hat{p}'')(1-\bar{f}(\hat{p}))].\end{aligned} \tag{8.51}$$

Again all functions are evaluated at the same position, same time and for the same energy E. Here we assume that inelastic scattering is much weaker than elastic, and only leads to the relaxation of the energy distribution.

Now let us consider what happens in the diffusive limit, where f can be expanded in spherical harmonics in the $\hat{p}$-dependence, see Eq. (4.11). Now the two included harmonics contain both the average and the fluctuating parts.

Proceeding as in pages 51-52, we note that the angular average of the Langevin source terms vanish, reflecting the fact that the number of electrons is conserved even in the presence of fluctuations. Therefore, Eq. (4.16) is unaltered by the Langevin term. However, we get an extra term for the p-wave term compared to Eq. (4.17). In the static case this is

$$\delta\vec{f} = -v\tau\nabla f_0 + 3\tau\int \hat{p}\xi d\hat{p}. \tag{8.52}$$

Therefore, the fluctuating part of the distribution function satisfies (in the zero-frequency limit where we can neglect the time derivative) an equation analogous to Eq. (4.18),

$$D\nabla^2 f_0 = l\nabla\cdot\int \hat{p}\xi d\hat{p} + I_{\rm inel}. \tag{8.53}$$

Assuming the inelastic scattering is much weaker than elastic, we may ignore the collision integral for the fluctuating part, and get

$$D\nabla^2\delta f_0 = l\nabla\cdot\int \hat{p}\xi d\hat{p}. \tag{8.54}$$

We can calculate the current density by integrating $\delta\vec{f}$ over the energy as in Eqs. (4.7), (4.19), and using the fluctuating distribution function instead of the average one. We get for the fluctuating part

$$\begin{aligned}\delta j =& -eN_0 v_F\int dE\delta(\delta\vec{f})\int d\hat{p}\hat{p}^2 = eN_0\int dE(-D\nabla\delta f_0 + \ell_{\rm el}\int d\hat{p}\hat{p}\xi)\\ \equiv&\sigma\delta\mu/e + \delta j^s.\end{aligned} \tag{8.55}$$

Here $\sigma = e^2N_0D$ is the Drude-Einstein conductivity. Thus, the total local current fluctuations consist of local potential fluctuations $\sigma\delta\mu/e$, and the fluctuations δj^s coming from the elastic scattering. The latter has a correlator

$$\langle\delta j^s(\vec{r},t)\delta j^s(\vec{r}',t')\rangle = 2\sigma\delta(\vec{r}-\vec{r}')\delta(t-t')\Lambda(\vec{r}) \tag{8.56}$$

with

$$\Lambda(\vec{r}) = \int dE\bar{f}_0(\vec{r},E)[1-\bar{f}_0(\vec{r},E)], \tag{8.57}$$

specified via the average distribution function $\bar{f}$.

In a quasi-one-dimensional geometry (where the system only changes in one direction, say, x), we can solve Eq. (8.54) and get

$$\delta f_0 = -\frac{c_1}{D}x + \frac{l}{D}\int_0^x dx\int d\hat{p}\hat{p}\xi + c_2, \tag{8.58}$$

where c_1 and c_2 are integration constants. If we assume that the sample is purely voltage-biased, the fluctuations of the distribution function vanish at the contacts, say at $x = 0$ and $x = L$. This implies $c_2 = 0$ and

$$c_1 = \frac{l}{L}\int_0^L dx\int d\hat{p}\hat{p}\xi.$$

The total local current fluctuation can be obtained from c_1 via

$$\delta j = -eN_0 \int dE c_1 = -\frac{1}{L}\int_0^L \delta j^s dx.$$

Now the correlator is obtained from

$$\begin{aligned} S(t,t') \equiv & 2\langle \delta I(t)\delta I(t')\rangle = 2\langle \int d\vec{y}\delta j(t,\vec{y}) \int d\vec{y}'\delta j(t',\vec{y}')\rangle \\ = & \frac{2}{L^2}\int_0^L dx \int_0^L dx' \int d\vec{y}d\vec{y}'\langle \delta j^s(x,\vec{y},t)\delta j^s(x',\vec{y}',t')\rangle \\ = & \frac{4A\sigma}{L^2}\int_0^L dx\Lambda(x)\delta(t-t'). \end{aligned} \tag{8.59}$$

Here we denoted the transverse coordinates by $\vec{y}$ and the cross section of the sample by A. We thus find white noise with the power spectral density

$$S(\omega \ll \omega^*) = \frac{4G_N}{L}\int_0^L dx \int dE \bar{f}_0(E,x)(1-\bar{f}_0(E,x)). \tag{8.60}$$

Here $G_N = A\sigma/L$ is the conductance of the wire. The frequency scale ω^* is related to the (inverse) time scales for particle diffusion through the sample or for charge relaxation, whichever is smaller: this analysis applies if the considered frequencies are much below these.

Equation (8.60) is applied in Exercise 8.5 to show, for example, that the Fano factor of a diffusive wire is 1/3, analogous to the result obtained from the scattering theory. This result illustrates that shot noise is not a coherent effect, and therefore for example phase-breaking scattering has no effect on it. Besides obtaining this Fano factor, the Boltzmann-Langevin theory is especially useful in studying the effect of inelastic scattering on shot noise. One of the important results is that shot noise is a mesoscopic effect: it vanishes if the sample size is much longer than the electron-phonon scattering length.

8.3.1 Langevin approach to noise in electric circuits

Typically noise measurements are done such that the noise from the studied object is coupled to the measurement device, where it is often amplified and transformed to an average current via some nonlinear element. Then the interesting topic is to find out how such a noise coupling can be quantified. It turns out that this can be done in a similar spirit as the Boltzmann-Langevin scheme following fairly simple circuit rules: in parallel with all the dissipative elements of the circuit, one writes a noise current source, as depicted in Fig. 8.2. When treating only fluctuations, the average voltage need not be written explicitely. The correlator of this noise current corresponds then to shot noise, thermal noise, etc., depending on the parameter regime one is interested. An example of such a noise description is given in Exercise 8.4.

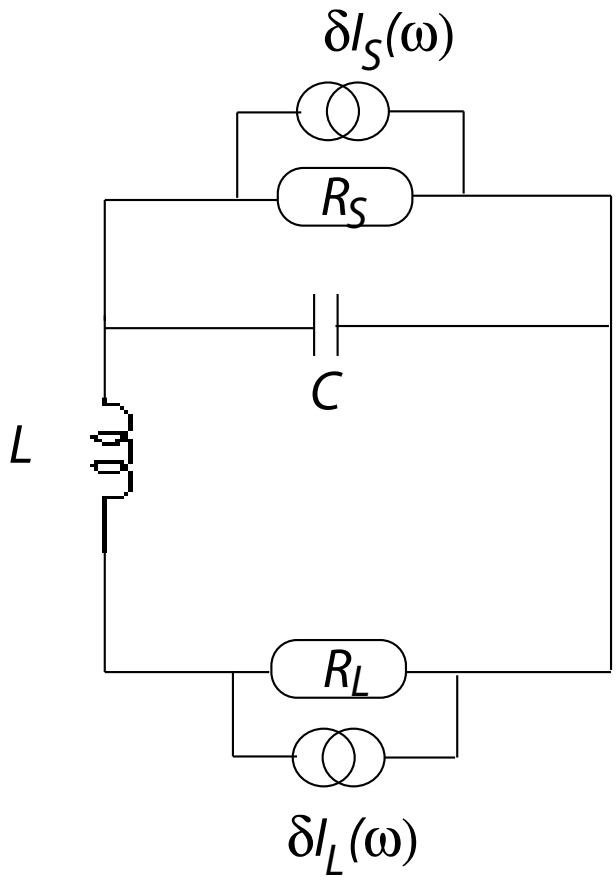


Figure 8.2: Example of a transformer circuit with the noise current sources.

8.4 Cross correlations

So far, we only concentrated on two-probe systems, where it is enough to concentrate on the autocorrelation function of the current in one of the electrodes. However, cross-correlations in mesoscopic systems have recently attracted a lot of attention, as there are a few simple principles describing especially the sign of the cross-correlations.[8]

The autocorrelators discussed above and the cross correlators are not completely independent from each other. In the limit of low frequencies, not only the average current, but also its fluctuations δI have to be conserved, i.e.,

[8] For a detailed treatment of cross-correlations, see M. Büttiker, Phys. Rev. B **46**, 12 485 (1992).

$\sum_\alpha \delta I_\alpha = 0$. Therefore, the fluctuations have to obey a "sum rule",[9]

$$\left\langle (\sum_\alpha \delta I_\alpha)^2 \right\rangle = \sum_\alpha \langle (\delta I_\alpha)^2 \rangle + \sum_{\alpha \neq \beta} \langle \delta I_\alpha \delta I_\beta \rangle = 0. \tag{8.61}$$

For example, for a two-probe system, this implies $S_{LL} = S_{RR} = -S_{LR} = -S_{RL}$, and one does not obtain more information from the cross-correlators than just from the autocorrelation function studied above. Thus cross-correlators are generally studied in multiprobe geometries. But there is something one can learn from this relation: As the autocorrelation functions are always positive, it seems that the cross-correlators should be negative. Thus, a current in one lead generally implies that there should be at the same time an opposite current in another lead. It turns out that such a property holds for a general multi-probe noninteracting fermion system, but positive cross-correlations can be found in systems containing Andreev reflection.

8.4.1 Equilibrium correlations

Consider first the case when there is no voltage applied over the sample. In this case, assuming that the scattering matrix is energy independent within a few $k_B T$ around the Fermi energy, we get

$$S^{\rm eq}_{\alpha\beta} = \frac{2e^2}{h} k_B T \sum_{\gamma\delta} {\rm Tr}[A^{\gamma\delta}(\alpha) A^{\delta\gamma}(\beta)], \tag{8.62}$$

independent of whether we consider fermions or bosons. This trace can be shown to yield

$$S^{\rm eq}_{\alpha\beta} = \frac{2e^2}{h} k_B T \sum_{\gamma\delta} {\rm Tr}[2 \cdot I_\alpha \delta_{\beta\alpha} - (s^{\alpha\beta})^\dagger s^{\alpha\beta} - (s^{\beta\alpha})^\dagger s_{\beta\alpha}]. \tag{8.63}$$

Here I_α is the unit (identity) matrix of size $M_\alpha \times M_\alpha$. For the autocorrelation function, i.e., $\alpha = \beta$, this yields simply the thermal noise $4k_B TG$. For the cross-correlators, we can define (positive definite) total transmission probabilities between α and β by $T_{\alpha\beta} \equiv {\rm Tr}[(s^{\alpha\beta})^\dagger s^{\alpha\beta}]$ and get

$$S^{\rm eq}_{\alpha\neq\beta} = -\frac{2e^2}{h} k_B T [T_{\alpha\beta} + T_{\beta\alpha}]. \tag{8.64}$$

Thus, equilibrium cross correlations are always negative, both for bosons and fermions.

8.4.2 Finite-voltage cross correlations

Now consider the cross correlations in the presence of a finite voltage and assume again that the scattering matrix is energy independent within the interesting

[9] This is proven in the paper of the above footnote.

region of energies. Now the sum over the lead indices γ and δ in Eq. (8.33) can be separated in two parts: one where the terms contain $f_\gamma(1 \mp f_\gamma)$, and another with cross-terms of the form $f_\gamma(1 \mp f_\delta)$, $\delta \neq \gamma$. The previous terms will yield a contribution proportional to the equilibrium fluctuations — these terms would vanish at $T = 0$.[10] For the cross-terms terms, $A_{\gamma\delta}(\alpha) = -(s^{\alpha\gamma})^\dagger s^{\gamma\delta}$. These yield

$$S^{\rm cross}_{\alpha\neq\beta} = \frac{2e^2}{h} \sum_{\gamma\neq\delta} \int dE {\rm Tr}[(s^{\alpha\gamma})^\dagger s^{\alpha\delta} (s^{\beta\delta})^\dagger s^{\beta\gamma}] f_\gamma(E)[1 \mp f_\delta(E)]. \tag{8.65}$$

It can be shown[11] that this equals

$$S^{\rm cross}_{\alpha\neq\beta} = \mp \frac{2e^2}{h} \int dE {\rm Tr} \left[\left(\sum_\gamma s^{\beta\gamma} (s^{\alpha\gamma})^\dagger f_\gamma(E) \right) \left(\sum_\delta s^{\alpha\delta} (s^{\beta\delta})^\dagger f_\delta(E) \right) \right]. \tag{8.66}$$

The upper sign is for fermions and the lower for bosons. But now we have a trace of a product of a matrix with its adjoint. Such a trace is always positive. Hence, for fermions, such "transport" cross-correlations are always negative, whereas for bosons they are positive. As the total correlations are a sum of these transport correlations and terms proportional to the equilibrium correlations, the sign of the cross-correlations for (noninteracting) fermions is always negative, but for bosons it can change the sign.[12]

The sign of the cross correlations is closely related to the statistics of the current carriers. The negative correlations for fermion systems is a sign of "antibunching" due to the Pauli principle that prohibits two electrons occupying the same state. On the contrary, bosons can "bunch", i.e., occupy the same quantum state, and therefore boson systems can show positive cross-correlations.

It has been suggested that also electron systems could show "bunching", i.e., positive cross-correlations, if they are measured with certain interacting systems. For example, if one of the leads in a three-terminal system is superconducting, one can in principle measure positive cross-correlations between the two other leads. The reason for such positive cross correlations can be schematically thought as a Cooper pair breaking into two excitations transmitting into different leads.

8.5 Full counting statistics

So far in this chapter, we have characterized the current fluctuations effectively through their variance, i.e., the second moment. We found that studying the nonequilibrium shot noise reveals information on the studied system not present

[10] For bosons in this limit, one should also take into account the effects due to Bose condensation. However, bosons are presented here just to illustrate a point about the correlations.

[11] See again the Phys. Rev. B by Marcus Büttiker.

[12] Negative cross-correlations for a noninteracting electron system have been measured for example by M. Henny, *et al.*, and W. D. Oliver, *et al.*, Science **284**, 296 and 299 (1998).

in the average current (essentially through the Fano factor). Further information can be obtained by studying the higher moments (cumulants) of charge transfer, or in general its full statistics.

8.5.1 Basic statistics

Let us briefly go through a few definitions concerning the statistical description of a stochastic phenomenon. Consider a process indexed by number n, that takes place with a certain probability $P(n)$. For example, this could be the probability that n charges have passed through a conductor in a given time. A *characteristic function* $\chi(\lambda)$ is defined by the (discrete or continuous) Fourier transform of this probability:

$$\chi(\lambda) \equiv \langle e^{i\lambda n} \rangle \equiv \sum_n e^{i\lambda n} P_n. \tag{8.67}$$

The mth derivative of the characteristic function is proportional to the m'th *raw moment* $\langle n^m \rangle$,

$$\left(\frac{d^m}{d\lambda^m} \chi(\lambda) \right)_{\lambda=0} = i^m \langle n^m \rangle. \tag{8.68}$$

If one first takes the logarithm and then differentiates $\chi(\lambda)$, one obtains the *cumulants* C_m,

$$C_m = \langle\langle n^m \rangle\rangle \equiv (-i)^m \left(\frac{d^m}{d\lambda^m} \ln \chi(\lambda) \right)_{\lambda=0}. \tag{8.69}$$

The lowest cumulants can be expressed through the *central moments* $M_m \equiv \langle (n - \langle n \rangle)^m \rangle$ as

$$\begin{aligned} C_1 &= \langle n \rangle = \text{average} \\ C_2 &= M_2 = \text{variance} \\ C_3 &= M_3 \\ C_4 &= M_4 - 3M_2^2 \\ C_5 &= M_5 - 10 M_2 M_3. \end{aligned} \tag{8.70}$$

The point in describing a probability distribution via cumulants is that they generally show how a given distribution deviates from Gaussian. Namely, for a Gaussian distribution, $C_m = 0$ for $m > 2$.

Besides Gaussian, two important distributions in terms of FCS are the Poisson and the binomial distributions, which have the characteristic functions $\chi_{\text{Poisson}}(\lambda) = e^{M_1(e^{i\lambda}-1)}$, and $\chi_{\text{Binomial}} = (1 - p + pe^{i\lambda})^n$, where M_1, p and n are values parameterizing the distributions.

8.5.2 FCS of charge transfer

The natural observable to study in mesoscopic systems is the statistics of the transferred charge in a measurement time t_0. In terms of the fluctuating current,

the cumulants of the number $n(t_0)$ of transmitted charges are expressed as[13]

$$\langle\langle n^m(t_0)\rangle\rangle = \frac{1}{e^m}\int_0^{t_0} dt_1 \dots dt_m \langle\langle I(t_1)I(t_2)\dots I(t_m)\rangle\rangle. \tag{8.71}$$

The first cumulant is $\langle\langle n(t_0)\rangle\rangle = \langle I\rangle t_0/e$, and the second in the limit of a long t_0 is proportional to the noise power,

$$\langle\langle n^2(t_0)\rangle\rangle = \frac{St}{2e^2}. \tag{8.72}$$

The characteristic function has been calculated for many types of mesoscopic systems, but here we only concentrate on the simplest: a normal-metal conductor between two normal-metal probes. In this case, the characteristic function obtained from a quantum-mechanical calculation is of the form $\chi(\lambda) = \exp(S(\lambda))$ with[14]

$$S(\lambda) = \frac{2t_0}{h}\sum_j \int dE \ln\left[1 + T_{jLR}(E)(e^{i\lambda}-1) + T_{jRL}(E)(e^{-i\lambda}-1)\right]. \tag{8.73}$$

Here $T_{jLR} = T_j f_L(E)(1-f_R(E))$ is the probability for tunnelling from the left to the right of the contact for the eigenchannel j, and T_{jRL} is the corresponding probability from the right to the left. The argument of the logarithm essentially summarizes the possible processes: no scattering, scattering from left to right, and scattering from right to left. In this expression, the *counting factors* $e^{\pm i\lambda}-1$ correspond to single-charge transfers between the contacts. For example, for Andreev reflection, which is a two-electron effect, the corresponding counting factor would be of the form $e^{\pm 2i\lambda}-1$.

For equilibrium (no voltage), the counting statistics is described by

$$S(\lambda) = -\frac{2t_0 k_B T}{h}\sum_j \arcsin^2\left(\sqrt{T_j}\sin\left(\frac{\chi}{2}\right)\right). \tag{8.74}$$

Thus, even at equilibrium the fluctuations are not Gaussian, except for a ballistic contact $T_j = 1 \forall j$, in which case $S(\lambda) = -t_0 k_B T \lambda^2/h$.

At a vanishing temperature, the characteristic function describes a series of binomial processes,

$$\chi(\lambda) = \prod_j \left(1 + T_j(e^{i\lambda}-1)\right)^N, \tag{8.75}$$

[13] In the quantum-mechanical case, the problem of the "correct" symmetrization of such cumulants grows with the index of the cumulant. The properties of different orderings of the current operators are not very well known, but it has recently been shown that different measurement schemes correspond to different types of orderings. The typically studied "quantum" definition of the characteristic function is $\chi(\lambda) = \langle \mathcal{T}_K e^{-\frac{i}{2e}\int_{C_K} dt\lambda(t)I(t)}\rangle$, where C_K is a "Keldysh contour" containing two branches, one going from $-\infty$ to the measurement time t_0 and one the opposite direction. The operator $\mathcal{T}_K$ corresponds to ordering of the operators within this contour, and $\lambda(t) = \lambda$ for the upper branch, and $\lambda(t) = -\lambda$ for the lower branch.

[14] For a summary of the research on FCS in mesoscopic systems, see for example W. Belzig, "Full Counting Statistics in Quantum Contacts", http://xxx.lanl.gov/abs/cond-mat/0312180.

where $N = 2et_0|V|/h$ is the *number of attempts* in a given time t_0. This has an easy interpretation of a particle in channel j trying to transmit through the scatterer for N times, the probability of success being T_j.

We can expand Eq. (8.75) in the cumulants of the charge transfer. The lowest cumulants are

$$\begin{aligned}
\langle\langle n\rangle\rangle &= N\sum_j T_j \\
\langle\langle n^2\rangle\rangle &= N\sum_j T_j(1-T_j) \\
\langle\langle n^3\rangle\rangle &= N\sum_j T_j(1-T_j)(1-2T_j) \\
\langle\langle n^4\rangle\rangle &= N\sum_j T_j(1-T_j)(1-6T_j(1-T_j)).
\end{aligned} \tag{8.76}$$

The first two are thus consistent with the formulae for the average current and noise derived earlier in the course.

Note that for a tunnel junction with $T_j \ll 1$, the counting distribution turns into a Poissonian, for which all the cumulants are of the form $\langle\langle n^m\rangle\rangle N\sum_j T_j$.

Whereas already quite much is known on the properties of the full counting statistics, or higher cumulants of different types of systems, the first higher (than second) cumulants have been measured only very recently.[15] Also, very little is known about the frequency dependence of the higher-order cumulants, or the proper ordering of the current operators when describing the statistics measurement. Full counting statistics is therefore both an experimentally and theoretically a rapidly evolving field.

8.6 Problems

1. Wiener-Khintchine theorem:
   Assume we measure random noise signal $\delta x(t)$ in a stationary system over a long time $t \to \infty$. As $\delta x(t)$ is a real quantity, its Fourier transform satisfies $\delta x(-\omega) = \delta x^*(\omega)$. Then one may write

   $$\delta x(t) = \int_0^\infty \frac{d\omega}{2\pi}[\delta x(\omega)e^{-i\omega t} + \delta x^*(\omega)e^{i\omega t}]. \tag{8.77}$$

   Now consider a spectrum analyzer which measures this signal. It contains a narrow-band band-pass filter which only leaves the (angular) frequencies $\omega \in [\omega_0 - \Delta\omega/2, \omega_0 + \Delta\omega/2]$, and an output detector that measures the mean square of this signal (corresponding to the signal power). Show that

[15] The third cumulant of a tunnel junction was measured by B. Reulet, *et al.*, Phys. Rev. Lett. **91**, 196601 (2003), and more recently by Yu. Bomze, *et al.*, Phys. Rev. Lett. **95**, 176601 (2005). Very recently a Swiss group claimed to have measured the full counting statistics of a quantum dot, see S. Gustavsson, *et al.*, http://xxx.lanl.gov/abs/cond-mat/0510269.

the averaged squared signal from the filtered band equals the spectral density of noise, $S(\omega_0)$ times the band width $\Delta f = \Delta\omega/2\pi$ (assuming that $S(\omega_0)$ does not essentially change within this band), i.e.,

$$\langle \delta x^f(t)\delta x^f(t)\rangle = S(\omega_0)\frac{\Delta\omega}{2\pi} = S(\omega_0)\Delta f. \tag{8.78}$$

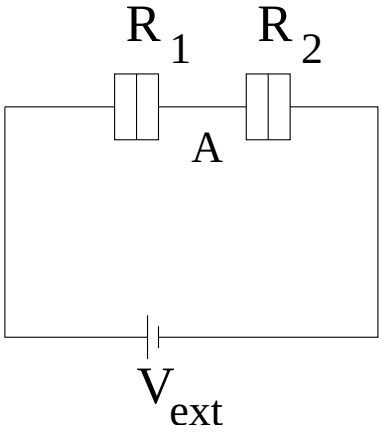
Here $\delta x^f(t)$ is the signal after the band-pass filtering,

$$\delta x^f(t) = \int_{\omega_0-\Delta\omega/2}^{\omega_0+\Delta\omega/2} \frac{d\omega}{2\pi}[\delta x(\omega)e^{-i\omega t} + \delta x^*(\omega)e^{i\omega t}]. \tag{8.79}$$

Hint: Write down the spectral density of noise for two times, t and t', and use the fact that in a stationary setup, it only depends on $t - t'$.

2. Prove Eq. (8.31).

3. Prove Eqs. (8.37) and (8.38).



4. Consider two tunnel barriers, of resistances R_1 and R_2 in series, connected to an ideal voltage source (see figure above) with voltage V_{ext}. Assume that the current noises through R_1 and R_2 are uncorrelated, such that the noise powers from these two sources can be simply added to obtain the total current noise in the system. Then, we can analyze the noise due to R_1 treating R_2 simply as a resistor and vice versa. The current fluctuations in the resistor R_1 produce voltage fluctuations $\Delta V(\omega)$ at the point A between the barriers. These result in additional fluctuations $\Delta V(\omega)/R_1$. Thus, the total current fluctuations due to the fluctuations in R_1 are

$$\Delta I(\omega) = \Delta V(\omega)/R_1 + \delta I(\omega), \tag{8.80}$$

where $\delta I(\omega)$ is the intrinsic (shot or thermal) noise over R_1. $\Delta V(\omega)$ may be easily calculated by requiring the fluctuations in the total voltage over the system to vanish, i.e.,

$$\Delta V(\omega) + R_2\Delta I(\omega) = 0. \tag{8.81}$$

a) By using the fact $S_{int}(\omega) = \langle \delta I(\omega)\delta I(-\omega)\rangle$, find the total current noise $S_I = \langle \Delta I(\omega)\Delta I(-\omega)$ (using an incoherent sum due to fluctuations in R_1 and those in R_2). b) Assuming $R_1 = R_2$, show that adding tunnel barriers but keeping the same total resistance decreases the shot noise (noise at $T = 0$, i.e., $S_{int} = 2e\langle I\rangle$), but the thermal noise $S^i_{int} = 4k_BT/R_i$ remains unaffected. c) Calculate the voltage noise $S_V = \langle \Delta V \Delta V\rangle$ at point A.

5. Using the Boltzmann-Langevin formalism and Eq. (8.60), consider the shot-noise limit where the temperature vanishes, $T = 0$, and there is a voltage V over the wire. Calculate S_I for a) nonequilibrium diffusive system where $L \ll l_{ee}, l_{eph}$ b) quasiequilibrium diffusive system with $l_{ee} \ll L \ll l_{eph}$ and c) equilibrium diffusive system with $l_{ee}, l_{eph} \ll L$. In each case, compare this to the average current,

$$I = \frac{A\sigma_N}{L}V \tag{8.82}$$

and find the corresponding Fano factor. Use the distribution functions calculated in the exercises of Ch. 4. Hint: Assuming the integrals converge, you may interchange the order of integration in Eq. (8.60).

The following integrals may be useful (the shorthand notation $f_L = f^0(E; 0, T)$, $f_R = f^0(E, eV, T)$ is used):

$$\int_{-\infty}^{\infty} dE f^0(1 - f^0) = k_BT \tag{8.83}$$

$$\int_{-\infty}^{\infty} dE(f_R - f_L)(1 - 2f_L) = \frac{(eV - 2k_BT)e^{eV/k_BT} + eV + 2k_BT}{e^{eV/k_BT} - 1} \tag{8.84}$$

$$\int_{-\infty}^{\infty} dE(f_R - f_L)^2 = \int_{-\infty}^{\infty} dE(f_R - f_L)(1 - 2f_L) \tag{8.85}$$

$$\int_0^1 \sqrt{x(1-x)} = \frac{\pi}{8}. \tag{8.86}$$

6. Consider the setup shown in Fig. 8.2. Calculate the total finite-frequency current noise across the "load", resistance R_L, by assuming that the intrinsic noise sources are uncorrelated and the autocorrelation function is frequency independent, i.e., $\langle \delta I_{S/L}(\omega)\delta I_{S/L}(-\omega)\rangle = S_{L/R}$. At the resonance frequency $\omega_0 = 1/\sqrt{LC}$, find L and C that "match" the resistance R_S of the source with the resistance R_L of the load, i.e., maximize the coupling between the total current noise across the load and the intrinsic noise across the source.

7. An alternative derivation of the Boltzmann-Langevin equation, Eq. (8.60) for a diffusive wire. Model the diffusive wire by a series of N tunnel barriers ($T_n \ll 1$), each with resistance $R_N = 1/(NG_N)$, and with an intrinsic noise source δI_i ($i = 1, \ldots N$) whose correlator is of the form of Eq. (8.37),

with f_L (f_R) given by the distribution functions at the left (right) of the junction. Find a recursion equation for the voltage fluctuations following a similar scheme as in Exercise 8.4. In the limit $N \to \infty$, show that the zero-frequency noise in this system can be modelled by Eq. (8.60).

8. Find the four lowest cumulants of the equilibrium fluctuations described by Eq. (8.74).

Chapter 9

Dissipation in quantum mechanics

The concept of noise and fluctuations is especially useful when describing the properties of small, spatially restricted quantum systems coupling to the outside world. In the elementary quantum-mechanics courses, we have learned that the time dependence of the state (density matrix) of a given system described with the Hamiltonian H is given by

$$\rho(t) = U(t)\rho(0)U^\dagger(t), \tag{9.1}$$

where $U(t) = T\exp(-\frac{i}{\hbar}\int_0^t H dt)$ is a unitary operator, i.e., $UU^\dagger = U^\dagger U = \mathbf{I}$. Because of unitarity, from the "final" state $\rho(t)$ we can easily find the initial state $\rho(0)$:

$$\rho(0) = U^\dagger(t)\rho(t)U(t). \tag{9.2}$$

Thus, for quantum systems it does not greatly matter whether we describe things advancing forward or backward in time.

From any everyday phenomenon we know that for the classical world there is a big difference between the future and the past — for example, when one thinks about classically allowed physical processes. A window is easy to break with a snowball, but it is quite unlikely that a reverse process would take place, a crashed window turning back into a single piece and the snowball flying back into the thrower's hand. In the world of electronics, the essential concept differentiating the backward-in-time processes from the forward-moving processes is *dissipation*: when a current flows through a resistor, it heats up the lattice of the resistor, and this heat is carried away by the phonons, which then somewhere else heat up the rest of the universe. The reverse rarely takes place: phonons emerging from the measurement setup, and making rise to a finite electric current.

In restricted quantum systems, the total energy of the system is conserved, and this essentially conserves unitarity. The way out is to describe the connec-

tion of this small system to the outside world via some *fluctuating* field.[1] That is, the Hamiltonian is of the form

$$H = H_0 + g\hat{A}_i\hat{f}(t), \tag{9.3}$$

where $\hat{f}(t)$ is a zero-average fluctuating field and g is a (scalar) coupling coefficient to this field. Generally $\hat{f}$ can still be an operator, but it is typically chosen such that it commutes with all the degrees of freedom of the interesting subsystem, i.e., all the operators $\hat{A}_i$. At the end of this chapter, we will describe a typical model for $\hat{f}(t)$, but at first we discuss the consequences of such a fluctuating field. The main effects are that generally $\hat{f}$ can make the subsystem (described by H_0) *relax* towards an equilibrium state, and also lose its coherent features. The latter process is called *dephasing*.

9.1 Relaxation

Assume for simplicity that the hamiltonian H_0 is time independent. Then we can describe the quantum system with the eigenvectors $|n\rangle$ of this H_0 with the eigenenergies E_n, satisfying

$$H_0|n\rangle = E_n|n\rangle. \tag{9.4}$$

Now assume that at some initial state $t = 0$, the occupation number of these states is described by the probabilities p_n. These p_n are the diagonal elements of the system density matrix at time $t = 0$. As by definition $|n\rangle$ are the eigenenergies of the Hamiltonian, without the perturbation term in Eq. (9.3), p_n would remain constant at later times as well. The fluctuating field thus can cause transitions in the system, and thus making p_n time dependent. The most convenient way to describe the system is using the interaction representation: we make the transformation to the Heisenberg picture in the subsystem, i.e., $\hat{A}_i(t) = \exp(iH_0t/\hbar)\hat{A}_i(0)\exp(-iH_0t/\hbar)$, and include the interaction term $g\hat{f}\hat{A}_i(t)$ perturbatively. In the first order in g we then get the relation between the wave function at $t = 0$ and at some later time t:

$$|\psi_I(t)\rangle = |\psi_I(0)\rangle - \frac{i}{\hbar}\int_0^t dt' g\hat{f}(t')\hat{A}_i(t)|\psi_I(0)\rangle. \tag{9.5}$$

Assuming the system initially at state $|n\rangle$, the probability amplitude that it has made a transition to the state $|m\rangle$, $m \neq n$ is then

$$\alpha_m = \langle m|\psi_I(t)\rangle = -\frac{ig}{\hbar}\int_0^t dt'\hat{f}(t')\langle m|\hat{A}_i(t)|n\rangle = -\frac{ig}{\hbar}\int_0^t dt'\hat{f}(t')e^{-i\omega_{mn}t}A_i^{mn}, \tag{9.6}$$

[1] There are many different ways to make a quantitative analysis of dissipation in quantum-mechanical systems, although most of them rely on a "system+environment" model similar to the one introduced here. The most widely used reference in this context is the book by U. Weiss, "Quantum dissipative systems", World Scientific, Second Edition, (1999).

where $A_i^{mn} = \langle m|\hat{A}_i(0)|n\rangle$ and the dynamic phase with $\omega_{mn} = (E_m - E_n)/t$ comes from the time evolution of the Heisenberg operator (or the states in the Schrödinger picture). The probability for the occupation of the state m is then

$$p_m(t) = |\alpha_m|^2 = \frac{g^2}{\hbar^2}|A_i^{mn}|^2 \int_0^t \int_0^t dt_1 dt_2 e^{-i\omega_{mn}(t_1-t_2)} \hat{f}(t_1)\hat{f}(t_2). \quad (9.7)$$

Let us now ensemble average this probability over all the fluctuations $\hat{f}$:

$$\bar{p}_m(t) = \frac{g^2}{\hbar^2}|A_i^{mn}|^2 \int_0^t \int_0^t dt_1 dt_2 e^{-i\omega_{mn}(t_1-t_2)} \langle \hat{f}(t_1)\hat{f}(t_2)\rangle. \quad (9.8)$$

Now let us perform a change of variables in the integrals, $\tau = t_1 - t_2$, and $T = (t_1 + t_2)/2$. Then we get

$$\bar{p}_m(t) = \frac{g^2}{\hbar^2}|A_i^{mn}|^2 \int_0^t dT \int_{-B(T)}^{B(T)} d\tau e^{-i\omega_{mn}\tau} \langle \hat{f}(T + \tau/2)\hat{f}(T - \tau/2)\rangle, \quad (9.9)$$

where $B(T) = T$ if $T < t/2$ and $B(T) = t - T$ otherwise.

Now assume the system describing the fluctuating field $\hat{f}(t)$ is stationary and is described by a correlation time τ_f. Then for $t \gg \tau_f$, we can take $B(T) \to \infty$ and get $\bar{p}_m(t) = \Gamma_{mn} t$ with

$$\Gamma_{mn} = \frac{g^2}{\hbar^2}|A_i^{mn}|^2 \int_{-\infty}^{\infty} d\tau e^{-i\omega_{mn}\tau} \langle \hat{f}(\tau)\hat{f}(0)\rangle = \frac{g^2}{\hbar^2}|A_i^{mn}|^2 S_f(-\omega_{mn}). \quad (9.10)$$

Here $S_f(\omega)$ is the (nonsymmetrized) autocorrelator of $\hat{f}$ and Γ_{mn} is the *relaxation rate* of the state n into a state m. For a general initial density matrix p_n, we can again construct a Master equation to describe the occupations, similar to the treatment of the charge state in Ch. 5:

$$\frac{dp_n}{dt} = \sum_m (\Gamma_{nm} p_m - \Gamma_{mn} p_n). \quad (9.11)$$

This master equation can, for example, be used to show how the quantum system reaches an equilibrium state, and the temperature of this equilibrium state can be found to be related to the properties of $S_f(\omega)$.

In the long-time limit, assuming that the perturbation is sufficiently weak,[2] the noise properties of the "outside world", described $\hat{f}$, determine the occupation probabilities of the state n.

9.2 Dephasing

For the full description of the properties of a quantum system, it is not sufficient to concentrate on the occupation numbers of the different energy levels — rather,

[2] It can be shown that the next order in the perturbation parameter g is related to the third cumulant of the fluctuations in f, see T. Ojanen and TTH, Phys. Rev. B **73**, 20501(R) (2006) and cond-mat/0609133.

all the quantum interference effects are described by the interference of different states. These are described by the off-diagonal elements of ρ. The fluctuations $\hat{f}$ will lead to a decay of these off-diagonal elements in time. For simplicity, we describe this process using a spin, i.e., a two-level system, as the small quantum system. The Hamiltonian for such a system can be specified by the Pauli spin matrices σ_i:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.12}$$

The simplest way to treat such a system is to describe it in the energy eigenbasis, such that the Hamiltonian is

$$H_0 = \frac{\epsilon}{2}\sigma_z. \tag{9.13}$$

The most general coupling of this system to the external world is described by a term of the form

$$H_{\rm int} = g_z \hat{f}_z \sigma_z + g_x \hat{f}_x \sigma_x. \tag{9.14}$$

A term proportional to σ_y could also be added, but it would not qualitatively change the outcome.

The term $\hat{f}_z\sigma_z$ commutes with the Hamiltonian, and does not lead to any relaxation (its matrix elements for $n \neq m$ would vanish). However, its presence does lead to dephasing. Let us simplify our problem a little more by assuming also $g_x = 0$. The time evolution of the density matrix can then be written as

$$\begin{aligned} \rho(t) = &\exp(-iH_0 t/\hbar)\mathcal{T}\exp\left[-\frac{i}{\hbar}g_z\int_0^t dt'\,\hat{f}_z(t')\sigma_z\right]\rho(0) \\ &\times\tilde{\mathcal{T}}\exp\left[\frac{i}{\hbar}g_z\int_0^t dt'\,\hat{f}_z(t')\sigma_z\right]\exp(iH_0 t/\hbar). \end{aligned} \tag{9.15}$$

Here $\mathcal{T}$ is the time ordering operator, and $\tilde{\mathcal{T}}$ is the anti-time ordering operator. Assume the initial state is of the form

$$\rho(0) = \begin{pmatrix} p & \alpha \\ \alpha & (1-p) \end{pmatrix} = \frac{1}{2}\sigma_0 + \left(p - \frac{1}{2}\right)\sigma_z + \alpha\sigma_x. \tag{9.16}$$

Here σ_0 is the 2×2 unit matrix. The term α thus describes the off-diagonal part containing the information on the interference of the two states. Plugging this into Eq. (9.15) and noting that σ_0 and σ_z commute with the perturbation and H_0, we get

$$\begin{aligned} \rho(t) = \frac{1}{2}\sigma_0 + (p-\frac{1}{2})\sigma_z + \alpha\exp(-iH_0 t/\hbar)\mathcal{T}\exp\left[-\frac{i}{\hbar}g_z\int_0^t dt'\,\hat{f}_z(t')\sigma_z\right]\sigma_x \\ \times\tilde{\mathcal{T}}\exp\left[\frac{i}{\hbar}g_z\int_0^t dt'\,\hat{f}_z(t')\sigma_z\right]\exp(iH_0 t/\hbar). \end{aligned} \tag{9.17}$$

Let us now concentrate on the σ_x-component of ρ, i.e., its off-diagonal part, and average it over the realizations of $\hat{f}_z$. We get

$$\rho_{od} = \frac{1}{2}\mathrm{Tr}[\sigma_x \rho(t)] = \frac{\alpha}{2}\mathrm{Tr}\left[\sigma_x \langle \mathcal{T} \exp(iA\sigma_z)\sigma_x \tilde{\mathcal{T}} \exp(-iA\sigma_z)\rangle\right], \tag{9.18}$$

where $A = -(\epsilon t/2 + g_z \int_0^t dt' \hat{f}_z(t'))/\hbar$. Using the fact that $\exp(a\sigma_z)\sigma_x = \sigma_x \exp(-a\sigma_z)$,[3] we finally get

$$\rho_{od} = \frac{\alpha}{2}\mathrm{Tr}[\langle \exp(2iA\sigma_z)\rangle] = \frac{\alpha}{2}\mathrm{Tr}\left[\exp(-i\epsilon t\sigma_z/\hbar)\langle \mathcal{T}_K \exp(-\frac{i}{\hbar}g_z \int_{C_K} dt' \hat{f}_z(t'))\rangle\right]. \tag{9.19}$$

Here the contour C_K is the so-called Keldysh contour: it first runs from the initial time $t = 0$ to the time t and then backwards in time from t to $t = 0$. The operator $\mathcal{T}_K$ orders the operators $\hat{f}_z$ along this contour.[4] The bracketed expression is thus the characteristic function of the fluctuations of $\int_0^t dt' \hat{f}_z(t')$ introduced in the context of full counting statistics. Now let us assume for simplicity that the noise is Gaussian, specified only by the variance which in this case is in the long-time limit (similar to what was done in Eqs. (9.8)-(9.10))

$$\int_0^t \int_0^t dt' dt'' \langle f_z(t') f_z(t'')\rangle = S_f(0)t,$$

where $S_f(0)$ describes the fluctuations in f_z at frequencies $\omega \lesssim 1/t$.

Finally, noting that $\mathrm{Tr}[\exp(-i\epsilon t/\hbar\sigma_z)] = \mathrm{Tr}[\cos(\epsilon t/\hbar) - i\sin(\epsilon t/\hbar)\sigma_z] = \cos(\epsilon t/\hbar)$, we get for the off-diagonal element of the density matrix the form

$$\rho_{od} = \frac{\alpha}{2}\cos\left(\frac{\epsilon t}{\hbar}\right)\exp\left(-\frac{g_z^2 S_f(0) t}{\hbar}\right). \tag{9.20}$$

We hence find that the fluctuations lead to the decay of the off-diagonal terms in the density matrix with the dephasing rate given by[5]

$$\Gamma_\phi^z = \frac{1}{\tau_\phi} = g_z^2 S_f(0)/\hbar. \tag{9.21}$$

Including the σ_x-term in the perturbation would have given us another dephasing rate,

$$\Gamma_\phi^x = g_x^2 S_x(-\epsilon)/(2\hbar), \tag{9.22}$$

where $S_x(\omega)$ is the noise power for the fluctuations of the field $f_x(t)$. The total dephasing rate is then the sum of these two contributions.[6]

[3] This can be proven by expanding the exponent and using the property $\sigma_z\sigma_x = -\sigma_x\sigma_z$.

[4] If $\hat{f}_z(t)$ commutes with itself at different time instants, this is simply twice the integral from $t = 0$ to t. This type of a model was actually used to define the quantum-mechanical characteristic function for the theory of full counting statistics.

[5] This terms is often called the "pure dephasing rate", as the total dephasing rate may also include the $S_x(\omega)$ term.

[6] As the dephasing and relaxation rates depend on the noise power of the environmental

9.3 Explicit model for dissipation: Caldeira-Leggett model

In the above section, we did not specify any given model for the fluctuating field. We also payed no attention to the back-action of the small quantum system on the noise source. An explicit model for dissipation in electric circuits was given by Amir Caldeira and the recent (2003) Nobel laureate Anthony Leggett.[7] They assumed that the quantum system under consideration is connected to a large bath of harmonic oscillators with different resonance frequencies ω_j, the latter mimicking the (almost infinitely many) microscopic degrees of freedom in the rest of the universe. The total Hamiltonian is then of the form $H = H_0 + H_{\rm CL}$, where H_0 again describes the unperturbed quantum system, and

$$H_{\rm CL} = \sum_{j=1}^{\infty} \left[\frac{\hat{p}_j(t)^2}{2m_j} + \frac{1}{2} m_j \omega_j^2 \left(\hat{x}_j(t) - \frac{\alpha_j}{m_j \omega_j^2} \hat{x}(t) \right)^2 \right]. \tag{9.23}$$

To fix the dimensions correctly, we now write the coordinate of the small quantum system as $\hat{x}(t)$ rather than $\hat{A}_i(t)$. In this model, α_j is the coupling parameter, which without a loss of generality can be chosen as $\alpha_j = m_j \omega_j^2$. Again we assume that the operators $\hat{x}_j$ commute with other operators $\hat{x}_k$ and with the operators $\hat{x}$, $\hat{p}$ of the small system.

Let us consider an electronic analogue of Eq. (9.23), using the correspondence Table 7.1 and the definition $\omega_j \equiv 1/\sqrt{L_j C_j}$. In a normal-metal circuit, one can also replace $2e$ by e. In this way, $H_{\rm CL}$ corresponds to an infinite set of LC oscillators in parallel to each other (see Fig. 9.1) and the circuit described by H_0 (for example, a Josephson junction).

From the total Hamiltonian $H_0 + H_{\rm CL} = \frac{\hat{q}(t)^2}{2C} + U(\phi(t)) + H_{\rm CL}$, we can find the equation of motion for the Heisenberg operator $\hat{\phi}(t)$. As now the exact time dependence of the environment operators $\hat{x}_j$ is known, the previous can be cast into the form

$$C\ddot{\phi}(t) + C \int_0^t ds \gamma(t-s)\dot{\phi}(s) + \nabla U(\phi) = \hat{\xi}(t). \tag{9.24}$$

Here the memory kernel $\gamma(t)$ depends on the oscillator bath through

$$\gamma(t) = \frac{\theta(t)}{C} \sum_{j=1}^{\infty} \frac{1}{L_j} \cos(\omega_j t). \tag{9.25}$$

---

fluctuations, they can be used for noise measurements. This is a quite novel idea perhaps first suggested in R. Schoelkopf, *et al.*, contribution to "Quantum Noise", edited by Yu. V. Nazarov, and Ya. M. Blanter, http://xxx.lanl.gov/abs/cond-mat/0210247. The first this kind of a measurement with a quantum bit was published by O. Astafiev, *et al.*, Phys. Rev. Lett. **93**, 267007 (2004).

[7] A. O. Caldeira and A. J. Leggett, Phys. Rev. Lett. **46**, 211 (1981); Physica A **121A**, 587 (1983); Phys. Rev. A **31**, 1059 (1985).

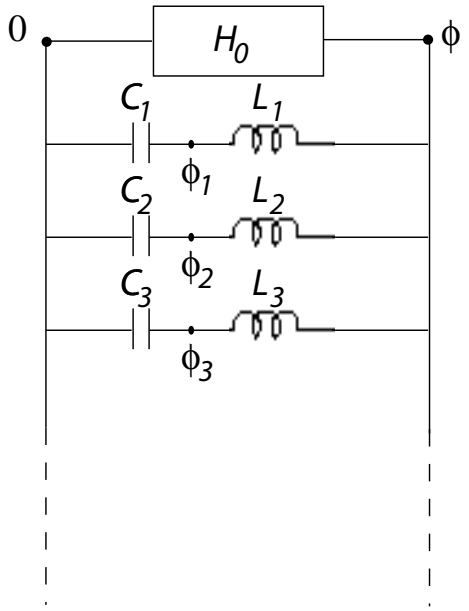


Figure 9.1: Circuit model of the Caldeira-Leggett Hamiltonian.

Hence, it is of a form of a Fourier cosine series. Therefore, fixing L_j and ω_j suitably, almost any type of a memory kernel can be modelled. The Langevin force term $\hat{\xi}(t)$ in Eq. (9.24) is defined through

$$\hat{\xi}(t) = \sum_{j=1}^{\infty} \frac{1}{L_j}[(\hat{\phi}_j(0) - \hat{\phi}(0))\cos(\omega_j t) + \frac{q_j(0)}{C_j \omega_j}\sin(\omega_j t)]. \tag{9.26}$$

As any proper Langevin term, its expectation value $\langle\hat{\xi}\rangle = \mathrm{Tr}[e^{-\beta H}\hat{\xi}] = 0$. Moreover, its second correlator satisfies $\langle\hat{\xi}(t)\hat{\xi}(t')\rangle \to k_B T\gamma(t-t')$ in the limit of a high temperature. This is one form of the fluctuation-dissipation theorem.

Now dissipative processes can be modelled by choosing a decaying memory kernel $\gamma(t)$, i.e., it makes the system "forget" its history. In general, $C\tilde{\gamma}(\omega)$ describes the admittance of the Caldera-Leggett element. Choosing $\gamma(t) \propto \delta(t)$ models a resistance. For example, with $\gamma(t) = Q^{-1}\delta(t)$ and $U(\phi)$ the tilted washboard potential Eq. (7.20), Eq. (9.24) becomes essentially Eq. (7.22) — the equation of motion for the phase ϕ in the RCSJ model with a fluctuating source term. Note that a decaying $\gamma(t)$ really requires an infinite number of oscillators, as a finite sum of $\cos(\omega_j t)$ functions would always show some type of an oscillatory behavior.[8]

[8] If we believe that the universe is finite, this would then mean that all the information about the past of any given system would "come back" to it at some later time. Anyway, this time may well exceed the age of the universe.

For an easier description of the heat bath, one often defines the density of oscillators,

$$J(\omega) \equiv \frac{\pi}{2}\omega \sum_{j=1}^{\infty} \frac{1}{L_j}\delta(\omega - \omega_j) = C\omega \mathrm{Re}[\tilde{\gamma(\omega)}]. \tag{9.27}$$

The Fourier transform of the memory kernel can be expressed through this density by

$$\tilde{\gamma}(\omega) = \frac{-2i\omega}{C\pi} \int_0^\infty d\omega' \frac{J(\omega')}{\omega'} \frac{1}{(\omega' - i0^+)^2 - \omega^2}. \tag{9.28}$$

The small imaginary part inserted in the denominator serves to yield a finite real part for $\tilde{\gamma}$ - this part then models dissipation. An "ohmic" density of oscillators, $J(\omega) = \omega/R$ describes the impedance $Z_{\mathrm{CL}} = R$, i.e., the effect of a classical resistor.[9]

The Caldera-Leggett model is only one possible model for the description of quantum dissipation. Its usefulness lies in the fact that it is exactly soluble, as the algebra of harmonic oscillators is relatively easy. One problem (or benefit, depending on the viewpoint) of this model is that it can only describe *linear environments*: for an oscillator, the linear response theory is exact, as the commutators between $\hat{p}$ and $\hat{x}$ are scalars. This feature can be traced to the fact that the oscillators essentially describe bosonic degrees of freedom. However, for example most models for $1/f$-noise are based on an ensemble of two-level systems, which are nonlinear. For modelling such environments, baths of fermions (essentially spins) have been considered. However, they are far beyond the scope of this course.

9.4 Problems

1. Assume a two-level system described by the Hamiltonian

$$H_0 = \epsilon\sigma_z + \Delta\sigma_x \tag{9.29}$$

is coupled to the environment through a force $f\sigma_z$. The spectrum of the fluctuations of f is given by a coupling constant times a thermal noise in a resistor:

$$S_f(\omega) = 2g^2 \frac{R_N}{R_Q} \hbar\omega \coth\left(\frac{\hbar\omega}{2k_BT}\right), \tag{9.30}$$

where g is a dimensionless constant and $R_Q = h/4e \approx 6.5k\Omega$. Calculate the dephasing and relaxation times of the qubit assuming that this is the only source of noise. Then, try to get an estimate of these times for $\Delta = 0.01\epsilon$, ϵ/k_B=1K, $g^2 = 0.01$ and $R_N = 50\Omega$ at temperature $T = 50$mK.

[9] In fact, often one also needs a cutoff frequency in Eq. (9.28) to avoid diverging integrals. Such a cutoff in the case of a resistor is given by the $1/(RC)$-frequency.

2. Consider a quantum system coupled to a "bath" with equilibrium fluctuations obeying the fluctuation-dissipation theorem with temperature T. Show that the steady-state occupation numbers of this system satisfy detailed balance,

$$\frac{p_m}{p_n} = \exp\left(\frac{E_n - E_m}{k_B T}\right). \tag{9.31}$$

What is the resulting steady-state energy distribution function of the quantum system? Why is it of this form?

Chapter 10

Summary of the main mesoscopic effects

In this course, we have chosen to take a fairly broad definition of a "mesocopic" effect, or a "nanoelectronic" system. Sometimes, the word mesoscopic is chosen to refer only to effects related with the finite number of quantum modes in the resistor and to the single-electron effects, and with this definition, a fairly practical limitation for a "nanoelectronic" system is a system with dimensions less than 100 nm. Here we have chosen to describe effects that one tends to first encounter when the electric circuits are miniaturized below the everyday scales. Therefore, we have viewed effects related to relaxation length scales, such as the nonequilibrium distribution functions and shot noise, also as mesoscopic effects. The aim of this chapter is to summarize the main effects and models discussed in this course.

10.1 Main mesoscopic effects

When we describe mesoscopic or quantum effects on electron transport, it is necessary to know what a "non-mesoscopic", or a "macroscopic" behavior is. In the absence of superconductivity, the electron transport in macroscopic wires can typically be described with a "billiard ball" model, such as the Drude model introduced in Subs. 1.1.1. In such a conductor, the properties of the system can be described via the thermodynamic quantities, such as the temperature T and the potential V. The current-voltage characteristics is given by the Ohm's law,

$$V = RI, \tag{10.1}$$

and the temperature and voltage independent R scales linearly with the length of the wire. Moreover, the fluctuations of the current at zero frequency are described by the fluctuation-dissipation relation

$$S_I(\omega) = \frac{4k_BT}{R}. \tag{10.2}$$

This description applies very well for most macroscopic normal conductors.

10.1.1 Interference effects

One of the landmark mesoscopic effects are those related with the wave-like character of the electron probability amplitude, and the related interference of the waves. Typically this interference is controlled via the magnetic field, that can turn constructive interference into destructive. The clearest interference effect is the *Aharonov-Bohm effect* realized in a multiply connected structure, see Sec. 3.1. There the interference of two paths the same electron may take shows up directly as an oscillation of the transmission probability ($\sim$ conductance).

The interference shows up also in a singly connected wire in the presence of impurities. Quantum-mechanically, the electron can take many different paths between two points in the wire. If the two points are distinct, the interference of the different paths averages out when the wire is sufficiently large or when one ensemble averages over different realizations of disorder. However, the interference still shows up in the *universal conductance fluctuations* (Sec. 3.3) between different realizations. Calculating the interference effect between closed paths, i.e., those that return to the same point, shows that the amplitude for the electron to backscatter from the impurities is enlarged compared to a simple billiard-ball model. As this localization (Sec. 3.2) is an interference effect, it vanishes in a magnetic field or when the phase breaking length becomes small.

10.1.2 Relaxation

In a nonequilibrium setting, the electron energy distribution function within a normal-metal wire may not follow the usual (equilibrium) Fermi-Dirac form. The convergence towards such an equilibrium shape is described by relaxation inside the wire. Only when the wire is much longer than certain (generally voltage-dependent) relaxation scales, one may for example talk about a well-defined local temperature. The main such relaxation effects are electron-electron and electron-phonon scattering. Below 1 K, the relaxation lengths corresponding to these scattering events may well exceed a few micrometers, and thus the nonequilibrium form of the distribution function is fairly straightforward to find.

This nonequilibrium form shows up especially in the shot noise of the current, which decreases as the electron-phonon coupling length becomes smaller (for example, due to an increasing temperature).

10.1.3 Single-electron effects

Although the Coulomb interaction between the electrons is the strongest force above the atomic scales, a surprisingly large part of the mesoscopic phenomena related to electron transport can be described with almost no account of the interactions. The interaction effects are mostly visible if one constructs an "island" where a relatively small metallic wire is connected to an outside world via small tunnel junctions with total capacitance C (this capacitance includes

also that between the island and the ground plain, therefore the island has to be sufficiently small). In this case, the interactions inside the island show up as an energy scale $E_c = e^2/(2C)$ which is required for adding or subtracting electrons from the island compared to the equilibrium charge (Sec. 5.1). In single-electron transistors, this equilibrium charge can be modulated by tuning a voltage in a gate electrode (Sec. 5.1.2).

The charging effects are quite robust and tunable and making the required systems with present lithographic techniques is relatively easy. Therefore they can be used in many types of applications, such as ultrasensitive charge detection or thermometry and highly accurate current pumping (Sec. 5.6).

10.1.4 Mesoscopic superconductivity

At the low temperatures where most mesoscopic phenomena are studied, many materials become superconductors. Whereas superconductivity is a quantum effect, it persists to macroscopic length scales and is not directly a mesoscopic effect. Many of the mesoscopic effects can be found also in (wholly or partially) superconducting systems, but superconductivity makes rise also to totally new types of phenomena, such as the Josephson effect and the superconducting proximity effect. For example, a combination of the single-electron and Josephson effects allows one to engineer and study an electrical analogue of almost any type of a model Hamiltonian. This fact can be used in realizing new types of applications,[1] but also for the study of the fundamental quantum-mechanical phenomena, such as relaxation or dephasing, or quantum measurements. The main illustrations of this are the harmonic oscillator limit in the case $k_BT, E_c \ll E_J$, and the two-state system realized either in the charge regime $E_c \gg E_J$, in the strongly biased case with $E_c \ll E_J$, or using closed three-junction SQUIDs (these are the charge, phase, and flux qubits, see Fig. 10.1). The use of such quantum two-state systems formed from Josephson elements for quantum computing is under an active study at present.[2]

10.1.5 Shot noise, correlations and statistics

One of the most fashionable research topics in the past few years are the electronic fluctuations in mesoscopic circuits. While sometimes they cause the harmful "noise" for the measurements, one can in principle also extract information from fluctuation measurements. This information is contained especially in the shot noise, whose zero-frequency autocorrelation function is typically

[1] Examples of this include the Cooper-pair pump, "sluice" studied in the Low Temperature laboratory and at the state research center VTT, see A. O. Niskanen, J. P. Pekola, and H. Seppä, Phys. Rev. Lett. **91**, 177003 (2003), where the pumping utilizes the fact that the Josephson energy can be tuned with an external flux, the SINIS cooler where superconductivity allows to cool a normal-metal island, or a supercurrent transistor where a supercurrent is controlled via the control of the nonequilibrium distribution function. The latter two are described for example in F. Giazotto, *et al.*, Rev. Mod. Phys. **78**, 217 (2006).

[2] See Yu. Makhlin, A. Shnirman, and G. Schön, Rev. Mod. Phys. **73**, 357 (2001).

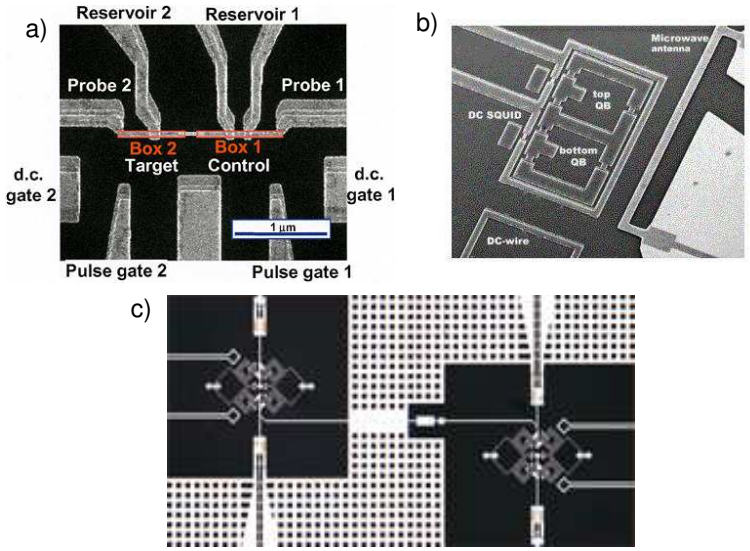


Figure 10.1: Josephson qubits in pairs: a) NEC charge qubit, b) Delft flux qubit, and c) NIST phase qubit. The figures have been obtained from a) http://www.riken.go.jp/engn/r-world/research/lab/frontier/quantum/coherence/result.html b) http://qt.tn.tudelft.nl/research/fluxqubit/fluxqubit.html and c) http://qdev.boulder.nist.gov/817.03/whatwedo/qcomputing/qcomputing.htm.

proportional to the average current $\langle I \rangle$,

$$S_I = 2eF\langle I \rangle. \tag{10.3}$$

The proportionality constant, the *Fano factor* F depends on the type of the sample. In the strictly normal-metal limit, it describes how the transmission probabilities in the different channels are distributed, but it can show the effective charge of current carriers in a condensed state, such as in superconductors, or with the quantum Hall effect.

Shot noise is a mesoscopic effect as it decreases when the sample size is increased beyond the electron-phonon length. As the latter is dependent on the local electron temperature, shot noise can also be used to study the electron-phonon coupling strength as a function of the temperature or the driving bias voltage.

In multiterminal setups, also the cross-correlations between the current fluctuations at different electrodes may yield information on the system. The most remarkable feature about the cross correlation is their sign, and its dependence on the statistics of the current carriers: for noninteracting fermions, the cross correlations are negative, which is a result of the fermion "antibunching", whereas for bosons (such as photons), the sign of the cross correlations can change between negative and positive values. A positive cross correlation

would then show a "bunching" effect, as a signature of many bosons occupying the same state.

Average current and shot noise describe the first and second moments of the distribution of transmission probabilities. The information on the full probability distribution can at least in principle be accessed by studying the "full counting statistics" of charge transport. Studies of this counting statistics for example show that the probability for the transmission of electrons through a normal-metal wire obeys a simple binomial distribution. Another feature one has found is that even the equilibrium fluctuations of mesoscopic samples is not Gaussian, but the fluctuations contain higher even cumulants than the second also present in a Gaussian distribution. One expects[3] that for large wires, the statistics tends to a Gaussian form as required by the central limit theorem.

10.2 Different theoretical models

Along with the mesoscopic effects, we have introduced a few theoretical models with which to describe them quantitatively. The major models are the scattering theory, Boltzmann and Boltzmann-Langevin theory, and the Orthodox theory for describing single-electron effects in the lowest order in the transmission.

10.2.1 Scattering theory

Scattering theory was introduced in Ch. 2. The main point of the theory is to introduce a mathematical object called leads between the studied scattering region and the electrodes. These leads define a set of transverse modes between which the scattering takes place. The scattering is described by a scattering matrix $s_{nm}^{\alpha\beta}$, where n and m index the modes and α, β the leads (electrodes). This scattering matrix then divides into parts describing transmission between two different leads and reflection back into the same lead. The eigenvalues of the part describing the transmission are the transmission probability amplitudes, and the square of their absolute value are the corresponding transmission probabilities, denoted by T_n.

From the scattering theory, one can find an expression for the average current and zero-frequency shot noise in terms of the transmission probabilities. In the two-probe case these are

$$\langle I \rangle = \frac{2e^2}{h} V \sum_n T_n \tag{10.4a}$$

$$S_I = \frac{4e^3}{h} V \sum_n T_n (1 - T_n). \tag{10.4b}$$

The multi-probe version of the current formula is given in Eq. (2.41), and that for the noise can be obtained from Eq. (8.34).

[3] Although this has not been strictly proven yet.

Scattering theory is especially useful when describing the general properties of arbitrary mesoscopic objects, ballistic or almost ballistic systems, systems with only a few transmission channels, or systems where the probability distribution of transmission eigenvalues is known. It cannot directly describe inelastic scattering or dephasing effects.[4]

10.2.2 Boltzmann (-Langevin) theory

The semiclassical Boltzmann theory describes the behavior of the electron energy distribution function $f(\vec{r}, E, t)$. By construction, it neglects all the quantum-mechanical interference effects, but it is especially useful for treating the effects of elastic and inelastic scattering. When the stochastic nature of the scattering is taken into account, one can formulate a Boltzmann-Langevin model, which then can also be utilized to describe the fluctuations. In a diffusive wire interference effects are relatively small and the Boltzmann approach is applicable. There, the fluctuating distribution function satisfies the kinetic equation

$$D\nabla^2 f = l\nabla \cdot \int \hat{p}\xi d\hat{p} + I_{\rm inel}, \tag{10.5}$$

where ξ is the zero-average fluctuating force, and $I_{\rm inel}$ is the collision integral for inelastic scattering. A typical assumption for the correlator of the fluctuating force is a Poissonian white noise, its energy integral given in Eq. (8.56). This correlator can be used in order to express the total current noise with the non-fluctuating part of the distribution function. Solving this non-fluctuating part $\bar{f}(\vec{r}, E)$, we can obtain the average current and zero-frequency noise from

$$\langle I \rangle = -\frac{G_N}{e} \int_{-\infty}^{\infty} \nabla \bar{f}(\vec{r}, E) \tag{10.6a}$$

$$S_I = \frac{4G_N}{L} \int_0^L dx \int dE \bar{f}_0(E, x)(1 - \bar{f}_0(E, x)). \tag{10.6b}$$

Here $G_N = e^2 N_0 D$ is the normal-state conductance of the wire, N_0 is the density of states at the Fermi level, and D is the diffusion constant. In the expression for the noise, the spatial integrals run between the ends of the diffusive wire, connected to large non-fluctuating electrodes.

10.2.3 Master equation approaches

For some systems typically weakly coupled to the external world, the most useful description of the system dynamics can be carried out by considering the occupation number P_n of a given energy state n (this is the diagonal entry of the system density matrix). This energy state can be a charge state in a single-electron transistor, the state of a qubit, or some other well-defined state.

[4] Although there are formulations for doing this by using fictitious probes, see Ya. Blanter and M. Büttiker, Phys. Rep. **336**, 1 (2000).

The coupling to the external world introduces transitions between the different levels n and m with the rate Γ_{mn} (see for example Eq. (5.18)). In the case of a single-electron transistor, these transitions are caused by the tunnelling of electrons between the island and the reservoirs. The time dependence of the occupation numbers are then described by a *master equation* of the form

$$\frac{dP_n}{dt} = \sum_m (\Gamma_{nm} P_m - \Gamma_{mn} P_n). \tag{10.7}$$

Typically only the static solution for which $\Gamma_{nm} P_m = \Gamma_{mn} P_n$ is required. When the transition rates are known (they are typically obtained from the Fermi Golden rule), P_n can be calculated. With the knowledge of these P_n, one can then obtain other system properties, for example the current through a single-electron transistor (see Eq. (5.21)).

10.3 Systems and effects not described in the course

A decision on what to include in a course on nanoelectronics is necessarily always also a decision on what to exclude from it. Whereas we have discussed many of the main effects studied still today, we have not considered (or at most have only mentioned) at least the following topics

- Properties of quantum dots with interactions. Especially with semiconductor systems (see Fig.1.4) and molecules one can realize so small islands, that the level spacing δE for the electron kinetic energy on the island becomes of the order of or larger than the temperature. In this case, our description of the single-electron effects in Ch. 5 has to be modified to include this additional energy scale. For example, the Coulomb diamond introduced in Exercise 6 gets distorted: for each added electron onto the island, one has to provide at least the energy $E_c + \delta E$. But as δE varies from level to level, the addition spectrum of quantum dots is not equispaced. Quantum dots are very versatile devices, there are for example suggestions on how to build quantum bits using double quantum dot systems.

- Kondo effect is a many-body effect related to the spin degree of freedom of the electrons. It was first discovered in bulk metals with magnetic impurities: there its effect is to increase resistance with a decreasing temperature. In quantum dots, it is a new type of a higher-order tunnelling effect.

- 0.7 anomaly. In Ch. 2, we discussed the conductance quantization in quantum point contacts and how the conductance can be altered by tuning the gate voltage. What was not mentioned was the fact that on a closer look to the measured curves, the conductance is not quite quantized at

the integer multiples of the conductance quantum, but it shows smaller-scale features. The first observation of these features showed a small conductance step at the value of 0.7 G_0, where the name arises. This effect has not been fully understood yet.

- Interacting one-dimensional electron systems form a new form of a correlated electron system, that cannot be described with the conventional Fermi liquid theory. For example, in these *Luttinger liquids*, the elementary excitations are bosonic in nature. It is believed that single-wall nanotubes are realizations of Luttinger liquids, but there is still some controversy on whether the existing observations prove this or if they can be explained via some other means.

- Placing two narrow ballistic wires near each other and driving a current through one of them makes rise to a current in the other wire. This is the *Coulomb drag effect*.

- Mesoscopic effects in exotic electron systems, such as strongly correlated fermion systems, quantum Hall systems, or high-temperature superconductors. These effects are mostly still not very well understood.

- Statistics of electronic heat transport has been recently described in addition to the statistics of the charge transport. The heat current noise is a relevant quantity for example in radiation detectors.

Many of these topics are not very well understood and they are under an active study at present. But also in the mesoscopic effects discussed in these notes there are many unsolved questions. Therefore, the field of mesoscopic electronics will still quite some time provide a fruitful platform, both for new types of applications, as well as for fundamental research.