

Lecture 19

NEGF Method

19.1. One-level Resistor

19.2. Multi-level Resistors

19.3. Conductance Functions for Coherent Transport

19.4. Elastic Dephasing

In the last Lecture I tried to provide a super-brief but hopefully self-contained introduction to the Hamiltonian matrix $[H]$ whose eigenvalues tell us the allowed energy levels in the channel. However, $[H]$ describes an isolated channel and we cannot talk about the steady-state resistance of an isolated channel without bringing in the contacts and the battery connected across it. In this Lecture I will describe the NEGF-based transport model that can be used to model current flow, given $[H]$ and the $[\Sigma]$'s (Fig.19.1).

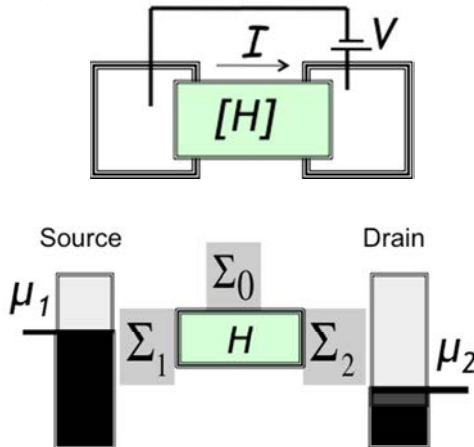
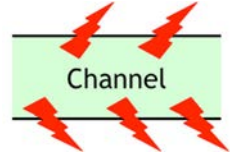


Fig.19.1. The NEGF-based quantum transport model described here allows us to model current flow given the Hamiltonian matrix $[H]$ describing the channel, the self-energy matrices $[\Sigma]$ describing the connection of the channel to the contacts, and $[\Sigma_0]$ describing interactions within the channel.

As I mentioned in Lectures 1 and 18, the NEGF method originated from the seminal works of Martin and Schwinger (1959), Kadanoff and Baym (1962), Keldysh (1965) and others who used the methods of many-body perturbation theory (MBPT) to describe the distributed entropy-generating processes along the channel which were believed to constitute the essence of resistance. Since MBPT is an advanced topic requiring many semesters to master, the NEGF method is generally regarded as an esoteric tool for specialists.



In our bottom-up approach we will start with elastic resistors for which energy exchange is confined to the contacts, and the problem of resistance can be treated within a one-electron picture by connecting contacts to the Schrödinger equation. Indeed our approach will be to start from the usual time-independent Schrödinger equation $E \{ \psi \} = [H] \{ \psi \}$ and add two terms to it representing the outflow and inflow from the contact

$$E \{ \psi \} = [H] \{ \psi \} + \underbrace{[\Sigma] \{ \psi \}}_{\text{OUTFLOW}} + \underbrace{\{ s \}}_{\text{INFLOW}}$$

These two terms arise from imposing open boundary conditions on the Schrödinger equation with an incident wave from the contact as shown in Chapters 8,9 of Datta (2005). Some readers may notice the similarity of the additional terms here with those appearing in the Langevin equation used to describe Brownian motion, but we will not go into it.

Using this modified Schrödinger equation, the wavefunction can be written as

$$\{ \psi \} = [EI - H - \Sigma]^{-1} \{ s \}$$

We will then argue that since the inflow from multiple sources $\{s\}$ are incoherent, one cannot superpose the resulting $\{ \psi \}$'s and it is more convenient to work in terms of quantities like (superscript '+' denotes conjugate transpose)

$$[G^n] \sim \{\psi\}\{\psi\}^+$$

$$[\Sigma^{in}] \sim \{s\}\{s\}^+$$

which can be superposed. Defining

$$G^R = [EI - H - \Sigma]^{-1} \quad (19.1)$$

$$\text{and } G^A = [G^R]^+$$

we can write

$$\{\psi\} = [G^R]\{s\}$$

so that

$$\underbrace{\{\psi\}\{\psi\}^+}_{G^n} = [G^R] \underbrace{\{s\}\{s\}^+}_{\Sigma^{in}} [G^A]$$

giving us the second NEGF equation

$$G^n = G^R \Sigma^{in} G^A \quad (19.2)$$

Though we have changed the notation somewhat, writing Σ for Σ^R , and (see Chapter 8, Datta 1995)

$$G^n \text{ for } -iG^< \quad \text{and} \quad \Sigma^{in} \text{ for } -i\Sigma^<$$

Eqs.(19.1, 19.2) are essentially the same as Eqs.(75-77) in Keldysh (1965), which is one of the seminal founding papers on the NEGF method that obtained these equations using MBPT. Although for simplicity we have only discussed the time-independent version here, a similar derivation could be used for the time-dependent version too (See Appendix, Datta (2005)).

How could we obtain these results using elementary arguments, without invoking MBPT? Because we are dealing with an elastic resistor where all entropy-generating processes are confined to the contacts and can be handled in a relatively elementary manner. But should we call this NEGF?

It seems to us that NEGF has two aspects, namely

- A. Eqs.(19.1), (19.2) and
- B. calculating $[\Sigma]$, $[\Sigma^{\text{in}}]$ that appear in Eqs.(19.1), (19.2).

For historical reasons, these two aspects, A and B, are often intertwined in the literature, but they need not be. Indeed these two aspects are completely distinct in the Boltzmann formalism (Lecture 7). The Boltzmann transport equation (BTE)

$$\frac{\partial f}{\partial t} + \vec{v} \cdot \vec{\nabla} f + \vec{F} \cdot \vec{\nabla}_p f = S_{op} f \quad (\text{same as Eq.(7.5)})$$

is used to describe semiclassical transport in many different contexts, but the evaluation of the scattering operator S_{op} has evolved considerably since the days of Boltzmann and varies widely depending on the problem at hand.

Similarly it seems to me that the essence of NEGF is contained in Eqs.(19.1), (19.2) while the actual evaluation of the $[\Sigma]$'s may well evolve as we look at more and more different types of problems. The original MBPT-based approach may or may not be the best, and may need to be modified even for problems involving electron-electron interactions.

Above all we believe that by decoupling Eqs.(19.1) and (19.2) from the MBPT method originally used to derive them, we can make the NEGF method more transparent and accessible so that it can become a part of the standard training of physics and engineering students who need to apply it effectively to a wide variety of basic and applied problems that require connecting contacts to the Schrödinger equation.

I should also note briefly the relation between the NEGF method applied to elastic resistors with the scattering theory of transport or the transmission formalism widely used in mesoscopic physics. Firstly, the scattering theory works directly with the Schrödinger equation with open

boundary conditions that effectively add the inflow and outflow terms we mentioned:

$$E \{ \psi \} = [H] \{ \psi \} + \underbrace{[\Sigma]}_{\text{OUTFLOW}} \{ \psi \} + \underbrace{\{s\}}_{\text{INFLOW}}$$

However, as we noted earlier it is then important to add individual sources incoherently, something that the NEGF equation (Eq.(19.2)) takes care of automatically.

The second key difference is the handling of dephasing processes in the channel, something that has no classical equivalent. In quantum transport randomization of the phase of the wavefunction even without any momentum relaxation can have a major impact on the measured conductance. The scattering theory of transport usually neglects such dephasing processes and is restricted to *phase-coherent elastic resistors*.

Incoherence is commonly introduced in this approach using an insightful observation due to Büttiker that dephasing processes essentially remove electrons from the channel and re-inject them just like the voltage probes discussed in Section 12.2 and so one can include them phenomenologically by introducing conceptual contacts in the channel.

This method is widely used in mesoscopic physics, but it seems to introduce both phase and momentum relaxation and I am not aware of a convenient way to introduce pure phase relaxation if we wanted to. In the NEGF method it is straightforward to choose $[\Sigma_0]$ so as to include phase relaxation with or without momentum relaxation as we will see in the next Lecture. In addition, the NEGF method provides a rigorous framework for handling all kinds of interactions in the channel, both elastic and inelastic, using MBPT. Indeed that is what the original work from the 1960's was about.

Let me finish up this long introduction by briefly mentioning the two other key equations in NEGF besides Eqs.(19.1) and (19.2). As we will see, the quantity $[G^n]$ appearing in Eq.(19.2) represents a matrix version of the electron density (times 2π) from which other quantities of interest

can be calculated. Another quantity of interest is the matrix version of the density of states (again times 2π) called the spectral function $[A]$ given by

$$\begin{aligned} A &= G^R \Gamma G^A = G^A \Gamma G^R \\ &= i[G^R - G^A] \end{aligned} \quad (19.3a)$$

where G^R , G^A are defined in Eq.(19.1) and the $[\Gamma]$'s represent the **anti-Hermitian** parts of the corresponding $[\Sigma]$'s

$$\Gamma = i[\Sigma - \Sigma^+] \quad (19.3b)$$

which describe how easily the electrons in the channel communicate with the contacts.

There is a component of $[\Sigma]$, $[\Gamma]$, $[\Sigma^{in}]$ for each contact (physical or otherwise) and the quantities appearing in Eqs.(19.1-19.3) are the total obtained summing all components. The current at a specific contact m , however, involves only those components associated with contact m :

$$\tilde{I}_m = \frac{q}{h} \text{Trace}[\Sigma_m^{in} A - \Gamma_m G^n] \quad (19.4)$$

Note that $\tilde{I}_m(E)$ represents the current per unit energy and has to be integrated over all energy to obtain the total current. In the following four Lectures we will look at a few examples designed to illustrate how Eqs.(19.1)-(19.4) are applied to obtain concrete results.

But for the rest of this Lecture let me try to justify these equations. We start with a one-level version for which all matrices are just numbers (Section 19.1), then look at the full multi-level version (Section 19.2), obtain an expression for the conductance function $G(E)$ for coherent transport (Section 19.3) and finally look at the different choices for the dephasing self-energy $[\Sigma_0]$ (Section 19.4).

19.1. One-level resistor

To get a feeling for the NEGF method, it is instructive to look at a particularly simple conductor having just one level and described by a 1×1 $[H]$ matrix that is essentially a number: $[H] = \mathcal{E}$.

Starting directly from the Schrödinger equation we will see how we can introduce contacts into this problem. This will help set the stage for Section 19.3 when we consider arbitrary channels described by $(N \times N)$ matrices instead of the simple one-level channel described by (1×1) “matrices.”

19.1.1. Semiclassical treatment

It is useful to first go through a semiclassical treatment as an intuitive guide to the quantum treatment. Physically we have a level connected to two contacts, with two different occupancy factors

$$f_1(\mathcal{E}) \text{ and } f_2(\mathcal{E})$$

Let us assume the occupation factor to be one for the source and zero for the drain, so that it is only the source that is continually trying to fill up the level while the drain is trying to empty it. We will calculate the resulting current and then multiply it by

$$f_1(\mathcal{E}) - f_2(\mathcal{E})$$

to account for the fact that there is injection from both sides and the net current is the difference.

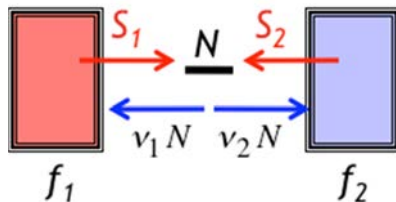


Fig.19.2.
Filling and emptying a level:
Semiclassical picture

With $f_1=1$ in the source and $f_2=0$ in the drain, the average number N of electrons ($N < I$) should obey an equation of the form

$$\frac{d}{dt}N = -(\nu_1 + \nu_2)N + S_1 + S_2 \quad (19.5)$$

where ν_1 and ν_2 represent the rates (per second) at which an electron escapes into the source and drain respectively, while S_1 is the rate at which electrons try to enter from the source. The steady state occupation is obtained by setting

$$\frac{d}{dt}N = 0 \rightarrow N = \frac{S_1 + S_2}{\nu_1 + \nu_2} \quad (19.6)$$

We can fix S_1 , by noting that if the drain were to be disconnected, N should equal the Fermi function $f_1(\mathcal{E})$ in contact 1, which we will assume one for this discussion. This means

$$\frac{S_1}{\nu_1} = f_1(\mathcal{E}) \quad \text{and} \quad \frac{S_2}{\nu_2} = f_2(\mathcal{E}) \quad (19.7)$$

The current can be evaluated by writing Eq.(19.5) in the form

$$\frac{dN}{dt} = (S_1 - \nu_1 N) + (S_2 - \nu_2 N) \quad (19.8)$$

and noting that the first term on the right is the current from the source while the second is the current into the drain. Under steady state conditions, they are equal and either could be used to evaluate the current that flows in the circuit:

$$I = q(S_1 - \nu_1 N) = q(\nu_2 N - S_2) \quad (19.9)$$

From Eqs.(19.6), (19.7) and (19.9), we have

$$N = \frac{\nu_1 f_1(\mathcal{E}) + \nu_2 f_2(\mathcal{E})}{\nu_1 + \nu_2} \quad (19.10a)$$

$$I = q \frac{V_1 V_2}{V_1 + V_2} (f_1(\varepsilon) - f_2(\varepsilon)) \quad (19.10b)$$

and

19.1.2. Quantum treatment

Let us now work out the same problem using a quantum formalism based on the Schrödinger equation. In the last Chapter we introduced the matrix version of the time-independent Schrödinger equation

$$E\{\psi\} = [H]\{\psi\}$$

which can be obtained from the more general time-dependent equation

$$i\hbar \frac{\partial}{\partial t} \{\tilde{\psi}(t)\} = [H]\{\tilde{\psi}(t)\} \quad (19.11a)$$

by assuming
$$\{\tilde{\psi}(t)\} = \{\psi\} e^{-iEt/\hbar} \quad (19.11b)$$

For problems involving steady-state current flow, the time-independent version is usually adequate, but sometimes it is useful to go back to the time-dependent version because it helps us interpret certain quantities like the self-energy functions as we will see shortly.

In the quantum formalism the squared magnitude of the electronic wavefunction $\tilde{\psi}(t)$ tells us the probability of finding an electron occupying the level and hence can be identified with the average number of electrons $N (< I)$. For a single isolated level with $[H] = \varepsilon$, the time evolution of the wavefunction is described by

$$i\hbar \frac{d}{dt} \tilde{\psi} = \varepsilon \tilde{\psi}$$

which with a little algebra leads to

$$\frac{d}{dt} (\tilde{\psi} \tilde{\psi}^*) = 0$$

showing that for an isolated level, the number of electrons $\tilde{\psi}\tilde{\psi}^*$ does not change with time.

Our interest, however, is not in isolated systems, but in channels connected to two contacts. Unfortunately the standard quantum mechanics literature does not provide much guidance in the matter, but we can do something relatively simple using the rate equation in Eq.(19.4) as a guide.

We introduce *contacts into the Schrödinger equation* by modifying it to read

$$i\hbar\frac{d}{dt}\tilde{\psi} = \left(\varepsilon - i\frac{\gamma_1 + \gamma_2}{2} \right) \tilde{\psi} \quad (19.12a)$$

so that the resulting equation for

$$\frac{d}{dt}\tilde{\psi}\tilde{\psi}^* = -\left(\frac{\gamma_1 + \gamma_2}{\hbar} \right) \tilde{\psi}\tilde{\psi}^* \quad (19.12b)$$

looks just like Eq.(19.5) except for the source term S_I which we will discuss shortly.

We can make Eq.(19.12b) match Eq.(19.5) if we choose

$$\gamma_1 = \hbar\nu_1 \quad (19.13a)$$

$$\gamma_2 = \hbar\nu_2 \quad (19.13b)$$

We can now go back to the *time-independent version* of Eq.(19.12a):

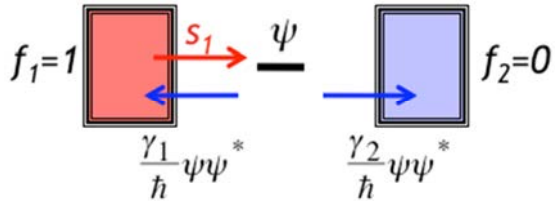
$$E\psi = \left(\varepsilon - i\frac{\gamma_1 + \gamma_2}{2} \right) \psi \quad (19.14)$$

obtained by assuming a single energy solution:

$$\tilde{\psi}(t) = \psi(E) e^{-iEt/\hbar}$$

Eq.(19.14) has an obvious solution $\psi=0$, telling us that at steady-state there are no electrons occupying the level which makes sense since we have not included the source term S_l . All electrons can do is to escape into the contacts, and so in the long run the level just empties to zero.

Fig.19.3.
Filling and emptying a level: Quantum picture



Introducing a source term into Eq.(19.11) and defining $\gamma \equiv \gamma_1 + \gamma_2$, we have

$$E\psi = \left(\varepsilon - i\frac{\gamma}{2} \right) \psi + s_1 \tag{19.15}$$

Unlike the semiclassical case (Eq.(19.5)) we are introducing only one source rather than two. The reason is subtle and we will address it later at the end of this Section. From Eq.(19.15), we can relate the wavefunction to the source

$$\psi = \frac{s_1}{E - \varepsilon + i(\gamma/2)} \tag{19.16}$$

Note that the wavefunction is a maximum when the electron energy E equals the energy ε of the level, as we might expect. But the important point about the quantum treatment is that the wavefunction is not significantly diminished as long as E differs from ε by an amount less than γ . This is an example of “broadening” or energy uncertainty that a semiclassical picture misses.

To obtain the **strength of the source** we require that the total number of electrons on integrating over all energies should equal our rate equation result from Eq.(19.5). that is,

$$\int_{-\infty}^{+\infty} dE \psi \psi^* = \frac{v_1}{v_1 + v_2} = \frac{\gamma_1}{\gamma_1 + \gamma_2} \quad (19.17)$$

where we have made use of Eq.(19.13). We now use Eqs.(19.16), (19.17) to evaluate the right hand side in terms of the source

$$\int_{-\infty}^{+\infty} dE \psi \psi^* = \int_{-\infty}^{+\infty} dE \frac{s_1 s_1^*}{(E - \varepsilon)^2 + \left(\frac{\gamma}{2}\right)^2} = \frac{2\pi s_1 s_1^*}{\gamma} \quad (19.18)$$

where we have made use of a standard integral

$$\int_{-\infty}^{+\infty} dE \frac{\gamma}{(E - \varepsilon)^2 + \left(\frac{\gamma}{2}\right)^2} = 2\pi \quad (19.19)$$

From Eqs.(19.17) and (19.18) we obtain, noting that

$$2\pi s_1 s_1^* = \gamma_1 \quad (19.20)$$

The strength of the source is thus proportional to the escape rate which seems reasonable: if the contact is well coupled to the channel and electrons can escape easily, they should also be able to come in easily.

Just as in the semiclassical case (Eq.(19.9)) we obtain the **current** by looking at the rate of change of N from Eq.(19.12b)

$$\frac{d}{dt} \tilde{\psi} \tilde{\psi}^* = (\text{Inflow from 1}) - \frac{\gamma_1}{\hbar} \tilde{\psi} \tilde{\psi}^* - \frac{\gamma_2}{\hbar} \tilde{\psi} \tilde{\psi}^*$$

where we have added a term “Inflow from 1” as a reminder that Eq.(19.12a) does not include a source term. Both left and right hand sides of this equation are zero for the steady-state solutions we are considering. But just like the semiclassical case, we can identify the current as either the first two terms or the last term on the right:

$$\frac{I}{q} = (\text{Inflow from 1}) - \frac{\gamma_1}{\hbar} \tilde{\psi} \tilde{\psi}^* = \frac{\gamma_2}{\hbar} \tilde{\psi} \tilde{\psi}^*$$

Using the second form and integrating over energy we can write

$$I = q \int_{-\infty}^{+\infty} dE \frac{\gamma_2}{\hbar} \psi \psi^* \quad (19.21)$$

so that making use of Eqs.(19.16) and (19.20), we have

$$I = \frac{q \gamma_1 \gamma_2}{\hbar 2\pi} \int_{-\infty}^{+\infty} dE \frac{1}{(E - \varepsilon)^2 + (\gamma/2)^2} \quad (19.22)$$

which can be compared to the semiclassical result from Eq.(19.10) with $f_1=1, f_2=0$ (note: $\gamma = \gamma_1 + \gamma_2$)

$$I = \frac{q \gamma_1 \gamma_2}{h \gamma_1 + \gamma_2}$$

19.1.3. Quantum broadening

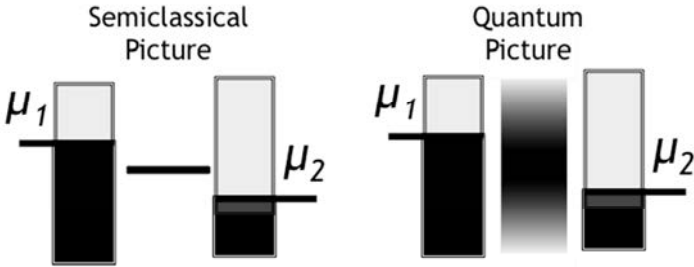
Note that Eq.(19.22) involves an integration over energy, as if the quantum treatment has turned the single sharp level into a continuous distribution of energies described by a density of states $D(E)$:

$$D = \frac{\gamma/2\pi}{(E - \varepsilon)^2 + (\gamma/2)^2} \quad (19.23)$$

Quantum mechanically the process of coupling inevitably spreads a single discrete level into a state that is distributed in energy, but integrated over all energy still equals one (see Eq.(19.15)). One could call it a consequence of the **uncertainty relation**

$$\gamma t \geq h$$

relating the length of time t the electron spends in a level to the uncertainty " γ " in its energy. The stronger the coupling, shorter the time and larger the broadening.



Is there any experimental evidence for this energy broadening (Eq.(19.23)) predicted by quantum theory? A hydrogen molecule has an energy level diagram like the one-level resistor we are discussing and experimentalists have measured the conductance of a hydrogen molecule with good contacts and it supports the quantum result (Smit et al. 2002). Let me elaborate a little.

Comparing Eq.(19.22) with Eq.(3.3) for elastic resistors we can write the conductance function for a one-level device including quantum broadening as

$$G(E) = \frac{q^2}{h} \frac{\gamma_1 \gamma_2}{(E - \varepsilon)^2 + \left(\frac{\gamma}{2}\right)^2}$$

If we assume (1) equal coupling to both contacts:

$$\gamma_1 = \gamma_2 = \frac{\gamma}{2}$$

and (2) a temperature low enough that the measured conductance equals $G(E=\mu_0)$, μ_0 being the equilibrium electrochemical potential, we have

$$G \approx G(E = \mu_0) = \frac{q^2}{h} \frac{(\gamma/2)^2}{(\mu_0 - \varepsilon)^2 + (\gamma/2)^2}$$

So the quantum theory of the one-level resistor says that the measured conductance should show a maximum value equal to the quantum of conductance q^2/h when μ_0 is located sufficiently close to ε . The experimentally measured conductance is equal to $2q^2/h$, the extra factor of 2 being due to spin degeneracy, since levels come in pairs and what we have is really a two-level rather than a one-level resistor.

19.1.4. Do Multiple Sources Interfere?

In our quantum treatment we considered a problem with electrons injected only from the source ($f_1 = 1$) with the drain empty ($f_2 = 0$) (Eq.(19.15)), unlike the semiclassical case where we started with both sources S_1 and S_2 (Eq.(19.5)).

This is not just a matter of convenience. If instead of Eq.(19.15) we start from

$$E\psi = \left(\varepsilon - i\frac{\gamma}{2} \right) \psi + s_1 + s_2$$

we obtain

$$\psi = \frac{s_1 + s_2}{E - \varepsilon + i\frac{\gamma}{2}}$$

so that

$$\psi\psi^* = \frac{1}{(E - \varepsilon)^2 + \left(\frac{\gamma}{2}\right)^2} (s_1s_1^* + s_2s_2^* + \underbrace{s_1s_2^* + s_2s_1^*}_{\substack{\text{Interference} \\ \text{Terms}}})$$

which has two extra interference terms that are never observed experimentally because the electrons injected from separate contacts have uncorrelated phases that change randomly in time and average to zero.

The first two terms on the other hand add up since they are positive numbers. It is like adding up the light from two light bulbs: we add their powers not their electric fields. Laser sources on the other hand can be coherent so that we actually add electric fields and the interference terms can be seen experimentally. Electron sources from superconducting contacts too can be coherent leading to Josephson currents that depend on interference. But that is a different matter.

Our point here is simply that normal contacts like the ones we are discussing are incoherent and it is necessary to take that into account in our models. The moral of the story is that we cannot just insert multiple sources into the Schrödinger equation. We should insert one source at a time, calculate bilinear quantities (things that depend on the product of wavefunctions) like electron density and current and add up the contributions from different sources. Next we will describe the non-equilibrium Green function (NEGF) method that allows us to implement this procedure in a systematic way and also to include incoherent processes.

19.2. Quantum transport through multiple levels

We have seen how we can treat quantum transport through a one-level resistor with a time-independent Schrödinger equation modified to include the connection to contacts and a source term:

$$E\psi = \left(\varepsilon - i\frac{\gamma}{2} \right) \psi + s$$

How do we extend this method to a more general channel described by an $N \times N$ Hamiltonian matrix $[H]$ whose eigenvalues give the N energy levels?

For an N -level channel, the wavefunction $\{\psi\}$ and source term $\{s_1\}$ are $N \times 1$ column vectors and the modified Schrödinger equation looks like

$$E\{\psi\} = [H + \Sigma_1 + \Sigma_2]\{\psi\} + \{s_1\} \quad (19.24)$$

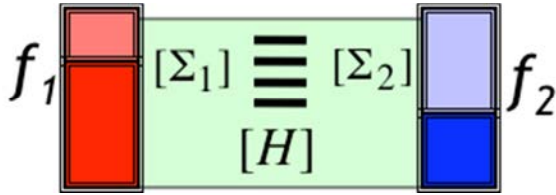
where Σ_1 and Σ_2 are $N \times N$ non-Hermitian matrices whose anti-Hermitian components

$$\Gamma_1 = i[\Sigma_1 - \Sigma_1^\dagger]$$

$$\Gamma_2 = i[\Sigma_2 - \Sigma_2^\dagger]$$

play the roles of $\gamma_{1,2}$ in our one-level problem.

Fig.19.4.
Transport model for multi-level conductor.



In Lecture 18 we discussed how for different structures we can write down the channel Hamiltonian $[H]$ and in the next few Lectures I will present examples to show how the $[\Sigma]$ are obtained.

For the moment, let us focus on how the basic NEGF equations summarized earlier (Eqs.(19.1)-(19.4)) follow from our contact-ed Schrödinger equation, Eq.(19.24).

19.2.1. Obtaining Eqs.(19.1)

From Eq.(19.24) it is straightforward to write

$$\{\psi\} = [G^R] \{s_1\}$$

where G^R is given by Eq.(19.1) with

$$\Sigma = \Sigma_1 + \Sigma_2 \tag{19.25}$$

19.2.2. Obtaining Eqs.(19.2)

The matrix electron density, G^n , defined as

$$G^n \rightarrow 2\pi \{\psi\} \{\psi\}^+ = 2\pi [G^R] \{s_1\} \{s_1\}^+ [G^A]$$

where the superscript "+" stands for conjugate transpose, and G^A stands for the conjugate transpose of G^R .

For the one-level problem $2\pi s_1 s_1^* = \gamma_1$ (see Eq.(19.20)): the corresponding matrix relation is

$$2\pi \{s_1\} \{s_1\}^+ = [\Gamma_1]$$

so that

$$G^n = [G^R][\Gamma_1][G^A]$$

This is for a single source term. For multiple sources, the electron density matrices, unlike the wavefunctions, can all be added up with the appropriate Fermi function weighting to give Eq.(19.2),

$$G^n = [G^R][\Sigma^{in}][G^A] \quad (\text{same as 19.2})$$

with Σ^{in} representing an incoherent sum of all the independent sources:

$$[\Sigma^{in}] = [\Gamma_1] f_1(E) + [\Gamma_2] f_2(E) \quad (19.26)$$

19.2.3. Obtaining Eq.(19.3)

Eq.(19.2) gives us the electron density matrix G^n , in terms of the Fermi functions f_1 and f_2 in the two contacts. But if both f_1 and f_2 are equal to one then all states are occupied, so that the matrix electron density becomes equal to the matrix density of states, called the spectral function matrix $[A]$. Setting $f_1 = 1$ and $f_2 = 1$, in Eq.(19.2) we have

$$[A] = [G^R][\Gamma][G^A] \quad (19.27)$$

since $\Gamma = \Gamma_1 + \Gamma_2$. This gives us part of Eq.(19.3). The rest of Eq.(19.3) can be obtained from Eq.(19.1) using straightforward algebra as follows:

$$[G^R]^{-1} = EI - H - \Sigma \quad (19.28a)$$

Taking conjugate transpose of both sides

$$[[G^R]^{-1}]^+ = [[G^R]^+]^{-1} = EI - H - \Sigma^+ \quad (19.28b)$$

Subtracting Eq.(19.28b) from (19.28a) (note that G^A stands for $[G^R]^+$) and making use of Eq.(19.3b)

$$[G^R]^{-1} - [G^A]^{-1} = i[\Gamma] \quad (19.28c)$$

Multiplying with $[G^R]$ from the left and $[G^A]$ from the right we have

$$i[[G^R] - [G^A]] = G^R \Gamma G^A$$

thus giving us another piece of Eq.(19.3). The final piece is obtained by multiplying Eq.(19.28c) with $[G^A]$ from the left and $[G^R]$ from the right.

19.2.4. Obtaining Eq.(19.4): The Current Equation

Like the semiclassical treatment and the one-level quantum treatment, the current expression is obtained by considering the time variation of the number of electrons N . Starting from

$$i\hbar \frac{d}{dt} \{\psi\} = [H + \Sigma] \{\psi\} + \{s\}$$

and its conjugate transpose (noting that H is a Hermitian matrix)

$$-i\hbar \frac{d}{dt} \{\psi\}^+ = \{\psi\}^+ [H + \Sigma^+] + \{s\}^+$$

we can write

$$\begin{aligned}
 i\hbar \frac{d}{dt} \{\psi\} \{\psi\}^+ &= \left(i\hbar \frac{d}{dt} \{\psi\} \right) \{\psi\}^+ + \{\psi\} \left(i\hbar \frac{d}{dt} \{\psi\}^+ \right) \\
 &= ([H + \Sigma] \{\psi\} + \{s\}) \{\psi\}^+ - \{\psi\} \left(\{\psi\}^+ [H + \Sigma^+] + \{s\}^+ \right) \\
 &= [(H + \Sigma)\psi\psi^+ - \psi\psi^+(H + \Sigma^+)] + [ss^+G^A - G^Rss^+]
 \end{aligned}$$

where we have made use of the relations

$$\{\psi\} = [G^R] \{s\} \quad \text{and} \quad \{\psi\}^+ = \{s\}^+ [G^A]$$

Since the trace of $[\psi\psi^+]$ represents the number of electrons, we could define its time derivative as a matrix current operator whose trace gives us the current. Noting further that

$$2\pi \{\psi\} \{\psi\}^+ = [G^n] \quad \text{and} \quad 2\pi \{s\} \{s\}^+ = [\Gamma]$$

we can write

$$I^{op} = \frac{[HG^n - G^nH] + [\Sigma G^n - G^n\Sigma^+] + [\Sigma^{in}G^A - G^R\Sigma^{in}]}{i2\pi\hbar} \quad (19.29)$$

We will talk more about the current operator in Lecture 22 when we talk about spins, but for the moment we just need its trace which tells us the time rate of change of the number of electrons N in the channel

$$\frac{dN}{dt} = \frac{-i}{\hbar} \text{Trace} \left([\Sigma G^n - G^n\Sigma^+] + [\Sigma^{in}G^A - G^R\Sigma^{in}] \right)$$

noting that $\text{Trace} [AB] = \text{Trace} [BA]$. Making use of Eq.(19.3b)

$$\frac{dN}{dt} = \frac{1}{\hbar} \text{Trace} \left[\Sigma^{in}A - \Gamma G^n \right]$$

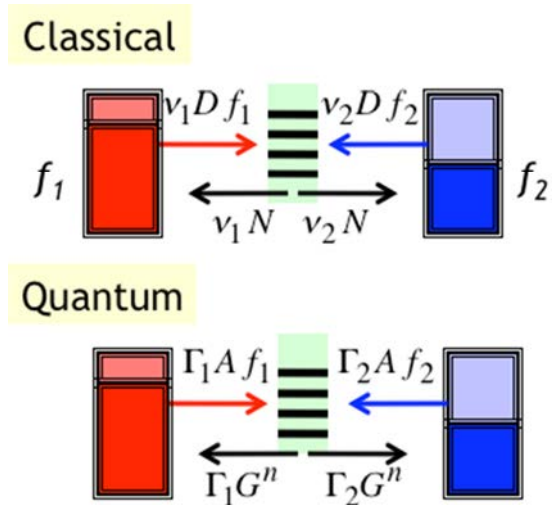
Now comes a tricky argument. Both the left and the right hand sides of Eq.(19.29) are zero, since we are discussing steady state transport with no time variation. The reason we are spending all this time discussing something that is zero is that the terms on the left can be separated into two parts, one associated with contact 1 and one with contact 2. They tell us the currents at contacts 1 and 2 respectively and the fact that they add up to zero is simply a reassuring statement of Kirchoff's law for steady-state currents in circuits.

With this in mind we can write for the current at contact m ($m=1,2$)

$$\tilde{I}_m = \frac{q}{h} \text{Trace} \left[\Sigma_m^{in} A - \Gamma_m G^n \right]$$

as stated earlier in Eq.(19.4). This leads us to the picture shown in Fig.19.5 where we have also shown the semiclassical result for comparison.

Fig.19.5:
Filling and emptying a channel: Classical and quantum treatment.



19.3. Conductance Functions for Coherent Transport

Finally we note that using Eqs.(19.2)-(19.3) we can write the current from Eq.(19.4) a little differently

$$\tilde{I}(E) = \frac{q}{h} \text{Trace}[\Gamma_1 G^R \Gamma_2 G^A] (f_1(E) - f_2(E))$$

which is very useful for it suggests a quantum expression for the conductance function $G(E)$ that we introduced in Lecture 3 for all elastic resistors:

$$G(E) = \frac{q^2}{h} \text{Trace} \left[\Gamma_1 G^R \Gamma_2 G^A \right] \quad (19.30)$$

More generally with multiterminal conductors we could introduce a self-energy function for each contact and show that

$$\tilde{I}_m = \frac{q}{h} \sum_r \bar{T}_{mr} (f_m(E) - f_n(E)) \quad (19.31)$$

with

$$\bar{T}_{mn} \equiv \text{Trace} \left[\Gamma_m G^R \Gamma_n G^A \right] \quad (19.32)$$

For low bias we can use our usual Taylor series expansion from Eq.(2.8) to translate the Fermi functions into electrochemical potentials so that Eq.(19.31) looks just like the Büttiker equation (Eq.(12.3)) with the conductance function given

$$G_{m,n}(E) \equiv \frac{q^2}{h} \text{Trace} \left[\Gamma_m G^R \Gamma_n G^A \right] \quad (19.33)$$

which is energy-averaged in the usual way for elastic resistors (see Eq.(3.1)).

$$G_{m,n} = \int_{-\infty}^{+\infty} dE \left(-\frac{\partial f_0}{\partial E} \right) G_{m,n}(E)$$

19.4. Elastic Dephasing

So far we have focused on the physical contacts described by $[\Sigma_{1,2}]$ and the model as it stands describes coherent quantum transport where electrons travel coherently from source to drain in some static structure described by the Hamiltonian $[H]$ without any interactions along the channel described by $[\Sigma_0]$ (Fig.19.1). In order to include $[\Sigma_0]$, however, no change is needed as far as Eqs.(19.1) through (19.4) is concerned. It is just that an additional term appears in the definition of Σ , Σ^{in} :

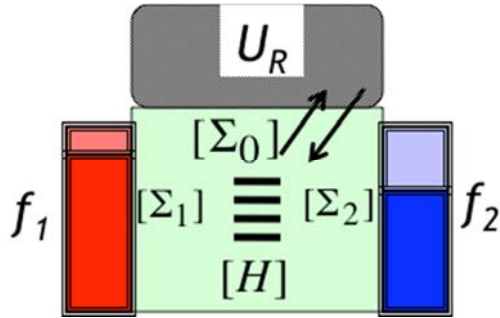
$$\Sigma = \Sigma_1 + \Sigma_2 + \Sigma_0$$

$$\Gamma = \Gamma_1 + \Gamma_2 + \Gamma_0$$

$$[\Sigma^{in}] = [\Gamma_1] f_1(E) + [\Gamma_2] f_2(E) + [\Sigma_0^{in}] \quad (19.34)$$

What does $[\Sigma_0]$ represent physically? From the point of view of the electron a solid does not look like a static medium described by $[H]$, but like a rather turbulent medium with a random potential U_R that fluctuates on a picosecond time scale. Even at fairly low temperatures when phonons have been frozen out, an individual electron continues to see a fluctuating potential due to all the other electrons, whose average is modeled by the scf potential we discussed in Section 18.2. These fluctuations do not cause any overall loss of momentum from the system of electrons, since any loss from one electron is picked up by another. However, they do cause fluctuations in the phase leading to fluctuations in the current. What typical current measurements tell us is an average flow over nanoseconds if not microseconds or milliseconds. This averaging effect needs to be modeled if we wish to relate to experiments.

Fig.19.6.
Quantum transport model with
simple elastic dephasing.



As we mentioned earlier, the NEGF method was originally developed in the 1960's to deal with the problem of including inelastic processes into a quantum description of large conductors. For the moment, however, we will look at simple elastic dephasing processes leaving more general interactions for Lecture 23.

For such processes the self-energy functions are given by

$$[\Sigma_0] = D \times [G^R] \quad (19.35a)$$

$$[\Sigma_0^{in}] = D \times [G^n] \quad (19.35b)$$

where \times denotes element by element multiplication. Making use of the relations in Eqs.(19.3), it is straightforward to show from Eq.(19.35a) that

$$[\Gamma_0] = D \times [A] \quad (19.35c)$$

The elements of the matrix $[D]$ represent the correlation between the random potential at location “i” and at location “j”:

$$D_{ij} = \langle U_{Ri} U_{Rj} \rangle \quad (19.36)$$

Two cases are of particular interest. The first is where the random potential is well-correlated throughout the channel having essentially the same value at all points “i” so that the every element of the matrix $[D]$ has essentially the same value D_0 :

$$\textbf{Model A:} \quad D_{ij} = D_0 \quad (19.37)$$

The other case is where the random potential has zero correlation from one spatial point i to another j , so that

$$\textbf{Model B:} \quad D_{ij} = D_0, \quad i = j \quad \text{and} \quad = 0, \quad i \neq j \quad (19.38)$$

Real processes are usually somewhere between the two extremes represented by models A and B.

To see where Eqs.(19.35) come from we go back to our contact-ed Schrödinger equation

$$E \{ \psi \} = [H + \Sigma_1 + \Sigma_2] \{ \psi \} + \{ s_1 \}$$

and noting that a random potential U_R should lead to an additional term that could be viewed as an additional source term

$$E \{ \psi \} = [H + \Sigma_1 + \Sigma_2] \{ \psi \} + U_R \{ \psi \} + \{ s_1 \}$$

with a corresponding inscattering term given by

$$\Sigma_0^{in} = 2\pi U_R U_R^* \{ \psi \} \{ \psi \}^+ = D_0 G^n$$

corresponding to Model A (Eq.(19.37)) and a little more careful argument leads to the more general result in Eq.(19.36). That gives us Eq.(19.35b). How about Eq.(19.35a) and (19.35c) ?

The simplest way to justify Eq.(19.35c) is to note that together with Eq.(19.35b) (which we just obtained) it ensures that the current at terminal 0 from Eq.(19.4) equals zero:

$$\begin{aligned} I_0 &= \frac{q}{h} \text{Trace}[\Sigma_0^{in} A - \Gamma_0 G^n] \\ &= \frac{q}{h} \text{Trace}[G^n \Gamma_0 - \Gamma_0 G^n] = 0 \end{aligned}$$

This is a required condition since terminal 0 is not a physical contact where electrons can actually exit or enter from.

Indeed a very popular method due to Büttiker introduces incoherent processes by including a fictitious probe (often called a Büttiker probe) whose electrochemical potential is adjusted to ensure that it draws zero current. In NEGF language this amounts to assuming

$$\Sigma_0^{in} = \Gamma_0 f_P$$

with the number f_P is adjusted for zero current. This would be equivalent to the approach described here if the probe coupling Γ_0 were chosen proportional to the spectral function $[A]$ as required by Eq.(19.35c).

Note that our prescription in Eq.(19.35) requires a “self-consistent evaluation” since Σ , Σ^{in} depend on G^R and G^n which in turn depend on Σ , Σ^{in} respectively (see Eqs.(19.1), (19.2)).

Also, Model A (Eq.(19.37)) requires us to calculate the full Green’s function which can be numerically challenging for large devices described by large matrices. Model B makes the computation numerically much more tractable because one only needs to calculate the diagonal elements of the Green’s functions which can be done much faster using powerful algorithms.

In these Lectures, however, we focus on conceptual issues using “toy” problems for which numerical issues are not the “show stoppers.” The important conceptual distinction between Models A and B is that the former destroys phase but not momentum, while the latter destroys momentum as well [Golizadeh-Mojarad et al. 2007].

The dephasing process can be viewed as extraction of the electron from a state described by $[G^n]$ and reinjecting it in a state described by $D \times G^n$. Model A is equivalent to multiplying $[G^n]$ by a constant so that the electron is reinjected in exactly the same state that it was extracted in, causing no loss of momentum, while Model B throws away the off-

diagonal elements and upon reinjection the electron is as likely to go on way or another. Hopefully this will get clearer in the next Lecture when we look at a concrete example.

Another question that the reader might raise is whether instead of including elastic dephasing through a self-energy function $[\Sigma_0]$ we could include a potential U_R in the Hamiltonian itself and then average over a number of random realizations of U_R . The answer is that the two methods are not exactly equivalent though in some problems they could yield similar results. This too should be a little clearer in the next lecture when we look at a concrete example.

For completeness, let me note that in the most general case D_{ijkl} is a fourth order tensor and the version we are using (Eq.(19.35)) represents a special case for which D_{ijkl} is non-zero only if $i=k, j=l$ (see Appendix E).

This page intentionally left blank

Lecture 20

Can Two Offer Less Resistance than One?

20.1. Modeling 1D Conductors

20.2. Quantum Resistors in Series

20.3. Potential Drop Across Scatterer(s)

In the next three Lectures we will go through a few examples of increasing complexity which are interesting in their own right but have been chosen primarily as “do it yourself” problems that the reader can use to get familiar with the quantum transport model outlined in the last Lecture. The MATLAB codes are all included in Appendix F.

In this Lecture we will use 1D quantum transport models to study an interesting question regarding multiple scatterers or obstacles along a conductor. Are we justified in neglecting all interference effects among them and assuming that electrons diffuse like classical particles as we do in the semiclassical picture?

This was the question Anderson raised in his 1958 paper entitled "Absence of Diffusion in Certain Random Lattices" pointing out that diffusion could be slowed significantly and even suppressed completely due to quantum interference between scatterers. “Anderson localization” is a vast topic and we are only using some related issues here to show how the NEGF model provides a convenient conceptual framework for studying interesting physics.

For any problem we need to discuss how we write down the Hamiltonian $[H]$ and the contact self-energy matrices $[\Sigma]$. Once we have these, the computational process is standard. The rest is about understanding and enjoying the physics.

20.1. Modeling 1D Conductors

For the one-dimensional examples discussed in this Lecture, we use the 1-D Hamiltonian from Fig.18.6, shown here in Fig.20.1. As we discussed earlier for a uniform wire the dispersion relation is given by

$$E(k) = \varepsilon + 2t \cos ka \quad (20.1a)$$

which can approximate a parabolic dispersion

$$E = E_c + \frac{\hbar^2 k^2}{2m} \quad (20.1b)$$

by choosing

$$E_c = \varepsilon + 2t \quad (20.2a)$$

and

$$-t \equiv t_0 \equiv \frac{\hbar^2}{2ma^2} \quad (20.2b)$$

It is straightforward to write down the $[H]$ matrix with ε on the diagonal and “ t ” on the upper and lower diagonals. What needs discussion are the **self-energy matrices**. The basic idea is to replace an infinite conductor described by the Hamiltonian $[H]$ with a finite conductor described by $[H + \Sigma_1 + \Sigma_2]$ assuming **open boundary conditions** at the ends, which means that electron waves escaping from the surface do not give rise to any reflected waves, as a good contact should ensure.

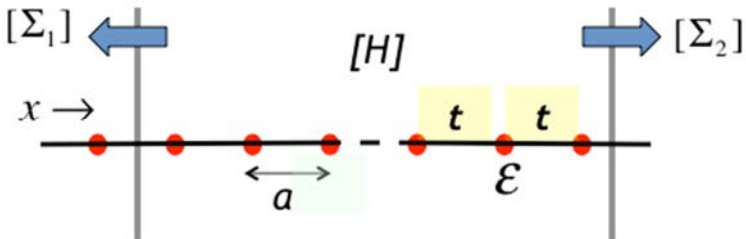


Fig.20.1. For the one-dimensional examples discussed in this Lecture, we use the 1-D Hamiltonian from Fig.18.6.

For a one-dimensional lattice the idea is easy to see. We start from the original equation for the extended system

$$E \psi_n = t \psi_{n-1} + \epsilon \psi_n + t \psi_{n+1}$$

and then assume that the contact has no incoming wave, just an outgoing wave, so that we can write

$$\psi_{n+1} = \psi_n e^{ika}$$

which gives
$$E \psi_n = t \psi_{n-1} + (\epsilon + t e^{ika}) \psi_n$$

In other words the effect of the contact is simply to add $t^* \exp(+ika)$ to H_{nn} which amounts to adding the self-energy

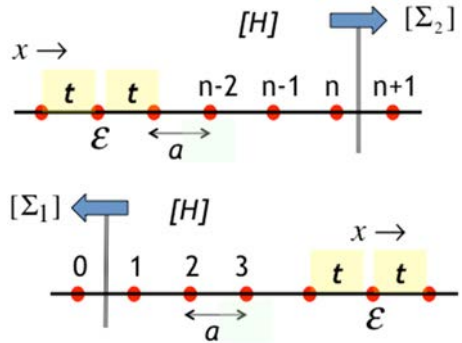
$$\Sigma_1 = \begin{bmatrix} t e^{ika} & 0 & 0 & \dots \\ 0 & 0 & 0 & \\ 0 & 0 & 0 & \\ \dots & & & \ddots \end{bmatrix} \quad (20.1a)$$

to the Hamiltonian. Note the only non-zero element is the (1,1) element. Similarly at the other contact we obtain

$$\Sigma_2 = \begin{bmatrix} \ddots & & \dots \\ & 0 & 0 & 0 \\ \dots & 0 & 0 & 0 \\ & 0 & 0 & t e^{ika} \end{bmatrix} \quad (20.1b)$$

Note the only non-zero element is the (n,n) element.

In short, the self-energy function for each contact has a single non-zero element corresponding to the point that is connected to that contact.



20.1.1. 1D ballistic conductor

A good test case for any theory of coherent quantum transport is the conductance function for a length of uniform ballistic conductor: If we are doing things right, the conductance function $G(E)$ should equal the quantum of conductance q^2/h times an integer equal to the number of modes $M(E)$ which is one for 1D conductors (neglecting spin). This means that the transmission (see Eq.(19.30))

$$\bar{T}(E) = \text{Trace} \left[\Gamma_1 G^R \Gamma_2 G^A \right] \quad (20.3)$$

should equal one over the energy range

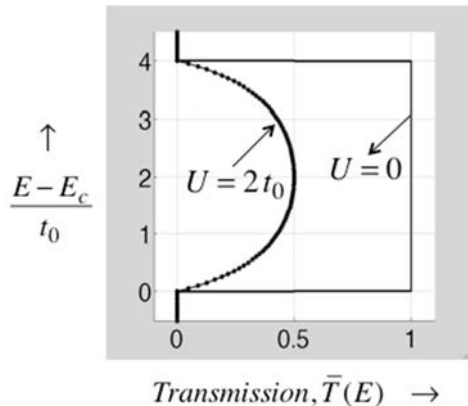
$$0 < E - E_c < 4t_0$$

covered by the dispersion relation

$$E = \varepsilon + 2t \cos ka = E_c + 2t_0(1 - \cos ka) \quad (20.4)$$

but zero outside this range (see Fig.20.2 below with $U=0$). This is a relatively simple but good example to try to implement numerically when getting started. Obtaining a constant conductance across the entire band is usually a good indicator that the correct self-energy functions are being used and things have been properly set up.

Fig.20.2. Transmission through a single point scatterer in a 1D wire.



20.1.2. 1D conductor with one scatterer

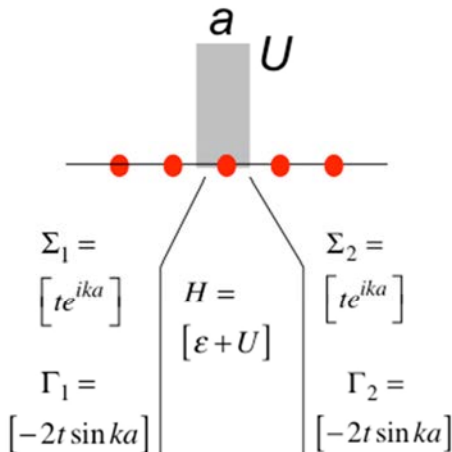
Another good example is that of a conductor with just one scatterer whose effect is included in the Hamiltonian $[H]$ by changing the diagonal element corresponding to that point to $\epsilon + U$:

$$H = \begin{bmatrix} \ddots & & & & & \\ & \dots & & & & \\ & & \epsilon & t & 0 & \\ \vdots & & t & \epsilon + U & t & \\ & & 0 & t & \epsilon & \\ & & & \dots & & \ddots \end{bmatrix}$$

Fig.20.2 shows the numerical results for $U=0$ (ballistic conductor) and for $U = 2t_0$. Actually there is a simple analytical expression for the transmission through a single point scatterer

$$\bar{T}(E) = \frac{(2t \sin ka)^2}{U^2 + (2t \sin ka)^2} = \frac{(\hbar v / a)^2}{U^2 + (\hbar v / a)^2} \quad (20.5)$$

that we can use to check our numerical results. This expression is obtained by treating the single point where the scatterer is located as the channel, so that all matrices in the NEGF method are (1x1) matrices, that is, just numbers:



It is easy to see that the Green's function is given by

$$G^R(E) = \frac{1}{E - (\varepsilon + U) - 2t e^{ika}} = \frac{1}{-U - i2t \sin ka}$$

making use of Eq.(20.2). Hence

$$\Gamma_1 G^R \Gamma_2 G^A = \frac{(2t \sin ka)^2}{U^2 + (2t \sin ka)^2}$$

giving us the stated result in Eq.(20.3). The second form is obtained by noting from Eq.(20.2) that

$$\hbar v = \frac{dE}{dk} = -2at \sin ka \quad (20.6)$$

Once you are comfortable with the results in Fig.20.2 and are able to reproduce it, you should be ready to include various potentials into the Hamiltonian and reproduce the rest of the examples in this Lecture.

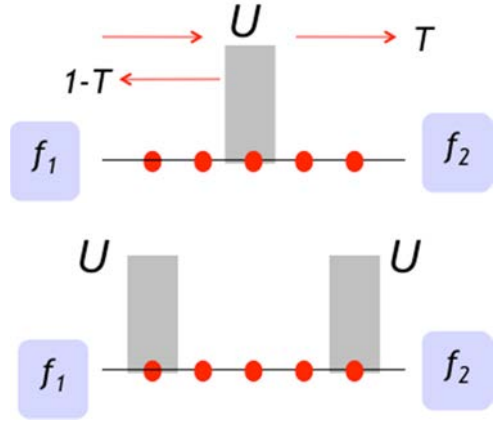
20.2. Quantum Resistors in Series

In Lecture 12 we argued that the resistance of a conductor with one scatterer with a transmission probability T can be divided into a scatterer resistance and an interface resistance (see Eqs.(12.1), (12.2))

$$R_1 = \frac{h}{q^2 M} \left(\underbrace{\frac{1-T}{T}}_{\text{scatterer}} + \underbrace{1}_{\text{interface}} \right)$$

What is the resistance if we have two scatterers each with transmission T ?

Fig.20.3.
Channel with (a) one scatterer, and
(b) two scatterers.



We would expect the scatterer contribution to double:

$$R_2 = \frac{h}{q^2 M} \left(2 \underbrace{\frac{1-T}{T}}_{\text{scatterer}} + \underbrace{1}_{\text{interface}} \right)$$

$$= \frac{h}{q^2 M} \frac{2-T}{T}$$

We can relate the two resistances by the relation:

$$R_2 = R_1(2-T)$$

If T is close to one we have the ballistic limit with $R_2 = R_1$: two sections in series have the same resistance as one of them, since all the resistance comes from the interfaces.

If $T \ll 1$, we have the Ohmic limit with $R_2 = 2R_1$: two sections have twice the resistance as one of them, since all the resistance comes from the channel.

But can R_2 ever be less than R_1 ? Not as long as electrons behave like classical particles. There is no way an extra roadblock on a classical highway can increase the traffic flow. But on a quantum highway this is quite possible due to wave interference.

We could use our 1D model to study problems of this type. Fig.20.4 shows the transmission functions $\bar{T}(E)$ calculated numerically for a conductor with one scatterer and a conductor with two scatterers.

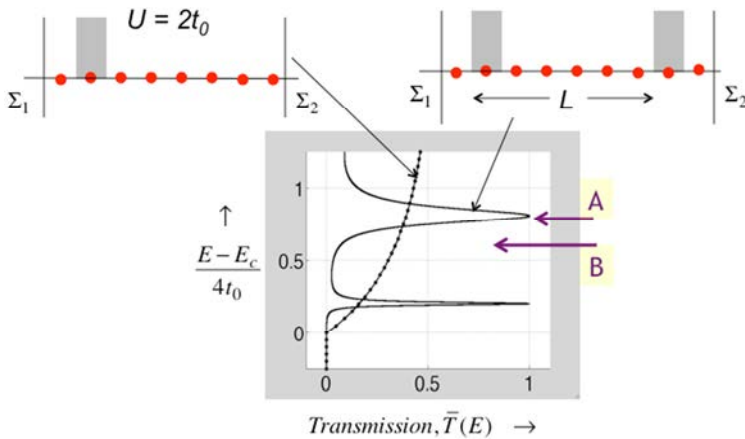
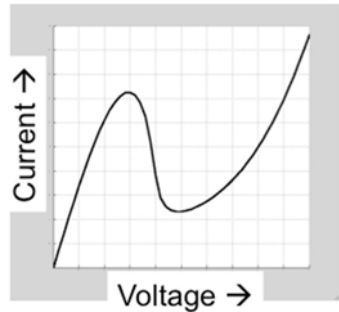


Fig.20.4. Normalized conductance for a wire with $M=I$ with (a) one scatterer, and (b) two scatterers.

If the electrochemical potential happens to lie at an energy like the one marked "B", R_2 will be even larger than the Ohmic result R_1 . But if the electrochemical potential lies at an energy like the one marked "A", R_2 is less than R_1 .

At such energies, the presence of the second scatterer creates a reflection that cancels the reflection from the first one, because they are spaced a quarter wavelength apart. Such quarter wave sections are widely used to create anti-reflection coatings on optical lenses and are well-known in the world of waves, though they are unnatural in the world of particles.

Actually there is a class of devices called resonant tunneling diodes that deliberately engineer two strategically spaced barriers and make use of the resulting sharp peaks in conductance to achieve interesting current-voltage characteristics like the one sketched here where over a range of voltages, the slope dI/dV is negative (“negative differential resistance, NDR”). We could use our elastic resistor model for the current from Eq.(3.3) and along with the conductance function from NEGF



$$G(E) \equiv \frac{q^2}{h} \bar{T}(E) = \frac{q^2}{h} \text{Trace} \left[\Gamma_1 G^R \Gamma_2 G^A \right]$$

to model devices like this, but it is important to include the effect of the applied electric field on the $[H]$ as mentioned earlier (see Fig.18.6). In these Lectures we will focus more on low bias response for which this aspect can be ignored.

Consider for example a resistor with scatterers distributed randomly throughout the channel. If we were to use the quantum formalism to calculate the conductance function for a single-moded wire with random scatterers we would find that once the classical transmission $M \lambda / L$ drops below one, the quantum conductance is extremely low except for occasional peaks at specific energies (Fig.20.5). The result marked semiclassical is obtained by calculating T for a single scatterer and then increasing the scatterer contribution by a factor of six:

$$R_6 = \frac{h}{q^2 M} \left(6 \underbrace{\frac{1-T}{T}}_{\text{scatterer}} + \underbrace{1}_{\text{interface}} \right) = \frac{h}{q^2 M} \frac{6-5T}{T}$$

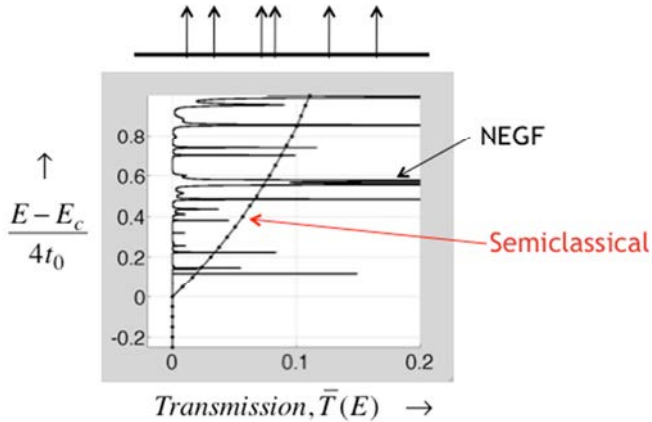


Fig.20.5. Normalized conductance for a wire with $M=1$ with six scatterers

Comparing the classical and quantum results suggests that such conductors would generally show very high resistances well in excess of Ohm's law, with occasional wild fluctuations. In a multi-moded wire too quantum calculations show the same behavior once the classical transmission $M\lambda/L$ drops below one. Such conductors are often referred to as being in the regime of strong localization. Interestingly, even when $M\lambda/L$ is well in excess of one, the quantum conductance is a little (\sim approximately one) less than the classical value and this is often referred to as the regime of weak localization.

However, localization effects like these are usually seen experimentally only at low temperatures. At room temperature there is seldom any evidence of deviation from Ohm's law. Consider for instance a copper wire with a cross-section of $10\text{nm} \times 10\text{nm}$ which should contain approximately 1000 atoms and hence should have $M \sim 1000$ (see discussion at end of Chapter 4). Assuming a mean free path of 40nm this suggests that a copper wire any longer than $M\lambda \sim 40\ \mu\text{m}$ should exhibit strange non-Ohmic behavior, for which there is no experimental evidence. Why?

The answer is that localization effects arise from quantum interference and will be observed only if the entire conductor is *phase-coherent*. A copper wire 40 μm long is not phase coherent, certainly not at room temperature. Conceptually we can think of the real conductor as a series of individual coherent conductors, each of length equal to the phase coherence length L_p and whether we see localization effects will depend not on $M\lambda/L$, but on $M\lambda/L_p$.

The bottom line is that to describe real world experiments especially at room temperature it is often important to include a certain degree of dephasing processes as described at the end of the last Lecture. Unless we include an appropriate degree of dephasing our quantum models will show interference effects leading to resonant tunneling or strong localization which under certain conditions may represent real world experiments, but not always. Just because we are using quantum mechanics, the answer is not automatically more "correct."

This can be appreciated by looking at the potential variation along the channel using NEGF and comparing the results to our semiclassical discussion from Lecture 12.

20.3. Potential Drop across Scatterer(s)

In Lecture 12 we discussed the spatial variation of the occupation factor which translates to a variation of the electrochemical potential for low bias. A conductor with one scatterer in it (Fig.20.6), can be viewed (see Fig.12.5) as a normalized interface resistance of one in series with a normalized scatterer resistance of $(1-T)/T$, which can be written as

$$(\text{Normalized}) R_{\text{scatterer}} = \left(\frac{Ua}{\hbar v} \right)^2 \quad (20.7)$$

using Eq.(20.5). The semiclassical potential profile in Fig.20.6 is then obtained by noting that since the current is the same everywhere, each section shows a potential drop proportional to its resistance.

The quantum profile is obtained using an NEGF model to calculate the effective occupation factor throughout the channel by looking at the ratio of the diagonal elements of G^n and A which are the quantum versions of the electron density and density of states respectively:

$$f(j) = \frac{G^n(j,j)}{A(j,j)} \quad (20.8)$$

For low bias, this quantity translates linearly into a local electrochemical potential as noted in Lecture 2 (see Eq.(2.8)). If we choose $f = 0$ at one contact, $f = 1$ at another contact corresponding to qV , then the $f(j)$ obtained from Eq.(20.8) is simply translated into an electrochemical potential μ at that point:

$$\mu(j) = qV f(j) \quad (20.9)$$

The occupation $f(j)$ shows oscillations due to quantum interference making it hard to see the potential drop across the scatterer (see solid black line marked NEGF).

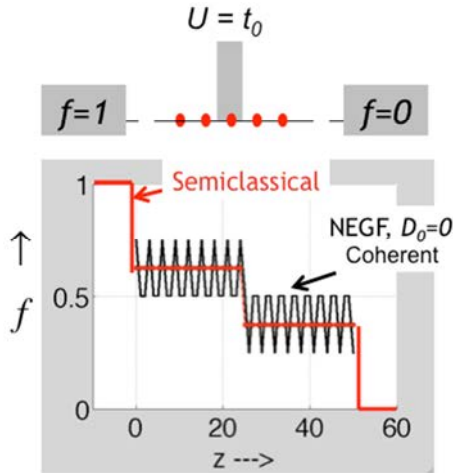


Fig.20.6. Potential drop across a scatterer calculated from the quantum formalism. (a) Physical structure, (b) Coherent NEGF calculation at $E = t_0$.

Experimentalists have measured profiles such as these using scanning probe microscopy (SPM) and typically at room temperature the quantum oscillations are not seen, because of the dephasing processes that are inevitably present at room temperature. This is another example of the need to include dephasing in order to model real world experiments especially at room temperature.

Indeed if we include pure phase relaxation processes (Eq.(19.37)) in the NEGF model we obtain a clean profile looking a lot like what we would expect from a semiclassical picture (see Fig.20.7a).

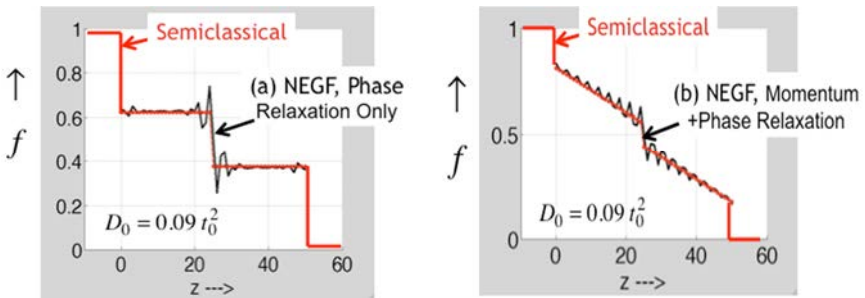


Fig.20.7. Potential drop for the structure in Fig.20.6 calculated from the NEGF method at $E=t_0$ with dephasing, (a) Phase-relaxation only, Eq.(19.37), (b) Phase and momentum relaxation, Eq.(19.38).

Interestingly, if we use a momentum relaxing model for Σ_0 (Eq.(19.38), the potential drops linearly across the structure (see Fig.20.7b), exactly what we would expect for a distributed classical resistor. The resistance per lattice site for this distributed resistor due to D_0 can be obtained by replacing U^2 with D_0 in Eq.(20.7):

$$\text{(Normalized) } R = \underbrace{\left(\frac{a}{\hbar v}\right)^2 D_0}_{\text{Resistance per lattice site}} \quad \frac{L}{a} \quad \text{\# of lattice sites}$$

Another interesting example is that of the two quantum resistors in series that we started with. We noted then that at energies corresponding to points A and B in Fig.20.2 we have constructive and destructive interference respectively. This shows up clearly in the potential profile for coherent transport with $D_0 = 0$ (see Fig.20.8). At $E=0.6t_0$ corresponding to destructive interference, the profile looks like what we might expect for a very large resistor showing a large drop in potential around it along with sharp spikes superposed on it. At $E=0.81t_0$ corresponding to constructive interference, the profile looks like what we expect for a ballistic conductor with all the drop occurring at the two contacts and none across the scatterers.

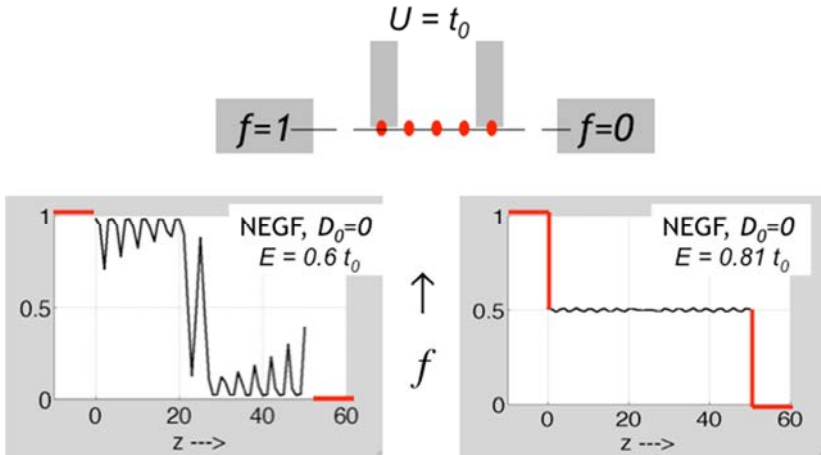


Fig.20.8. Potential drop across two scatterers in series calculated from the NEGF method without dephasing at two energies, $E = 0.81 t_0$ and $E = 0.6 t_0$ corresponding to points marked “A” and “B” respectively in Fig.20.4.

Clearly at $E=0.81t_0$ the answer to the title question of this Lecture is yes, two scatterers can offer less resistance than one. And this strange result is made possible by quantum interference. And once we introduce sufficient phase relaxation into the model using a non-zero D_0 , the profile at both energies look much the same like any semiclassical resistor (Fig.20.9).

Before we move on, let me note that although it is straightforward to include dephasing into toy calculations like this, for large devices described by large matrices, it can be numerically challenging. This is because with coherent NEGF ($D_0 = 0$) or with the momentum relaxing model (Eq.(19.38)), it is often adequate to calculate just the diagonal elements of the Green's functions using efficient algorithms. But for pure phase relaxation (Eq.(19.37), it is necessary to calculate the full Green's function increasing both computational and memory burdens significantly.

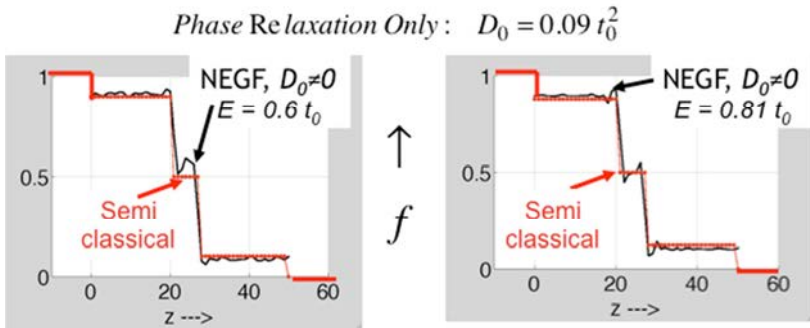


Fig.20.9. Potential drop across two scatterers in series calculated from the NEGF method with pure phase relaxation at two energies, $E = 0.81 t_0$ and $E = 0.6 t_0$ corresponding to points marked “A” and “B” respectively in Fig.20.4.

So a natural question to ask is whether instead of including dephasing through Σ_0 we could include the potential U_R in the Hamiltonian itself and then average our quantity of interest over a number of random realizations of U_R . Would these results be equivalent?

For short conductors like the one shown in Fig.20.4, this seems to be true, but for long conductors like the one in Fig.20.5 this may not be true. With a conductor in the regime of strong localization (Fig.20.5) it is hard to see how averaging the coherent quantum result over many configurations can lead to the semiclassical result.

NEGF with dephasing does not just average over many configurations, it also averages over different sections of the same configuration and that is why it is able to capture the semiclassical result which often describes real world experiments at room temperature quite well.

But could NEGF capture the localization effects observed at low temperatures through a proper choice of Σ_0 ? We believe so, but it would involve going beyond the simple dephasing models (technically known as the self-consistent Born approximation) for evaluating Σ_0 described in Section 19.4.