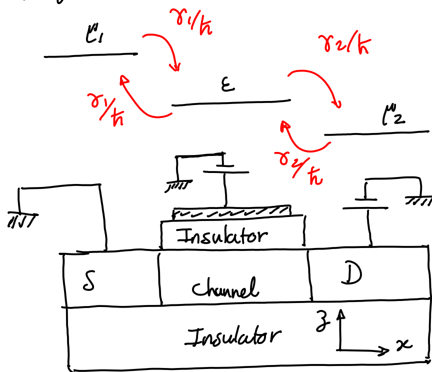
Device Hamiltonian

We are going to derive the Hamiltonian [H] for our FET. This matrix H will replace the single energy ε.



- Where the energy levels in the channel are coming from and how to model them?
- How the coupling to the contacts works?
- How can we calculate [H] for the channel?

We can use our method of band structure calculation. If there are b basis per unit cell and N unit cell in the channel, we will have a bN x bN [H] matrix. For example, Silicon has 2 atoms per unit cell and if we use sp^3s^* model, i.e. 5 orbitals basis functions, [H] will be 10N x 10N, And individual H_{nm} matrices of size 10 x 10.

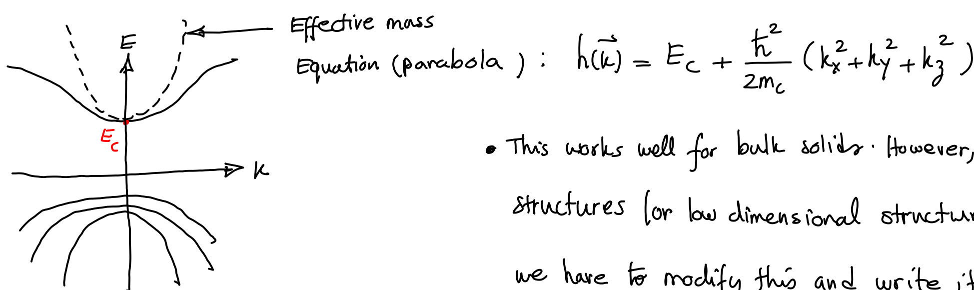
However, we learned that for periodic structures, we can use the ansatz $\{\varphi_m\} = \{\varphi_0\} e^{i\vec{k}\cdot\vec{d}_m}$ and the large atomistic [H] may be solved relatively easily:

$$[h(\vec{k})] = \sum_m [H_{nm}] e^{i\vec{k}\cdot(\vec{d}_m - \vec{d}_n)}$$

The eigenvalues of $[h(\vec{k})]$ are the familiar $E(\vec{k})$ band diagrams.

Alternative to atomistic Hamiltonian: Effective mass Equation

A widely used alternative to the atomistic [H] is the effective mass approximation. This comes from the idea that the important part for conduction and electronic properties is the bottom of the conduction band (or top of the valance band) that can be approximated by a parabola:



Effective mass Equation (parabola): $h(\vec{k}) = E_c + \frac{\hbar^2}{2m_c}(k_x^2 + k_y^2 + k_z^2)$

- This works well for bulk solids. However, for inhomogenous structures (or low dimensional structures) like our FET, we have to modify this and write it in a differential form.

Convert the effective mass equation to a differential equation:

Recall $P_x = \hbar k_x = -i\hbar \frac{\partial}{\partial x} \rightarrow k_x = -i \frac{\partial}{\partial x}$ so we can go from $h(\vec{k}) = E_C + \frac{\hbar^2}{2m_c}(k_x^2 + k_y^2 + k_z^2) \rightarrow$

to:

$$E\psi = \left[E_c - \frac{\hbar^2}{2m_c}\left(\frac{\partial}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) \right] \psi$$

Conceptually this is the equation that electron obeys in a solid. Notice the two parameters E_c and m_c in the equation. Basically all the atomic interactions due to U in the Sch. equation are modeled by the effective mass m_c and the conduction band edge E_c.

Also, notice that any external bias potential U_{ext} can be applied to this equation by shifting E_c:

$$E_c \longrightarrow E_c + U_{ext} \qquad (U_{ext} = -e V_{ext})$$

What have we lost by this HUGE approximation? Answer: our results will only be accurate close to the edge of conduction band. So if the Fermi energy is more inside the band or at high Temperature that electrons can go to higher energy, the parabolic band model fails.

Secondly, we assumed a single effective mass "m_c" which means we assumed an isotropic parabolic approximation. This is not often sufficiend for valance bands.

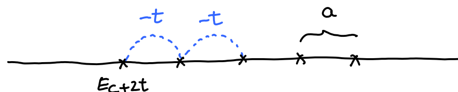
Finite Difference applied to the effective mass equation:

We can use our FD method to solve the effective mass equation:

$$E\psi = \Big(E_c - \frac{\hbar^2}{2m} \underbrace{\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right)}_{\nabla^2} \Big) \psi$$

As an example, in 1D:

$$t = \frac{\hbar^2}{2m_c a^2}$$



$$[H] = \begin{bmatrix} E_c+2t & -t & & \\ -t & E_c+2t & -t & 0 \\ 0 & -t & E_c+2t & \\ & & & \ddots \end{bmatrix}$$

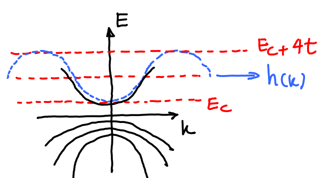
And at presence of an external potential:

$$[H] = \begin{bmatrix} E_c+2t+U_{ext}(x_1) & -t & & \\ -t & E_c+2t+U_{ext}(x_2) & -t & 0 \\ 0 & -t & E_c+2t+U_{ext}(x_3) & \\ & & & \ddots \end{bmatrix}$$

Here "a" need not be atomistic distance and only depeds on the energy range over which accuracy is desired.

We once calculated the bandstructure of this 1D lattice and it was: $E(k) = E_c + 2t(1 - \cos ka)$.

obviously, cosine is not a prabola. Our parabolic band model is acceptable only at small energy range.



* Since $t = \frac{\hbar^2}{2m_c a^2}$ for a given energy range, we can use a lagy 'a' if the effective mass m_c is small.

Self_consistent solution with the external Potential

The effective mass Hamiltonian [H] must be solved self-consistently with the Poisson's equation to find U.

Effective mass Hamiltonian:

$$[H] = \begin{bmatrix} E_c+2t+U(x_1) & -t & & 0 \\ -t & E_c+2t+U(x_2) & -t & \\ 0 & -t & E_c+2t+U(x_3) & \\ & & & \ddots \end{bmatrix}$$

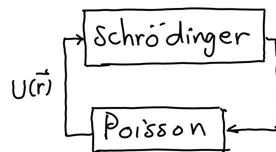
Poisson's equation:

From our Electromagnetic Course: $\nabla^2 V = -\frac{\rho}{\varepsilon}$

In our energy form, $U=-eV$: $-\nabla^2 U = \frac{e^2 n}{\varepsilon}$

And if ε also changes with $\vec{r}$: $-\nabla\cdot(\varepsilon \nabla U) = e^2 n$

Self consistent Solution:



$n(\vec{r}) \rightarrow$ How do we get $n(r)$ from [H]?

Answer: Density matrix $\rho = f_0(H-\mu I)$

and the diagonal elements are n at that point:

$$\rho(r,r') = \sum_\alpha \phi_\alpha(r)\phi^*_\alpha(r') f_0(\varepsilon_\alpha - \mu) \rightarrow n(r) = \sum_\alpha |\phi_\alpha(\vec{r})|^2 f_0(\varepsilon_\alpha - \mu)$$

Electron density in a homogeneous material:

$$\left(E_c - \frac{\hbar^2}{2m_c}\nabla^2\right)\psi_\alpha = E_\alpha \psi_\alpha \rightarrow \psi_k(\vec{r}) = \frac{e^{ik_x x}}{\sqrt{L_x}} \frac{e^{ik_y y}}{\sqrt{L_y}} \frac{e^{ik_z z}}{\sqrt{L_z}}$$

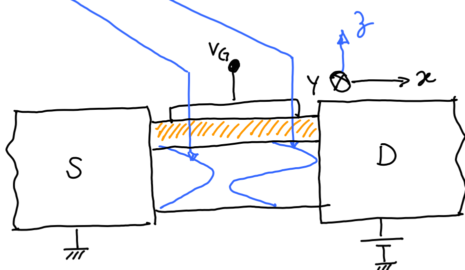
$$\rightarrow n(r) = \sum_{k_x k_y k_z} \frac{1}{L_x L_y L_z} f_0(\varepsilon_k - \mu)$$

Boundary conditions in our FET:

We will assume x and y dimensions are infinite. So we can use Periodic BC for them.

On z direction, the channel is narrow and we use box B.C.:

$\Phi_m(z)$: 1st and 2nd eigenfunctions in z direction



Plain wave as there is no limit on x and y direction

$$\rightarrow \psi_k(\vec{r}) = \frac{e^{ik_x x}}{\sqrt{L_x}} \frac{e^{ik_y y}}{\sqrt{L_y}} \phi_m(z)$$

quantized states due to box B.C. on z direction

If the potent is separable in x, y, and z, Schrodinger equ. can be broken into three coordinates:

In order to find $\Phi_m(z)$, we must solve the z-part of the Schrodinger equation:

$$\left[E_c - \frac{\hbar^2}{2m_c} \frac{d^2}{dz^2} + U(z) \right] \phi_\alpha(z) = E \phi_\alpha(z)$$

In x and y direction, of course, the plain waves satisfy the following:

$$-\frac{\hbar^2}{2m_c}\frac{d^2}{dx^2} \underset{\uparrow \, \Phi(x)}{e^{ik_x x}} = \frac{\hbar^2 k_x^2}{2m_c} e^{ik_x x} \quad \text{and same for } y.$$

So for our FET we have:

$$n(r) = \sum_{k_x k_y m} \frac{1}{L_x L_y} |\Phi_m(z)|^2 f_0(\overbrace{\mathcal{E}_m + \mathcal{E}_{xy}}^{\text{total energy}} - \mu)$$

$$\mathcal{E}_{xy} = \mathcal{E}_x + \mathcal{E}_y = \frac{\hbar^2 k_x^2}{2m} + \frac{\hbar^2 k_y^2}{2m}$$

$$n(\vec{r}) = \sum_m |\Phi_m(z)|^2 \underbrace{\sum_{k_x k_y} \frac{1}{L_x L_y} f_0(\mathcal{E}_m - \mathcal{E}_{xy} - \mu)}_{f_{2D}(\mathcal{E}_m - \mu)}$$

We can sum over k_x and k_y as we will see shortly. Let's call it $f_{2D}(\mathcal{E}_m - \mu)$.

$$n(z) = \sum_m |\Phi_m(z)|^2 f_{2D}(\mathcal{E}_m - \mu)$$

So putting everything together, we have to solve the following set of equations self-consistently:

- Schrödinger

$$\left[E_c - \frac{\hbar^2}{2m} \frac{d^2}{dz^2} + U(z) \right] \Phi_m(z) = \mathcal{E}_m \Phi_m(z)$$

↓ $\Phi_m(z)$

$$n(z) = \sum_m |\Phi_m(z)|^2 f_{2D}(\mathcal{E}_m - \mu)$$

↓ $n(z)$

- Poisson

$$-\frac{d}{dz}\left(\varepsilon \frac{d}{dz} U\right) = e^2 n(z)$$

$U(z)$ (fed back to Schrödinger)

Let's calculate f_{2D} now:

$$f_{2D} = \sum_{k_x k_y} \frac{1}{L_x L_y} f(\mathcal{E}_m + \mathcal{E}_{xy} - \mu) = \frac{1}{L_x L_y} \sum_{k_x k_y} \frac{1}{1 + \exp\left(\frac{\mathcal{E}_m - \mu + \mathcal{E}_{xy}}{k_B T}\right)} \quad \text{where } \mathcal{E}_{xy} = \frac{\hbar^2}{2m}(k_x^2 + k_y^2)$$

$$\exp\left(\frac{\mathcal{E}_m - \mu + \mathcal{E}_{xy}}{k_B T}\right) = \exp\left(\frac{\mathcal{E}_m - \mu}{k_B T}\right) \exp\left(\frac{\mathcal{E}_{xy}}{k_B T}\right)$$

We know that k_x (also k_y) values are separated by $\frac{2\pi}{L_x}$ (or $\frac{2\pi}{L_y}$ for k_y). So for large devices, the k_x values are very close together and we can convert the sum to integral:

$$\frac{1}{L_xL_y}\sum_{k_xk_y}\frac{1}{1+\underbrace{e^{\frac{\varepsilon_m-\mu}{k_BT}}}_{A}\,e^{\frac{\varepsilon_{xy}}{k_BT}}} = \frac{1}{L_xL_y}\int\frac{dk_x\,dk_y}{\frac{2\pi}{L_x}\frac{2\pi}{L_y}}\frac{1}{1+Ae^{X}}$$

($\frac{\varepsilon_{xy}}{k_BT} \to X$)

$$= \frac{1}{4\pi^2}\int\frac{dk_x\,dk_y}{1+Ae^X}$$

($dk_x\,dk_y = k_{xy}\,d\theta\,dk_{xy}$ in polar coordinate) where $k_{xy}=\sqrt{k_x^2+k_y^2}$

$$= \frac{1}{4\pi^2}\underbrace{\int_0^{2\pi}d\theta}_{2\pi}\int_0^\infty dk_{xy}\frac{k_{xy}}{1+Ae^X} = \frac{1}{2\pi}\int_0^\infty\frac{k_{xy}\,dk_{xy}}{1+Ae^X}$$

$$X=\frac{\hbar^2}{2mk_BT}\overbrace{(k_x^2+k_y^2)}^{k_{xy}^2} \to dX=\frac{\hbar^2}{mk_BT}k_{xy}\,dk_{xy}$$

$$= \frac{1}{2\pi}\frac{mk_BT}{\hbar^2}\int_0^\infty\frac{dx}{1+Ae^x} = \frac{mk_BT}{2\pi\hbar^2}\int_0^\infty\frac{\overbrace{e^{-x}dx}^{-du}}{\underbrace{A+e^{-x}}_{U}} = \frac{mk_BT}{2\pi\hbar^2}\ln U\Big|_A^{A+1} = \frac{mk_BT}{2\pi\hbar^2}\ln\left(1+\frac{1}{A}\right)$$

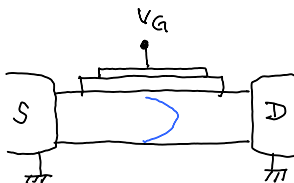
$$= \frac{mk_BT}{2\pi\hbar^2}\ln\left(1+e^{-\frac{\varepsilon_m-\mu}{k_BT}}\right)$$

$$\to f_{2D}(E)=\frac{mk_BT}{2\pi\hbar^2}\ln\left(1+e^{-E/k_BT}\right)$$ where $E=\varepsilon_m-\mu$ for us.

Quantum versus Semiclassical

Quantum

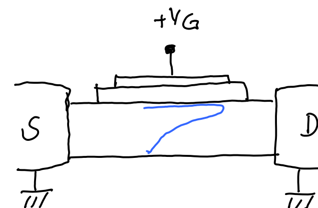
$\Phi(z)$ goes to zero at the boundaries, Therefore electron density is zero at the boundaries:



For big devices the QM solution results in various fourier components which add up together and result in a distribution like that of semiclassics.

Semiclassical

There is no concept of wavefunction here. Electron density has a pick at the boundary as the positive gate voltage pulls up the Electrons to the surface.



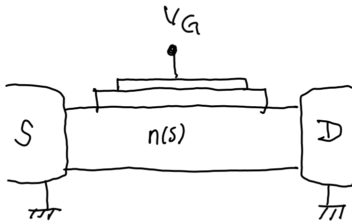
Electron Density as a function of gate voltage

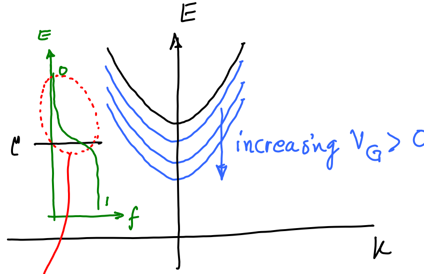
We learned how to calculate $n(z)$. We want to calculate $n(s)$ now which is the electron density per unit area:

$$n(z) = \sum_m |\phi_m(z)|^2 f_{2D}(\varepsilon_m - \mu)$$

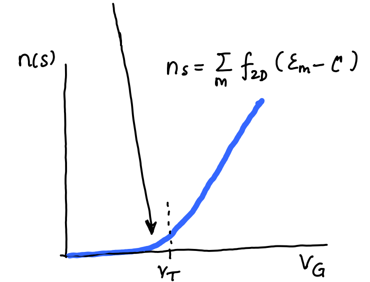
$$\rightarrow \quad n(s) = \int n(z)\, dz = \sum_m f_{2D}(\varepsilon_m - \mu) \underbrace{\int |\phi_m(z)|^2 dz}_{=1} = \sum_m f_{2D}(\varepsilon_m - \mu)$$

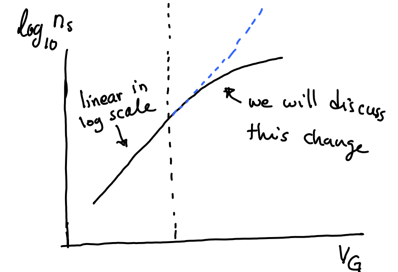
Our motivation for doing this is that if we think the device as a capacitor, $n(s)$ is the charge density per unit area and we have a capacitance per unit area.





n_s and hence current starts to increase as the bottom of conduction band passes μ





In this region $f(E) \simeq e^{-E/k_BT}$

So the slope of $n(s)$ will also be as e^{-E/k_BT}.

Or in log scale as linear:

As the potential U is changed inside the channel, there will be more electrons filling the bottom of conduction band:

$$n_s \propto e^{\frac{-E_c + U}{k_BT}} \qquad \text{assume } U = -eV_G \rightarrow$$

or say $\quad n_s = A\, e^{\frac{-E_c + eV_G}{k_BT}}$; A is some constant

$$\rightarrow \quad \ln n_s = \underbrace{\ln A - \frac{E_c}{k_BT}}_{\equiv C} + \frac{eV_G}{k_BT}$$

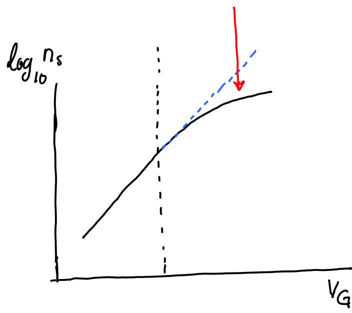
At room temperature: $\ln n_s = C + \dfrac{eV_G}{26 \text{ mev}}$ or $\log_{10} n_s = C' + \dfrac{eV_G}{2.3 \times 26 \text{ mev}}$

$$\rightarrow \quad \log_{10} n_s \equiv C' + \frac{eV_G}{60 \text{ mev}}$$

So electron density changes at rate of one decade per 60 mev.

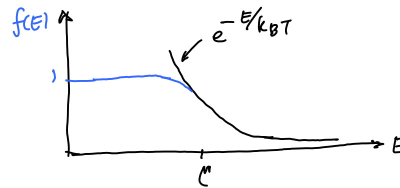
Is there anyway to make the device faster? We like to use smaller gate voltage to increase the electron density (or to turn on the transistor). The answer is as long as our FET operates on the basis of electrostatic principles, we're stuck with 60 mev per decade!

Why does the slope change here?



There are two reasons:

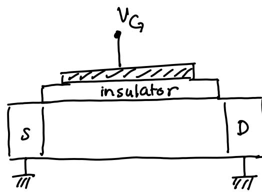
1) Fermi function does not increase as e^{-E/k_BT} in the whole energy range. It saturates to one as $E \ll \mu$.



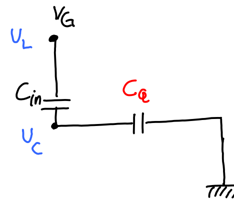
2) We assumed $U = -eV_G$. But as the channel gets more conductive, the potential in the channel becomes smaller and at metallic limit U becomes zero (note: potential in the metal is always zero)

Quantum Capacitance

We can model the variation of the channel potential V_c versus the Gate voltage V_G by associating a quantum capacitance C_Q to the electron density in the channel:



$\equiv$



$C_{in} = \frac{\varepsilon}{d_{in}}$ F/unit area

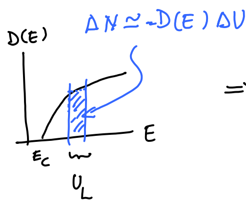
d_{in} ↳ insulator thickness

From our model:

$$U_C = U_L \frac{C_{in}}{\underbrace{C_{in}+C_Q}_{C_E}} = U_L \frac{C_E - C_Q}{C_E} = U_L - \frac{C_Q}{C_E} U_L$$

From the physics:

we can estimate ΔN by linearizing $D(E)$ at small change in energy →



$$\Rightarrow U_C = U_L + \frac{e^2}{C_E}\Delta N = U_L - \frac{e^2}{C_E} D U_L$$

$$\boxed{C_Q = e^2 D}$$

Let's consider two limitting case of:

(1) Low DOS in channel: C_Q is small so the voltage accross C_Q (i.e. V_c) is large (capacitor acts as inverse resistor).

(2) High DOS in channel: C_Q is large so V_c is small. So C_Q has small effect. This is the case for large FET where DOS is large in the channel.

In the past, FET transistors where large and C_Q was large. So its effect was not significant. However, in nano-FETs C_Q becomes very small ($C_Q < C_{in}$), hence it has a significant effect on the behavior of the transistor.

Example: Quantum Capacitance in 2D

$C_Q = e^2 D$

In 2D: $D = \frac{m_c}{\pi \hbar^2} \rightarrow C_Q = \frac{e^2 m_c}{\pi \hbar^2}$ This can be smaller than C_{in} if $d_{in} \simeq 10\,\text{Å}$ $\left(C_{in} = \frac{\varepsilon}{d_{in}}\right)$.

CHAPTER 8:

ENERGY LEVEL BROADENING

Coupling of the gate channel to the contacts results in what is known as level broadening.

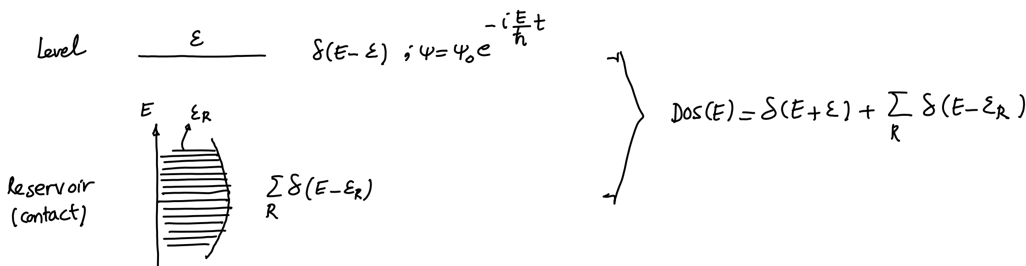
For a closed or in equilibrium system, we can write down the Hamiltonian [H] and from that find the density matrix and calculate the electron density, which are the diagonal elements of the density matrix in real space representation: $[\rho] = f_0([H] - \mu[I])$

E, contact, ε level, E, D(E), $\gamma = \frac{\hbar}{\tau}$

But in open systems, for example when the channel is in contact with very large contacts relations are not that simple.

In an isolated system (closed system), the electron and the electron wavefunction will stay in the level for ever.

The DOS for a isolated single energy level and an isolated reservoir (contact) can be written as:



When they are in touch, the electron escapes from the level to the reservoir. Or we can say the electron wavefunction has a decay time (or life time). In formulation this is a factor of $e^{-t/\tau}$ to the wavefunction:

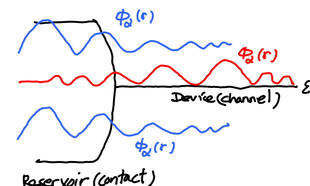
$$\psi_0 e^{-i\frac{E}{\hbar}t} \longrightarrow \psi_0 e^{-i\frac{E}{\hbar}t} e^{-\frac{t}{\tau}}$$

If we look at DOS as Fourier transform of the wavefunction, DOS is a delta functio $\delta(E-\varepsilon)$ when isolated (no decay). When connected to a reservoir DOS looks like a broadened as the electron wavefunction has a decay time:

Stronger Coupling $\longrightarrow$ shorter lifetime $\longrightarrow$ Bigger Broadening

Local Density of states & Spectral Function

When the channel is in touch with the contact, the contact wavefunctions do not stay only in the contact and "leak" into the channel. Similarly, the channel (or device) wavefunction will "leak" slightly into the contact.



The "local density of states" can be written as:

$$D(\vec{r}, E) = \sum_\alpha |\phi_\alpha(r)|^2 \delta(E-\varepsilon_\alpha)$$

We can look at this more generally as the diagonal elements of **the spectral function A** (deviden by 2π; this is just a convention):

$$A(\vec{r},\vec{r}', E) = 2\pi \sum_\alpha \phi_\alpha(\vec{r})\delta(E-\varepsilon_\alpha)\phi_\alpha^*(\vec{r}')$$

We can compare $D(E)$ to $A(r,r',E)$ with the electron density $n(\vec{r})$ to the density function $\rho(r,r')$:

Density Matrix

$$n(r) = \sum |\phi_\alpha(r)|^2 f(\varepsilon_\alpha - \mu) = \rho(r,r)$$

$$\rho(r,r') = \sum_\alpha \phi_\alpha(\vec{r}) f(\varepsilon_\alpha - \mu)\phi_\alpha^*(r')$$

$$[\rho] = f([H]-\mu[I])$$

Spectral Function

$$D(\vec{r},E) = \sum |\phi_\alpha(r)|^2 \delta(E-\varepsilon_\alpha) = A(r,r,E)/2\pi$$

$$A(r,r',E) = 2\pi \sum_\alpha \phi_\alpha(r)\delta(E-\varepsilon_\alpha)\phi_\alpha^*(r')$$

$$[A] = 2\pi\,\delta(E[I]-[H])$$

Now, let's confirm that even after broadening, the total number of electrons a level holds is 1 (conservation).
We must show $\int_{-\infty}^{\infty} dE\, D(\vec{r},E) = 1$. Using the spectral function $\int_{-\infty}^{\infty} dE\, D(\vec{r},E)$ are the diagonal elements of $\int_{-\infty}^{\infty} \frac{A}{2\pi} dE$. For calculating this, A in real space can be taken in eigenstate space using V and V^+ as before:

$$A(r,r',E) = 2\pi \begin{bmatrix} A(r_1,r_1,E) & A(r_2,r_1,E) & \cdots & A(r_N,r_1,E) \\ A(r_1,r_2,E) & A(r_2,r_2,E) & & \\ \vdots & & \ddots & \\ A(r_1,r_N,E) & & & A(r_N,r_N,E) \end{bmatrix} = 2\pi V \underbrace{\begin{bmatrix} \delta(E-\varepsilon_1) & & 0 \\ & \delta(E-\varepsilon_2) & \\ & & \ddots \\ 0 & & \delta(E-\varepsilon_N) \end{bmatrix}}_{A \text{ in eigenstate space}} V^+$$

$$\int_{-\infty}^{\infty} \frac{A}{2\pi} dE = 2\pi V \begin{bmatrix} \int \delta(E-\varepsilon_1)dE & & 0 \\ & \ddots & \\ 0 & & \int \delta(E-\varepsilon_N)dE \end{bmatrix} V^+ = VIV^+ = VV^+ = I$$

Density Matrix in terms of Spectral Function

spectral function can be also used to rewrite the density matrix in a new form:

$$[\rho] = f_0([H]-\mu[I])$$

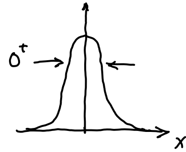
$$= \int dE\, f_0(E-\mu)\,\delta(E[I]-[H])$$

$$= \int \frac{dE}{2\pi} f_0(E-\mu)[A(E)] \quad \rightarrow \quad [\rho] = \int \frac{dE}{2\pi} [A(E)]\, f_0(E-\mu)$$

For a closed system the old expression $f_0([H]-\mu[I])$ is easier. But for open systems like our nano-FET we have to also consider the coupling effect. For this case we can use the spectral function for looking at the DOS inside the channel (device) and consider the effect of coupling to the contacts (reservoir).

Green's function introduction

We can write the delta function in the form of an infinite height and area of 1:



$$2\pi\delta(x) = \frac{2\,0^+}{x^2 + 0^{+2}} = i\left(\frac{1}{x + i0^+} - \frac{1}{x - i0^+}\right)$$

Since $A = 2\pi\delta(EI - H)$, we can write it as: $[A(E)] = i\Big[\underbrace{\left((E + i0^+)I - H\right)^{-1}}_{G} - \underbrace{\left((E - i0^+)I - H\right)^{-1}}_{G^+}\Big]$

where we define the Green's function G as:

$$G = \left[(E + i0^+)I - H\right]^{-1}$$

Such that:

$$[A(E)] = i\,[G - G^+]$$

When we also consider the contact (reservoir), the Green's function for the channel (device) becomes:

$$G = \left[EI - H - \Sigma\right]^{-1}$$

Σ is the self energy that gives the effect of coupling to the contacts. As we will see, its imaginary part gives us the lifetime of the electrons in the channel.

H and Σ here are of the size of the channel and not that of reservoir. So to describe the channel, we don't have to inverse a huge matrix H for the whole system of channel and the contacts. We just take the Green's function describing the channel and work with that.