Charge Transport in Semiconductors

EE 698D, Advanced Semiconductor Physics
Debdeep Jena (djena@nd.edu)
Department of Electrical Engineering
University of Notre Dame
(Fall 2004)

Contents

1 Introduction 2
1.1 Basic semiconductor physics 2
1.2 Example illustrating a Matrix element evaluation 4

2 Transport of Bloch electrons in perfect semiconductors 5

3 Effective Mass Approximation, Envelope Functions 7

4 Transport theory - Real Semiconductors 10
4.1 Setting up the Quantum Mechanical Problem 11
4.2 Fermi's Golden Rule 12
4.3 Example illustrating application of Fermi's Golden Rule 17

5 Case study of scattering in semiconductors 19
5.1 Example - Ionized impurity scattering in bulk semiconductors and 2DEGs . 20
5.2 Other scattering mechanisms and where to read about them 21
5.3 High field transport 21
5.4 Transport Regimes 23

6 Formal Transport theory 24
6.1 Boltzmann transport equation 24
6.1.1 Electric field 26
6.2 Mobility- basic theory 28
6.3 Statistics for two- and three-dimensional carriers 30
6.4 Screening: Semiclassical Theory 31
6.5 Screening by 2D/3D Carriers: Formal Theory 35
6.6 Mobility of two- and three-dimensional carriers 36
6.6.1 Two-dimensional carriers 36
6.6.2 Three-dimensional carriers 39
6.7 Material properties of III-V nitrides relevant to transport 42

7 Current research, future directions 44

1 Introduction

1.1 Basic semiconductor physics

This is a brief recap of semiconductor physics that we will need for the transport theory that will follow. As we all know, a semiconductor has atoms arranged in a crystal lattice, with all atoms tetrahedrally bonded. Such a solid will allow only certain bands of allowed energies, separated by forbidden gaps. The gap separating the highest filled band (valence band) from the lowest unoccupied band (conduction band) is the well known bandgap of the semiconductor.

Semiconductors have to be doped to get sufficient carriers to do useful electronics. These dopants are typically substitutional impurity atoms that have energy levels very close to the conduction band edge (donors) or the valence band edge (acceptors). The electrons (in case of donors) and holes (acceptors) contributed by the dopants are bound to the parent atoms at very low temperatures. However, at room temperature, they have sufficient energy to escape the binding energy of the parent dopant and become free to wander around the whole crystal. How fast do these free mobile charges move? That is the question transport theory aims to answer. It is of fundamental importance, since it determines how fast your transistor can go, how well your laser will lase, and so on.

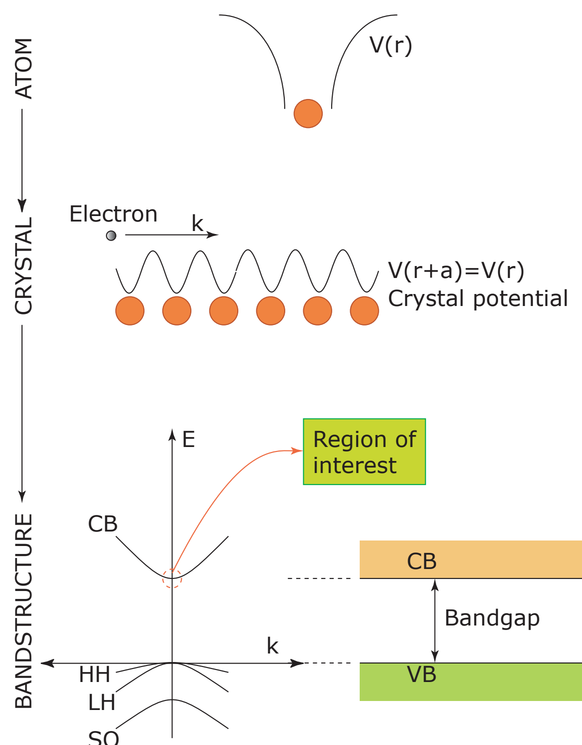

Figure 1: Evolution of bands and gaps in semiconductors

Electrons (or holes) that are free to wander around the crystal experience the Coulombic potential variation of the atoms in the crystal (see Figure 1). In a perfect crystal, this

potential variation is periodic in nature. How is this problem solved? Bloch's theorem provides the vital link to the solution of the problem. It turns out that in a perfect crystal, you can remove the whole periodic crystal potential and lump it into the wonderful concept of the 'effective mass' of the electron. By doing so, you have freed the electron from the shackles of widely fluctuating Coulombic potentials and have reduced the problem to that of an electron in free space, albeit with a different mass [1]. Solving the problem for the electron taking the periodicity of the crystal into account yields the bandstructure of the semiconductor, which stores a wealth of information. The Brillouin Zone, the bandstructure, and the simplified bandstructure around the CB minima and VB maxima for GaN, which is the region of interest for most problems is shown in Figure 2. It is important to note that the bandstructure is nothing but the allowed eigenvalues of the quantum mechanical problem for the perfect semiconductor. The k-values are quantized; the quantization is so fine that for all practical purposes, it is assumed continuous.

We now direct our attention towards the problem at hand - how to understand transport theory. So the model we start with is really the first problem of an undergraduate quantum mechanics class - you have a particle (electron, effective mass m^*) in a macroscopic crystal (say a GaN crystal). Let us assume the shape of the crystal to be a cube of side L [2]. Since the particle is confined in a box, we model it as an infinite quantum well in 3 dimensions. It is the good old particle in a box problem, which has well known solutions. The wavefunction of the electron will be

$$\psi(x,y,z) = \frac{1}{\sqrt{L^3}} e^{i(k_x x + k_y y + k_z z)} = \frac{1}{\sqrt{L^3}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{1}$$

which when subjected to the boundary conditions [3] yields the limits on the wavevectors -

$$\mathbf{k} = (k_x, k_y, k_z) = \frac{2\pi}{L}(n_x, n_y, n_z) \tag{2}$$

where n_x, n_y, n_z are allowed only integral values (...,-1,0,1,...). The energy is $E(k) = \frac{\hbar^2|\mathbf{k}|^2}{2m^*}$ and the momentum is $\mathbf{p} = \hbar\mathbf{k}$. The wavevector $\mathbf{k}$ for a free electron (not in a box) is related to its wavelength by the de Broglie relation $\lambda = \frac{h}{p}$; it is easy to see that it translates to $k = \frac{2\pi}{\lambda}$. When the electron is confined in a box, it can have only those wavelengths that can fit in the box; so, the wavelengths are 'quantized', according to $\mathbf{k}$. Electrons with wavevectors $k \to 0$ have very small momentum and are delocalised over the whole crystal, since their $\lambda \to \infty$. On the other hand, electrons with large wavevectors have really small wavelengths, and can be fitted into small microscopic areas. We will see how this is crucial in the phenomenon of screening, a very important phenomenon in transport and indeed in all of semiconductor physics.

[1] This is the reason why you see that effective mass is not a scalar, but has dependence on the wavevector k. Roughly, since all directions are not the same in a crystal, neither should the periodic potential (and hence the effective mass in that direction). For our purpose though, we will be happy to deal with a direct gap semiconductor that has an isotropic effective mass at the Γ point of the conduction band.

[2] L can be 1 cm; it is a macroscopic length, not microscopic!

[3] The periodic boundary condition requires $\psi(x+L,y,z) = \psi(x,y+L,z) = \psi(x,y,z+L) = \psi(x,y,z)$

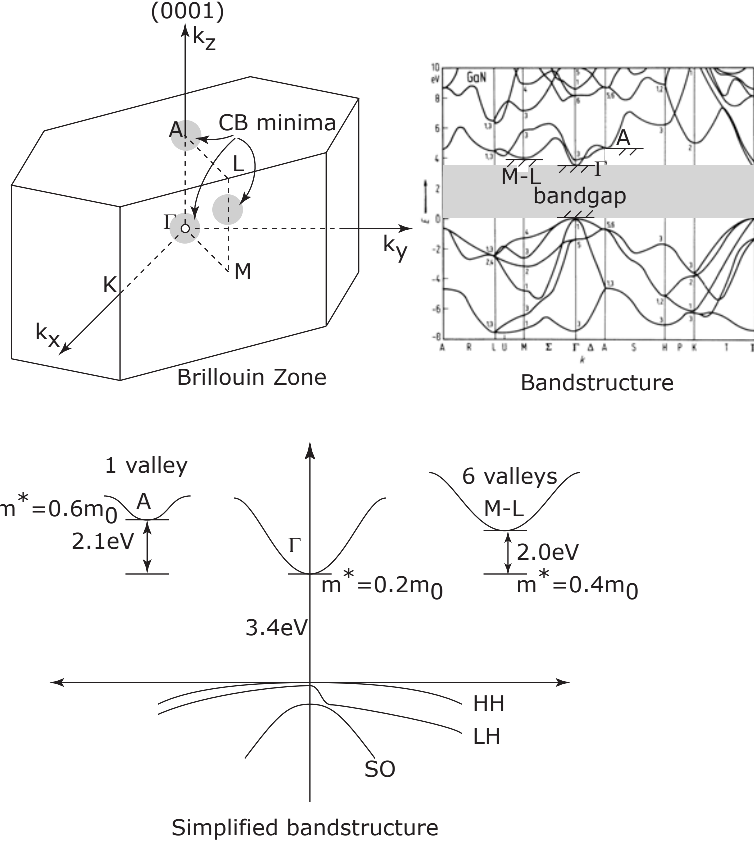

Figure 2: The Brillouin Zone of Wurtzite GaN crystal structure, The calculated bandstructure for different directions in the Brillouin Zone, and the Kane model of Bandstructure around the Γ point.

1.2 Example illustrating a Matrix element evaluation

Evaluate $\langle \mathbf{k}'|\mathbf{k}\rangle$ for the free electron in the box of volume $V = L_x \cdot L_y \cdot L_z$.

$$|\mathbf{k}\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{3}$$

$$\langle \mathbf{k}'|\mathbf{k}\rangle = \frac{1}{V}\int d^3r e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} = \frac{1}{V}\int_0^{L_x} dx e^{i(k_x-k'_x)x} \int_0^{L_y} dx e^{i(k_y-k'_y)y} \int_0^{L_z} dz e^{i(k_z-k'_z)z} \tag{4}$$

$$\langle \mathbf{k}'|\mathbf{k}\rangle = (\frac{e^{i(k_x-k'_x)L_x}-1}{i(k_x-k'_x)L_x}) \cdot (\frac{e^{i(k_y-k'_y)L_y}-1}{i(k_y-k'_y)L_y}) \cdot (\frac{e^{i(k_z-k'_z)L_z}-1}{i(k_z-k'_z)L_z}) \tag{5}$$

Now we use a very useful result that k_x, k'_x are quantized by the wave-fitting procedure as in Equation 2. So $(k_x - k'_x)L_x = 2\pi n$ where n is an integer. The same holds for k_y, k_z as

well. Using this, we see that unless $k_x = k'_x, k_y = k'_y, k_z = k'_z$ all hold true, the value of the integral would be zero. If they do hold, the value is ONE (prove it!). Thus, the result is

$$\langle \mathbf{k}' | \mathbf{k} \rangle = \delta_{k_x, k'_x} \delta_{k_y, k'_y} \delta_{k_z, k'_z} = \delta_{\mathbf{k}, \mathbf{k}'} \tag{6}$$

2 Transport of Bloch electrons in perfect semiconductors

Consider a semiconductor crystal that is perfect - i.e., there are no defects or impurity doping of any kind. Let us follow the motion of an electron in the valence band in the presence of a force (due to electric field), pointing to the right ($+k$) (see Figure 3). A force imparts momentum to the electron, moving it into the next nearest state on the right, the electron in the next nearest state to its nearest state on right, and so on, according to the 'Newton's law' in **k**-space,

$$\mathbf{F} = \hbar \frac{d\mathbf{k}}{dt}. \tag{7}$$

The force may be due to an externally applied electric field ($\mathbf{E}$), magnetic field $\mathbf{B}$, or a combination of both. In general, one writes the equation of motion as

$$\mathbf{F} = (-e) \cdot [\mathbf{E} + \mathbf{v} \times \mathbf{B}] = \hbar \frac{d\mathbf{k}}{dt}. \tag{8}$$

The electron at the end of the Brillouin Zone (BZ) in the VB "flips over", and re-enters the 1st BZ from the left, by acquiring a reciprocal lattice vector $\mathbf{G}$ from the lattice; such a process is called an Umklapp[4] process. Umklapp processes occur so long as the field is not too high to cause the transition of an electron from one band to the other. If the force is too strong, instead of Umklapp processes, an electron can 'tunnel' through the forbidden gap to the next nearest empty band - this is the famous 'Zener' tunneling process, which is used for commercially available Zener diodes.

The role of the force (through the electric field) is to impart the electron momentum; the velocity and acceleration of the electrons are determined by the bandstructure. We have seen already that the velocity is given by

$$\mathbf{v} = \frac{1}{\hbar} \frac{\partial E(k)}{\partial k} = \frac{1}{\hbar} \nabla_k E(k), \tag{9}$$

and the acceleration is given by

$$\mathbf{a} = \frac{\mathbf{F}}{m^\star}. \tag{10}$$

Keeping these facts in mind, we plot the bandstructure, the effective mass $m^\star$, the velocity, and the acceleration of the electron in the valence band in Figure 3.

Many important features can be inferred from this figure. First, if one has a valence band filled with electrons, there can be NO NET CURRENT in a perfect semiconductor, even in

[4]The German verb *umklappen* means to flip over.

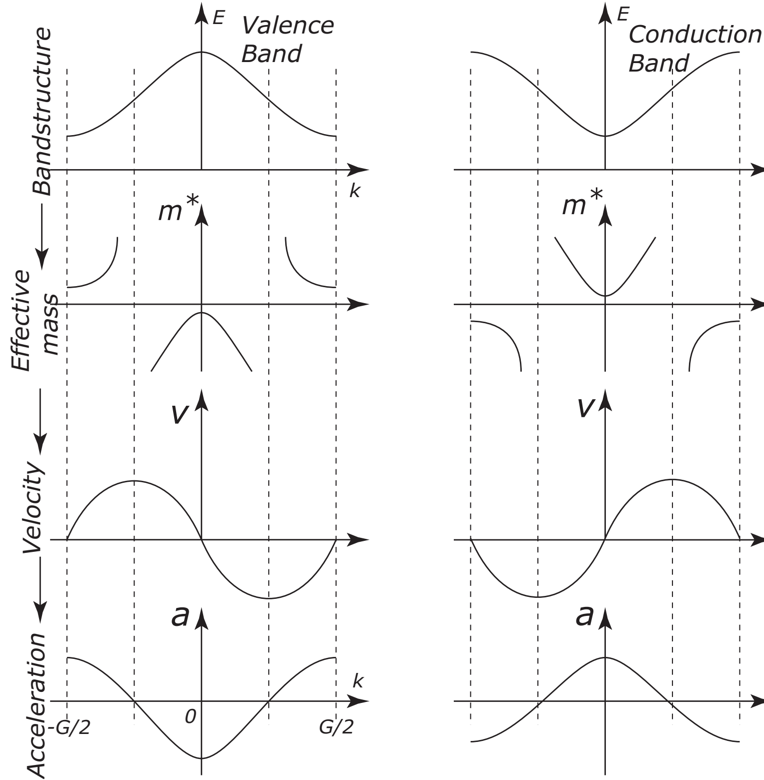

Figure 3: Bandstructure, effective mass, velocity, and acceleration of Bloch electrons in a valence and a conduction band.

the presence of an electric field. This happens since the electrons move within the bands in endless cycles, and a sum of the velocities due to all electrons is zero, since for every electron with a positive velocity, there is one with an equivalent negative velocity.

Strange things happen if one considers only ONE electron moving in a empty conduction band in a perfect semiconductor crystal. From Figure 3, it is clear that similar to the motion of an electron in the filled valence band, the electron in the CB will undergo oscillations in k-space, as well as real space. This is because after each cycle of period $T = \hbar G/F$, the electron returns to where it started from in both k-space (obvious), and in real space as well (since area under velocity -time curve is zero)! Such oscillations are called Bloch oscillations, and they remain a source of great interest to this day, as a possible source of microwave radiation.

However, we know very well that in *real* semiconductors, when one applies a force on carriers (via electric field), there is a current. The reason is the presence of *scattering*. In real semiconductors, there will always be impurities present, which will scatter carriers, and Bloch oscillations are stopped far before carriers can reach the edge of the Brillouin zones. Carrier scattering forces electrons back to the bottom of the CB or holes back to top of the VB; thus, we can conclude that characteristic scattering times (such as those that enter in the expression for mobility, $\mu = e\tau/m^{\star}$) are far smaller that the Bloch oscillation period.

This leads us naturally to study transport in real semiconductors, which is dominated by scattering off impurities, defects, and phonons, all of which are *perturbations* from the perfect crystal.

3 Effective Mass Approximation, Envelope Functions

Before we jump into considering real semiconductors with impurities and corresponding perturbations from perfect periodic potentials, it is worthwhile to develop a very powerful formalism that greatly simplifies our treatment of transport properties. So long as the perturbations of the crystal potential is not drastic, one can re-cast the Schrödinger equation in a form that is very useful for discussing transport and device applications. One runs into a fundamental problem in dealing with a particle location in real space and its momentum at the same time. To do that, the concept of a wave packet is necessary. Wave packets, unlike pure Bloch-eigenstates, have a finite spread both in the momentum and real space. A wave packet is nothing but a linear combination of Bloch eigenstates for small $k-$values around a region of interest in the Brillouin zone. For most cases, it suffices to investigate properties of electrons and holes located very close to the band extrema in the $k-$space; therefore, one collects Bloch eigenstates around such points, and creates a wavepacket by taking their linear combinations.

To illustrate this, let us consider the 1-dimensional case. We construct a wavepacket by taking a linear combination of Bloch eigenstates $\phi_{nk}(x)$ from the nth band with wavevector k. The sum is over the whole BZ.

$$\psi(x) = \sum_n \sum_k C(k)\phi_{nk}(x) = \sum_n \int \frac{dk}{2\pi} C(k)\phi_{nk}(x) \tag{11}$$

We now make two crucial approximation -

a) We assume that wavefunctions from only one band play a part in the wavepacket, and thus drop the sum over all bands.

b) We assume that in the single band we are interested in, wavevectors from a small region say around $k_0 = 0$ are important (see Figure 4).

Then, Bloch functions can be written as $\phi_{nk}(x) = e^{ikx}u_{nk}(x) \approx u_{n0}(x)e^{ikx} = \phi_{n0}(x)e^{ikx}$. Then the wavepacket takes the form

$$\psi(x) \approx \phi_{n0}(x) \int \frac{dk}{2\pi} C(k)e^{ikx} = \underbrace{\phi_{n0}}_{Bloch} \cdot \underbrace{C(x)}_{envelope}, \tag{12}$$

where the integral term is identified as the Fourier transform of the weights $C(k) \leftrightarrow C(x)$. The real-space function $C(x)$ which is a Fourier transform of the weights of the wavepacket is called as the **envelope** function; since the weights $C(k)$ are over a small region in $k-$space, $C(x)$ is spread over real space. It is typically a smooth function spreading over several lattice constants. This is illustrated in Figure 5.

How does the wavepacket behave when we apply the periodic crystal Hamiltonian H_0 on it? Since $\phi_{nk}(x)$ are Bloch-eigenfunctions of this Hamiltonian, $H_0\phi_{nk}(x) = E_n(k)\phi_{nk}(x)$,

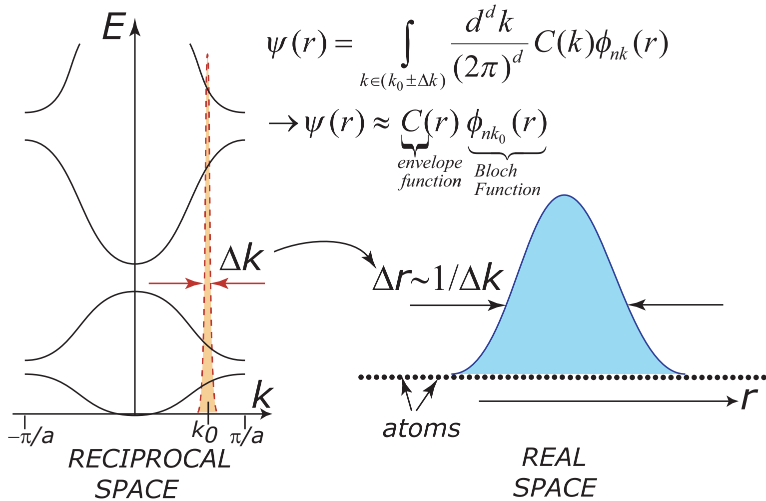

Figure 4: A wavepacket is constructed by taking Bloch functions from a small region of the reciprocal space, and summing them with weights. The weights $C(k)$ have a small extent Δk in reciprocal space; when carried over to real space, the spread is large, since $\Delta r \sim 1/\Delta k$; thus the wavepacket has a finite spread in real space, and represents the wavefunction of a particle. If we restrict the sum in reciprocal space to 1% of the BZ, the wavepacket spreads over $1/0.01 = 100$ atoms in real space. The real space wavefunction is given by the Bloch wavefunction at the k_0 point, modulated by an envelope function $C(r)$, which is the Fourier transform of the weights $C(k)$.

and we recover

$$H_0\psi(x) = \int \frac{dk}{2\pi} C(k)E_n(k)\phi_{nk}(x) \approx \phi_{n0}(x) \int \frac{dk}{2\pi} C(k)E_n(k)e^{ikx}. \tag{13}$$

We now write out the energy eigenvalues as a Taylor-series of small wavevectors around $k = k_0 = 0$,

$$E_n(k) = \sum_m a^m k^m \tag{14}$$

and Schrödinger equation becomes

$$H\psi(x) \approx \phi_{n0}(x) \sum_m \int \frac{dk}{2\pi} C(k)k^m e^{ikx}. \tag{15}$$

We now use a property of Fourier transforms - if $f(k) \leftrightarrow f(k)$, then $kf(k) \leftrightarrow (-id/dx)f(x)$, and in general, $k^m f(k) \leftrightarrow (-id/dx)^m f(x)$. Thus,

$$\int \frac{dk}{2\pi} k^m C(k)e^{ikx} \leftrightarrow (-i\frac{d}{dx})^m C(x), \tag{16}$$

and the Schrödinger equation is recast as

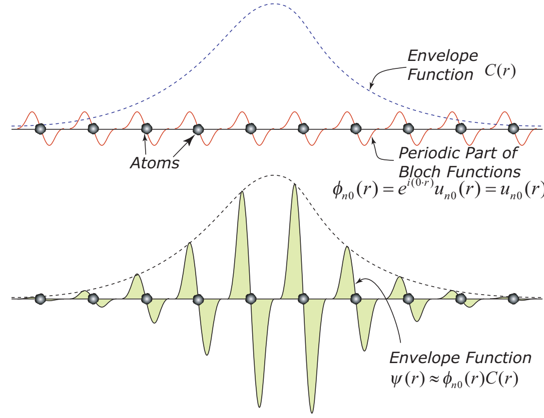

Figure 5: Envelope function $C(r)$ modulates the Bloch function $\phi_{n0}(x)$ to produce the wavefunction of the wavepacket $\psi(x)$

$$H\psi(x) \approx \phi_{n0}(x)E_n(-i\nabla)C(x), \tag{17}$$

which can be generalized to the 3-D case. Thus, in the energy term, we make the substitution $k \to i\partial/\partial r$, making it an operator that acts on the envelope function only. This step is crucial - the Bloch function part has been pulled out as a coefficient; no operators act on it.

Now, instead of the periodic potential Hamiltonian, if we have another potential (say a perturbation) $V(r)$ present, Schrödinger equation becomes

$$H_0\phi_{n0}(r)C(r) + V(r)\phi_{n0}(r)C(r) = E\phi_{n0}(r)C(r), \tag{18}$$

and using Equation 17, it becomes

$$[E_n(-i\nabla) + V(r)]C(r) = EC(r), \tag{19}$$

where the Bloch functions do not appear at all! Furthermore, if we already know the bandstructure of the semiconductor, then we can write the energy around the point $k_0 = 0$ of interest in terms of the effective mass, and the operator $E_n(-i\nabla)$ thus becomes

$$E_n(k) \approx E_c(r) + \frac{\hbar^2 k^2}{2m^\star} \to E_n(-i\nabla) \approx E_c(r) - \frac{\hbar^2}{2m^\star}\nabla^2, \tag{20}$$

and the Schrödinger equation takes the enormously simplified form

$$[-\frac{\hbar^2}{2m^\star}\nabla^2 + V_{imp}(r)]C(r) = [E - E_c(r)]C(r), \tag{21}$$

which is the celebrated "Effective Mass Approximation". Take a moment to note what has been achieved. The Schrodinger equation has been re-cast into a much simpler problem of a particle of mass $m^\star$, moving in a potential $E_c(r) + V(r)$! All information about the bandstructure and crystal potential has been lumped into the effective mass $m^\star$. The wavefunctions are envelope functions $C(r)$, from which one recovers the real wavefunction of

the wavepacket by multiplying with the Bloch function - $\psi(r) \approx \phi_{n0}(r)C(r) = u_{n0}(r)C(r)$. The envelope functions $C(r)$ can be actually determined for any potential - it amounts to solving the Schrödinger equation for a particle in the potential $E_c(r) + V(r)$. Note that the envelope function in the absence of any impurity potential $V(r) = 0$ is given by

$$C(r) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{22}$$

and the corresponding eigenvalues of the Schroödinger equation are given by

$$E = E_c(r) + \frac{\hbar^2|\mathbf{k}|^2}{2m^\star}. \tag{23}$$

If we consider electrons at the bottom of the conduction band, $E_c(r)$ is the spatial variation of the conduction band edge - exactly what one draws in band diagrams. An impurity potential can now be included as a perturbation to the periodic crystal, and the new energy eigenvalues can be found. As an example, consider an ionized impurity, which has a Coulomb potential. The effective mass equation takes the form

$$[-\frac{\hbar^2}{2m^\star}\nabla^2 - \frac{e^2}{4\pi\epsilon r}]C(r) = (E - E_c)C(r), \tag{24}$$

which is identified as the same as the classic problem of a hydrogen atom, albeit with two modifications - the mass term is an effective mass instead of the free electron mass, and the dielectric constant is that of the semiconductor. Then, the new energy levels that appear are given by

$$E - E_c = E_\infty \frac{m^\star}{\epsilon_r^2}, \tag{25}$$

and the effective Bohr-radius is given by

$$a_B^\star = a_B \frac{\epsilon_r}{m^\star} \tag{26}$$

In bulk semiconductors, the band-edge variation in real space can be varied by applying electric fields, or by doping variations. In semiconductor heterostructures, one can further engineer the variation of the band-edge $E_c(r)$ in space by quasi-electric fields - the band edge can behave as quantum-wells, wires, or dots, depending upon composition of the semiconductor. The effective mass approximation is a natural point of departure, where analysis of such low-dimensional structures begins.

4 Transport theory - Real Semiconductors

If you apply an electric field and measure the drift velocity of electrons in a bulk semiconductor, you will see what is called the velocity-field characteristic. It is schematically shown in the left part of Figure 6. Drift velocity is proportional to the applied field for low fields; this is the 'ohmic' regime characterized by the electron mobility $\mu = v_d/F$. As you ramp up the field, the velocity will approach a saturation value - this is the 'hot-electron' regime

of transport. If you run Hall measurements on mobile carriers in a bulk semiconductor and a two-dimensional electron gas at varying temperatures and measure the mobility, you will see behavior depicted in the right part of Figure 6. Most of our study in transport theory is geared towards understanding from a quantum-mechanical viewpoint the reasons for this kind of behavior.

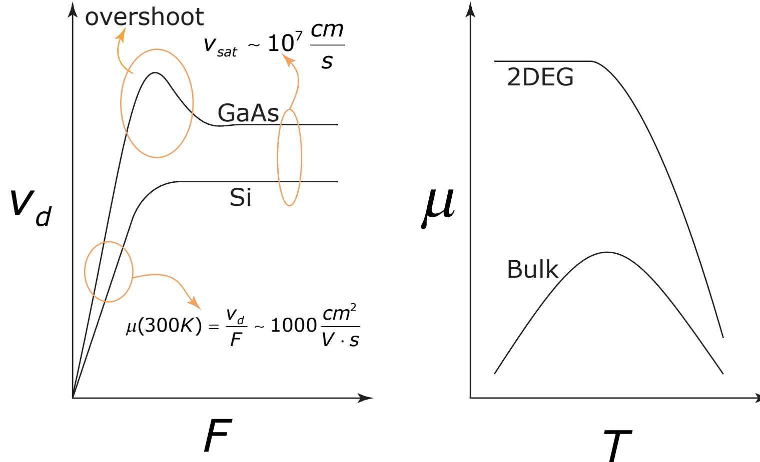

Figure 6: Velocity field curves (left) of Si and GaAs showing some definitions. Mobility versus temperature for generic semiconductor bulk and two-dimensional electron gas (2DEG). The figure is for motivating the study of transport theory

4.1 Setting up the Quantum Mechanical Problem

We begin with a completely general time-dependent Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = (H_0 + H_1)\Psi(\mathbf{r}, t) \tag{27}$$

We split the Hamiltonian into two parts. The first part,

$$H_0 = -\frac{\hbar^2}{2m^\star} \nabla^2 + E_C(\mathbf{r}) \tag{28}$$

is the Effective Mass Hamiltonian that we derived earlier (Equation 21). For a bulk semiconductor, H_0 is the crystal potential, which upon solution yields the bandstructure as the eigenvalues. We then concentrate on only the small region around the band extrema (CB minima or VB maxima).

The second part of the Hamiltonian is a small perturbation to the CB minimum introduced by impurities such as charged dopants, vacancies, or dislocations -

$$H_1 = Perturbation \tag{29}$$

Solving the exact part gives us the set of eigenfunctions $\Phi_m(\mathbf{r})$ and eigenvalues E_m corresponding to the Hamiltonian of the pure system, H_0 -

$$H_0\Phi_m(\mathbf{r}) = E_m\Phi_m(\mathbf{r}) \tag{30}$$

For the time dependent solution of the problem (with the scattering potentials), we expand the time dependent wavefunction in terms of the unperturbed eigenfunctions -

$$\Psi(\mathbf{r}, t) = \sum_m \psi_m(t)\Phi_m(\mathbf{r}) \tag{31}$$

The time-dependent equation (Equation 27) then becomes

$$i\hbar\frac{d}{dt}\psi_m = \sum_n \langle m|H|n\rangle\psi_n \tag{32}$$

Since the H_0 part has been solved exactly, we can use it to get the relation

$$\frac{d}{dt}\psi_m(t) + \frac{iE_m}{\hbar}\psi_m = \sum_n \frac{\langle m|H_1|n\rangle}{i\hbar}\psi_n \tag{33}$$

For a homogenous semiconductor, the solutions of the pure Hamiltonian are planewaves

$$\Phi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}}e^{i\mathbf{k}\cdot\mathbf{r}}, \tag{34}$$

where $V = L_xL_yL_z$ is the sample volume. The quantization of wavevectors is given by $k_i = n_i\frac{2\pi}{L_i}, i = x, y, z$. The corresponding eigenvalues are given by

$$E_{\mathbf{k}} = E_C + \frac{\hbar^2|\mathbf{k}|^2}{2m^\star} \tag{35}$$

4.2 Fermi's Golden Rule

Evaluation of scattering rates requires us to use one of the most useful results of time dependent perturbation theory in quantum mechanics. Don't throw your hands up in despair, since it is not only simple, it is extremely useful too. Knowing it will help you figure out a lot of stuff in your research, no matter which field you are thinking about working in. The result is known by the glorified name of 'Fermi's Golden Rule'[5].

We will derive Fermi's Golden Rule (FGR[6]) with the following problem in mind. We have an incident electron with wavevector along the z direction (Figure 7) in a uniform bulk semiconductor of dimensions $V = L_xL_yL_z$. The semiconductor is pure, except for a small impurity with a time-independent potential $U(r)$ located somewhere in the bulk. The incident beam wavefunction is given by e^{ikz} where $k = |\mathbf{k}|$.

We have a detector at an angular position (θ, ϕ) and would like to find the probability $P_S(\theta, \phi)$ of finding the portion of the beam scattered by the potential $U(r)$ at that angular

[5] Golden is no exaggeration; however, refer to Kroemer's book, pg. for what he has to say about licensing the rule!

[6] I hate acronyms myself when I come across them. Please bear with me for this one three letter word!

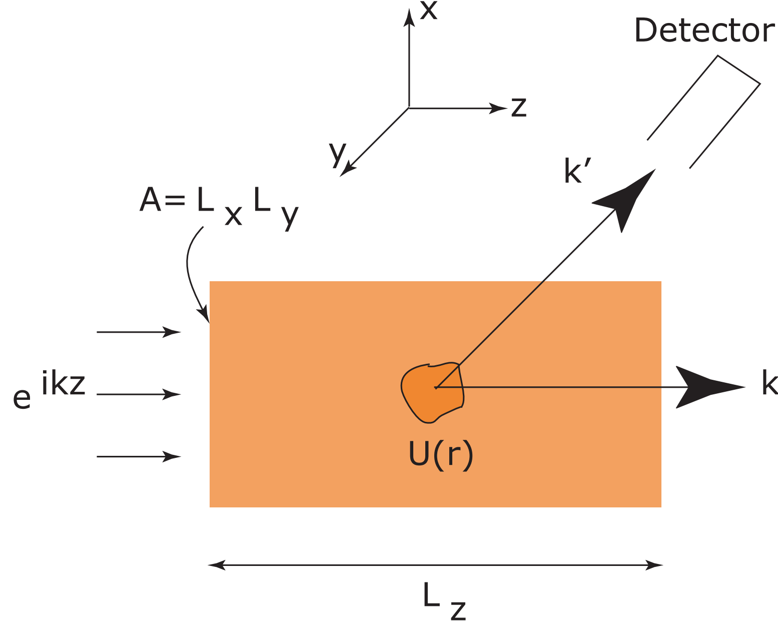

Figure 7: Figure for setting up the problem of scattering. An incident electron is scattered by a scattering potential $U(r)$ in an otherwise pure semiconductor. A detector is placed to find the angular dependence of scattering.

location. The result we derive is valid only in a weak scattering limit, i.e., the total scattering probability is much less than unity

$$P_S = \int_0^{\pi} d\theta \sin(\theta) \int_0^{2\pi} d\phi P(\theta, \phi) \ll 1 \tag{36}$$

In Equation 33, we saw how the time-coefficients of the solved wavefunction are linked to impurity potentials. A perturbation always causes transitions between allowed states of a exactly solved problem with known eigenfunctions and eigenvalues. In this problem, the exactly solved problem is again of the semiconductor bandstructure with allowed states indexed by the wavevectors $|\mathbf{k}\rangle$. The eigenfunction and eigenvalues are given by Equation 34 and Equation35 respectively.

To evaluate the probability of a state in the eigenstate $|\mathbf{k}\rangle$ to scatter into any state $|\mathbf{k}'\rangle$, we have to find the amplitude of that state in the eigenfunction expansion and square it. The amplitudes are governed by Equation 33. We now revert to the quantum number index $(m, n) \to (k, k')$ since the allowed states have been solved. Starting with the equation

$$\frac{d}{dt}\psi_{k'}(t) + \frac{iE_{k'}}{\hbar}\psi_{k'} = \sum_{k'' \neq k'} \frac{\langle k''|H_1|k'\rangle}{i\hbar}\psi_{k''} \tag{37}$$

The electron is incident with a wavevector along z at $t = 0$. So, the initial wavefunction is

$$\Psi(\mathbf{r}, t=0) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{38}$$

From Equation 31, we get $\Psi(\mathbf{r}, t=0) = \sum_{k'} \psi_{k'}(t=0)\Phi_{k'}(\mathbf{r})$. So $\psi_{k'}(t=0) = 1$ for $k' = k$ and 0 for all other k'. To find the amplitudes $\psi_{k'}(t)$ at all other times, we have to solve Equation 33. If the scattering potential is absent ($U(r) = 0$), then the equation is solved easily, giving

$$\psi^0_{k'}(t) = \psi^0_{k'}(0)e^{-i\frac{E_{\mathbf{k'}}t}{\hbar}}. \tag{39}$$

We label the solution by '0' to emphasize that it is a zero order solution. From Equation 31, it is easy to see that only the amplitude of state $|\mathbf{k}'\rangle = |\mathbf{k}\rangle$ evolves with time, whereas the amplitude of all other states is 0. When there is a nonzero scattering potential, there are transitions between different eigenstates, so the amplitudes of $|\mathbf{k}'\rangle \neq |\mathbf{k}\rangle$ will not be zero anymore. However, the solution of Equation 33 is not simple anymore, and we have to take the route of successive approximations now. It is crucial to understand this part to appreciate the limitations of FGR.

We start by substituting the solution of the zero order solution on the RHS of Equation 33 and solve for the first order solution. This process is repeated for higher order solutions. However, if the scattering potential happens to be weak, it turns out that the first order solution is adequate and leads to Fermi's Golden Rule. Doing that, our time evolution equation becomes

$$\frac{d}{dt}\psi_{k'}(t) + \frac{iE_{k'}}{\hbar}\psi_{k'} = \frac{\langle k'|U(r)|k\rangle}{i\hbar} e^{-i\frac{E_{\mathbf{k'}}t}{\hbar}} \tag{40}$$

Subject to the initial condition Equation 39, the solution gives the amplitude of state $|\mathbf{k}'\rangle$ at time T

$$\psi_{k'}(T) = e^{-i\frac{E_{\mathbf{k'}}T}{\hbar}} \frac{\langle k'|U(r)|k\rangle}{i\hbar} \int_0^T e^{-i\frac{(E_{\mathbf{k}}-E_{\mathbf{k'}})T}{\hbar}} dt \tag{41}$$

Thus, the probability of finding an electron scattered into the final state $|\mathbf{k}'\rangle$ is the square of the amplitude of that state. That is exactly what we have determined in Equation 41. So,

$$P(\mathbf{k}', \mathbf{k}) = |\psi_{k'}(T)|^2 = \frac{|\langle \mathbf{k}'|U(r)|\mathbf{k}\rangle|^2}{\hbar^2} |\int_0^T e^{-i\frac{(E_{\mathbf{k}}-E_{\mathbf{k'}})T}{\hbar}} dt|^2 \tag{42}$$

T can be seen as the time taken for the electron to traverse the length L_z; it is thus given by $T = \frac{L_z}{v} = \frac{m^\star V}{\hbar k A}$, where $A = L_x L_y$ is the cross sectional area. The probability can be recast in the form

$$P(\mathbf{k}', \mathbf{k}) = T\frac{|\langle \mathbf{k}'|U(r)|\mathbf{k}\rangle|^2}{\hbar^2} g(E_{\mathbf{k}'} - E_{\mathbf{k}}) \tag{43}$$

where

$$g(\varepsilon) = \frac{1}{T} |\frac{\sin(\varepsilon T/2\hbar)}{(\varepsilon T/2\hbar)}|^2 \tag{44}$$

Using a special property of the function $\int_{-\infty}^{+\infty} d\varepsilon g(\varepsilon) = 2\pi\hbar$, we see that in the limit of a long time $T \to \infty$, the function g becomes a delta function. Thus, $Lim_{T\to\infty} g(\varepsilon) = 2\pi\hbar\delta(\varepsilon)$. This allows us to write the probability as

$$P(\mathbf{k}', \mathbf{k}) = TS(\mathbf{k}', \mathbf{k}) \tag{45}$$

where

$$S(\mathbf{k}', \mathbf{k}) = \frac{2\pi}{\hbar} |\langle \mathbf{k}' | U(r) | \mathbf{k} \rangle|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}}) \tag{46}$$

This is the traditional way of stating the famous FGR for time-independent stationary potentials. Note that this is not yet a measurable probability, since the measured probability will be a sum over all $|\mathbf{k}'\rangle$ states. Thus, to find a measurable probability, we need to sum the probability over all such final states.

$$P_S = \sum_{\mathbf{k}'} P(\mathbf{k}', \mathbf{k}) \tag{47}$$

Using the recipe $\sum_{\mathbf{k}'}(...) \to \frac{V}{(2\pi)^3} \int d^3k'(...)$, we get

$$P_S = \frac{V}{(2\pi)^3} \int_0^\infty \int_0^\pi \int_0^{2\pi} k'^2 dk' \sin(\theta) d\theta d\phi P(\mathbf{k}', \mathbf{k}) \tag{48}$$

Remembering the problem we started out with, we see that the angular probability can now be evaluated. Comparing with Equation 36, we write

$$P(\theta, \phi) = \frac{V}{(2\pi)^3} \int_0^\infty dk' k'^2 P(\mathbf{k}', \mathbf{k}) \tag{49}$$

Using FGR for $P(\mathbf{k}', \mathbf{k})$, we get

$$P(\theta, \phi) = \frac{VT}{4\pi^2\hbar} \int dk' k'^2 |\langle \mathbf{k}' | U(r) | \mathbf{k} \rangle|^2 \delta(E_{\mathbf{k}'} - E_{\mathbf{k}}) \tag{50}$$

We now use two important techniques. We transform the variable of integration to the argument of the delta function ($k' \to E_{\mathbf{k}'}$) using the identity $dk' = \frac{dk'}{dE_{\mathbf{k}'}} dE_{\mathbf{k}'} = \frac{m^\star k'}{\hbar^2} dE_{\mathbf{k}'}$. We do this to take advantage of a very important property of a delta function -

$$\int_a^b d\varepsilon \delta(\varepsilon - \varepsilon_0) F(\varepsilon) = F(\varepsilon_0) \tag{51}$$

if $a < \varepsilon_0 < b$, and is 0 otherwise. Using this useful result, we get

$$P(\theta, \phi) = \frac{VT}{4\pi^2\hbar} \frac{m^\star k}{\hbar^2} |\langle \mathbf{k}' | U(r) | \mathbf{k} \rangle|^2 \tag{52}$$

The delta function enforced energy conservation in the scattering process ($E_{\mathbf{k}'} = E_{\mathbf{k}}$), and as a result, also required us to conserve the magnitude of momentum ($k' = k$). Using the explicit form of the matrix element $\langle \mathbf{k}' | U(r) | \mathbf{k} \rangle = \frac{1}{V} \int d^3 U(\mathbf{r}) e^{i(\mathbf{k}'-\mathbf{k})\cdot r}$, we write the probability in the form

$$P(\theta,\phi)=\frac{(m^\star)^2}{4\pi^2\hbar^4 A}|\int d^3U(\mathbf{r})e^{i(\mathbf{k}'-\mathbf{k})\cdot r}|^2 \tag{53}$$

The angular dependence is hidden in the $(\mathbf{k}'-\mathbf{k})$ term. We will encounter it when studying ionized impurity scattering that has a strong angular dependence.

This result can be recast in a visually more appealing form by rewriting $P(\theta,\phi)$ as $P(\theta,\phi)=\sigma(\theta,\phi)/A$, where

$$\sigma(\theta,\phi)=\frac{(m^\star)^2}{4\pi^2\hbar^4}|\int d^3U(\mathbf{r})e^{i(\mathbf{k}'-\mathbf{k})\cdot r}|^2 \tag{54}$$

is referred to as the 'scattering cross section'. The total scattering scattering cross section σ_S is evaluated by integrating over all angular contributions

$$\sigma_S=P_S A=\int_0^\pi d\theta\sin(\theta)\int_0^{2\pi}d\phi\sigma(\theta,\phi) \tag{55}$$

Scattering cross section lets us visualize the scatterer as a solid obstacle of cross sectional area σ_S, which will scatter an electron with the probability σ_S/A. The typical problem in transport is electrons moving in a semiconductor with a density N_I of scattering centers (or impurities. In time t, the total number of scatterers an electron encounters will be $N_I Avt$ where v is the velocity. Thus, the probability that the electron will be scattered in time t is

$$Pr(t)=\frac{t}{\tau}=\frac{t}{\frac{1}{\sigma_S N_I v}} \tag{56}$$

This relation holds only for short intervals ($t\ll\tau$). For large t, it is $Pr(t)=1-e^{-\frac{t}{\tau}}$ (prove it!). Thus using $T=\frac{L_z}{v}=\frac{m^\star V}{\hbar kA}$, $P(\mathbf{k}',\mathbf{k})=TS(\mathbf{k}',\mathbf{k})$, and $P_S=\sum_{\mathbf{k}'}P(\mathbf{k}',\mathbf{k})$, the mean free time between two scattering events (collisions) can be written as

$$\frac{1}{\tau_c}=\sigma_S v N_I=\sum_{\mathbf{k}'}(N_I V)S(\mathbf{k},\mathbf{k}') \tag{57}$$

This relation is easy to understand. $\sum_{\mathbf{k}'}S(\mathbf{k},\mathbf{k}')$ is the rate at which an electron of momentum $\mathbf{k}$ is scattered by a single impurity. Multiplying by the total number of impurities $N_I V$, the total scattering rate is obtained.

To find the scattering time that goes into determining mobility, we have to use the momentum scattering rate, τ_m. Experimental mobility is related to the momentum scattering time by $\mu=e\tau_m/m^\star$. The scattering rate we just derived is the inter-collision time. Collision rate is $\frac{1}{\tau_Q}$ where τ_Q, the collision time determined by FGR is also called the quantum lifetime. From Figure 8, we note that some scattering events can change the angle of the $\mathbf{k}$ vector by very large amounts. Others have a weak effect. If there is small angle scattering, the current is higher, and hence the mobility is higher, so the momentum scattering *rate* is smaller. If the angle of scattering is large, momentum scattering rate will be smaller than the quantum scattering rate. This can be heuristically built into the formalism by writing

$$\frac{1}{\tau_m}=\frac{1}{\tau_Q}[1-\cos(\theta)]=\frac{1}{\tau_Q}[1-\frac{\mathbf{k_0}\cdot\mathbf{k_1}}{|k_0|^2}] \tag{58}$$

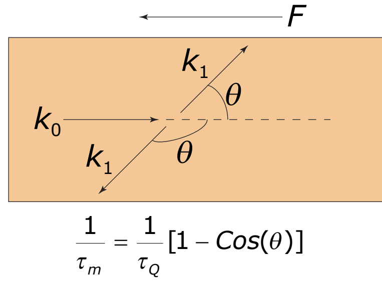

Figure 8: Figure to show the difference between the quantum lifetime τ_Q and the transport lifetime τ_m. If there is a small angle scattering, current is higher, and thus mobility is higher. For large angle scattering, current is lower, and mobility is lower. Thus, heuristically, we can say that the momentum lifetime and quantum lifetime (which is time between collisions) are related by $\frac{1}{\tau_m} = \frac{1}{\tau_Q}(1 - \cos(\theta))$

This result can be got by a rigorous analysis, so we believe in it and go ahead and use it to calculate mobility. Thus, momentum scattering rate is given by incorporating this small change to Equation 37 -

$$\frac{1}{\tau_c} = \sigma_S v N_I = \sum_{\mathbf{k}'} (N_I V) S(\mathbf{k}, \mathbf{k}')(1 - \cos(\theta)) \tag{59}$$

where $\cos(\theta) = \frac{\mathbf{k} \cdot \mathbf{k}'}{|\mathbf{k}|^2}$.

4.3 Example illustrating application of Fermi's Golden Rule

To get comfortable with Fermi's Golden rule we apply it to study scattering of a wavelike particle from a rectangular barrier of height $\triangle$ and thickness d.

Consider the problem of a wave incident on the rectangular barrier. The classical probability of reflection R is 1 for energy $E < \triangle$ and 0 for $E > \triangle$. Quantum mechanically, the behavior is more interesting. The reflection probability is given by

$$R = \frac{2R - 2R\cos(2k_1 d)}{1 + R^2 - 2R\cos(2k_2 d)} \tag{60}$$

where $R = |r|^2$ is the reflection probability for a single barrier, and $r = \frac{k_1 - k_0}{k_1 + k_0}$ is the reflection *amplitude* for a single barrier. Let us see how we can apply Fermi's Golden Rule to solve the same problem.

Fermi's Golden Rule will tell us the probability of scattering due to a small perturbation to the exactly solved Hamiltonian potential. Here, the exactly solved Hamiltonian is the free electron Hamiltonian. The solved wavefunctions are planewave states, given by

$$|\mathbf{k}\rangle = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} \tag{61}$$

If the particle is reflected from the barrier, we say that it has been scattered. FGR will tell us the probability of this event -

$$P_S = T\sum_{k'} S(k',k) = \sum_{\mathbf{k}'} \frac{2\pi T}{\hbar}|U(\mathbf{k}',\mathbf{k})|^2\delta(E_{\mathbf{k}'} - E_{\mathbf{k}'}) \quad (62)$$

The barrier is in the z direction; the initial momentum vector of the electron points in the z direction - $k_z = k_0, k_x = k_y = 0$. So the matrix element of the barrier potential is given by

$$\langle\mathbf{k}'|\triangle|\mathbf{k}\rangle = \frac{\triangle}{V}\int_0^{L_x} dx e^{i(k_x-k_x')x}\int_0^{L_y} dx e^{i(k_y-k_y')y}\int_0^d dz e^{i(k_z-k_z')z} \quad (63)$$

which, on using results of Example 1 evaluates to

$$\langle\mathbf{k}'|\triangle|\mathbf{k}\rangle = \frac{\triangle}{L_z}\delta_{k_x,k_x'}\delta_{k_y,k_y'}(\frac{e^{i(k_z-k_z')d}-1}{i(k_z-k_z')}) \quad (64)$$

Thus, the scattered electrons have the same transverse momenta as the incident $(k_x,k_y) = (k_x',k_y')$. So, the sum over all scattered $\mathbf{k}'$ runs over only k_z'. Thus, we have

$$P_S = \sum_{k_z'} \frac{2\pi T}{\hbar}|U(\mathbf{k}',\mathbf{k})|^2\delta(E_{\mathbf{k}'} - E_{\mathbf{k}'}) \quad (65)$$

where

$$|U(\mathbf{k}',\mathbf{k})|^2 = \frac{4\triangle^2}{L_z^2(k-k_z')^2}\sin^2(\frac{(k-k_z')d}{2}) \quad (66)$$

Converting the sum over k_z' to an integral using the recipe $\sum_{k_z'}(...) \rightarrow \frac{L_z}{2\pi}\int dk_z'(...)$ and using the transformation $dk_z' = dE_{\mathbf{k}'}\frac{dk_z'}{dE_{k'}}$ we get

$$P_S = \frac{L_z T}{\hbar}\int dE_{\mathbf{k}'}\delta(E_{\mathbf{k}'} - E_{\mathbf{k}})\frac{dk_z'}{dE_{\mathbf{k}'}}|U_{\mathbf{k},\mathbf{k}'}|^2 \quad (67)$$

The delta function forces energy conservation; $\mathbf{k}' = \mathbf{k} \rightarrow k_z' = k_0$. Also, $\frac{dk_z'}{dE_{k'}} = \frac{m^\star}{\hbar^2 k_z'}$. Using this, the scattering probability becomes

$$P_S = \frac{m^\star T\triangle^2}{\hbar^3 k_0^3 L_z}\sin^2(k_0 d) \quad (68)$$

Using $T = \frac{L_z}{v_0} = \frac{m^\star L_z}{\hbar k_0}$, we get

$$P_S = (\frac{\triangle m^\star}{\hbar^2 k_0^2})^2\sin^2(k_0 d) \quad (69)$$

This is the final scattering rate. It certainly does not look like the quantum mechanical result for probability of reflection as in Equation 38. The difference brings in focus the assumptions that are inherent in FGR. Consider the QM result. Let us see how the expression looks for *small* R. Then, neglecting R and R^2 in the D^r, the expression for reflection probability Equation 38 reduces to

$$R_B \simeq 2R - 2R\cos(2k_2d) = 4R\sin^2(k_2d) = \frac{4(k_2-k_1)^2}{(k_2+k_1)^2}\sin^2(k_2d) \tag{70}$$

If R is small, $k_2 \simeq k_1 = k_0$, and the expression becomes

$$R_B \simeq (\frac{k_2-k_1}{k_0})^2\sin^2(k_0d) \tag{71}$$

Since $k_2 = \sqrt{\frac{2m^*(E-E_C-\triangle)}{\hbar^2}}$ and $k_2 = \sqrt{\frac{2m^*(E-E_C)}{\hbar^2}}$, $k_2^2 - k_1^2 = \frac{2m^\star\triangle}{\hbar^2} \simeq 2k_0(k_2-k_1)$. Using this, the result becomes

$$R_B \simeq (\frac{\triangle m^\star}{\hbar^2 k_0^2})^2\sin^2(k_0d) \tag{72}$$

This is exactly the result we got from the FGR calculation. So the assumption in FGR of a *weak* scattering potential is very important and should always be borne in mind when applying the rule.

FGR has widespread customers. A short list of it's clients other than semiconductor transport are optics, atomic processes, spectroscopy (Raman, NMR, etc) etc. Wherever there is a time-dependent transition between quantum mechanical states is involved, FGR rules.

5 Case study of scattering in semiconductors

Low field mobility is calculated by doing the following -

1) Identify the defect (perturbation) and choose a suitable model for it's potential $U_i(r)$ for the ith type of impurity.

2) Calculate the momentum scattering rate for the particular type of impurity and sum it over all available final states $\mathbf{k}'$.

3) Calculate the total scattering rate using the Matheissen's rule - $\frac{1}{\tau} = \sum_i \frac{1}{\tau_i}$.

4) Average the momentum scattering rate for drift velocity (Eqn 74).

5) Calculate the mobility using $\mu = e\tau_m/m^\star$.

We have gone through the whole formalism and have armed ourselves now with all the tools needed for doing the above steps. The following sections would be really short since we have done all the hard work earlier.

5.1 Example - Ionized impurity scattering in bulk semiconductors and 2DEGs

Consider an ionized impurity in a bulk semiconductor. We find the quantum scattering time τ_Q for scattering from state $|\mathbf{k}\rangle$ to a general final state $|\mathbf{k}'\rangle$. The impurity may be an intentional donor or acceptor, or a charged vacancy state or any such impurity. Let the charge on the impurity be Z Coulombs. Equation 34, Equation 67 yield

$$\sigma(\theta,\phi) = \frac{4L_D^4}{a_0^2}\frac{1}{(1+qL_D)^2} \tag{73}$$

where $a_0 = \frac{4\pi\epsilon_0 K\hbar^2}{m^\star q^2}$, the effective Bohr radius. Since Coulombic scattering is elastic, $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ can be written as

$$q^2 = 2k^2[1-\cos(\theta)] \tag{74}$$

where $k = |\mathbf{k}| = |\mathbf{k}'|$. Using this, writing $\gamma = 2kL_D$ and integrating over all angles (θ,ϕ)[7], we get the total scattering cross section as

$$\sigma_S = \frac{\pi}{k^4 a_0^2}\frac{\gamma^4}{1+\gamma^2} \tag{75}$$

Thus, from Eq. 37, the quantum scattering rate is given by

$$\frac{1}{\tau_Q(k)} = N_I \cdot \frac{\hbar k}{m^\star} \cdot \frac{\pi}{k^4 a_0^2}\frac{\gamma^4}{1+\gamma^2} \tag{76}$$

The momentum relaxation rate can be got similarly by using Equation 39. The result is

$$\frac{1}{\tau_m(k)} = N_I \cdot \frac{\hbar k}{m^\star} \cdot \frac{2\pi}{k^4 a_0^2}[ln(1+\gamma^2) - \frac{\gamma^2}{1+\gamma^2}] \tag{77}$$

The total momentum scattering time is found by an averaging procedure. It is easy to convert $\tau_m(k) \rightarrow \tau_m(E)$, since the dispersion $E(k) = \frac{\hbar^2 k^2}{2m^\star}$ is known. For drift velocity along an applied field F_z (along the z axis), the averaging[8] procedure for a Fermi-Dirac distributed carrier population is given by

$$\langle\tau_m\rangle = -\frac{2}{3}\int_0^\infty \frac{\tau_m(E)\frac{\partial f_0}{\partial(E/k_BT)}(E/k_BT)^{3/2}d(E/k_BT)}{\int_0^\infty f_0(E/k_BT)^{1/2}d(E/k_BT)} \tag{78}$$

Using this recipe, we get the mobility as

$$\mu_{imp} = \frac{e\langle\tau_m\rangle}{m^\star} = \frac{2^{7/2}(4\pi\epsilon_0 K)^2(k_BT)^{3/2}}{\pi^{3/2}Z^2e^3(m^\star)^{1/2}N_I[ln(1+\gamma_0^2) - \frac{\gamma_0^2}{1+\gamma_0^2}]} \tag{79}$$

where Z is the total electron charges in the impurity and $\gamma_0 = 2\frac{m^\star}{\hbar}(\frac{2}{m^\star}3k_BT)^{1/2}L_D$. The dependence of this mobility on temperature T and impurity density N_I is seen to be

[7] From now on, you need to evaluate all integrals yourself and verify the results!

[8] Seeger, Semiconductor Physics, Pg 50

$$\mu_{imp} \propto \frac{T^{\frac{3}{2}}}{N_I} \tag{80}$$

In bulk semiconductors at low temperatures, ionized impurities set the mobility limits. This is the mobility dependence in Figure 4 at low temperatures (right curve, increasing mobility).

For 2DEGs, the phenomena of quantization and degeneracy of conducting electrons (all conducting electrons have $k \simeq k_F = \sqrt{2\pi n_{2D}}$) change the averaging procedure[9] and make the problem easy. Also, the remote location of impurities (modulation doping) introduces an exponential decay of the Coulomb potential. This results in much higher mobilities. An important fact is there is no temperature activation of carriers required for a 2DEG; the 2DEG density is independent of temperature. Thus, the ionized impurity limited mobility is *independent of temperature* for a 2DEG. This results in the 'flat' mobility saturation at low temperatures for 2DEGs (right plot, Figure 3).

5.2 Other scattering mechanisms and where to read about them

At room temperature, the mobility in a reasonably pure semiconductor is almost always limited by phonon scattering. Phonon scattering is very important for operational devices and is treated well in Seeger's book (Semiconductor Physics). In fact, all scattering mechanisms are very well treated in this book. I highly recommend it if you want to learn more on transport and scattering. However, Seeger's book is an 'old' book, it does not include quantum structures like quantum wells, etc. For a newer treatment (though not as comprehensive as Seeger), John Davies' book (Physics of low dimensional semiconductors) is recommended. Another book which does a good and compact treatment of transport properties is Wolfe, Holonyak and Stillman's 'Physical properties of semiconductors'. For a comprehensive review of 2DEG transport properties, and every conceivable property of the 2DEG, refer to the 'Bible' of 2DEGs - a journal article that can be downloaded from the web. The article is by Ando, Fowler and Stern; it appeared in Reviews of Modern Physics, (Rev. Mod. Phys. 54, 437-672 (1982)).

Alloy disorder scattering is important for transport in ternary (or even binary - SiGe) alloys. For 2DEGs, interface roughness scattering becomes severe at high 2DEG densities. Dislocation scattering is proving to be a hindrance for the III-V nitride semiconductors making it bigger than what they already are.

5.3 High field transport

Consider an electron moving in the presence of impurity scattering with an increasing electric field. In the 'ohmic' regime we defined a mobility as the ratio of drift velocity and field. Now in the microscopic picture, the electron starts from zero velocity, accelerates by gaining energy from the field, and then scatters. This is the way it traverses the length of a semciconductor device. It is depicted in Figure 9.

[9] John Davies, Pg 356

After every scattering event, the electron starts gaining momentum (and hence energy) from the applied field. The energy it acquires before the next collision increases as the electric field increases. At a particular value of the electric field, the energy gained between two collisions becomes larger than the optical phonon energy $\hbar\omega_{op}$. When this happens, the probability of the electron losing all it's energy to the lattice by the emission of an optical phonon becomes large. The emission of a phonon results in no increase in current; it just heats the device. Any further increase of the field will cause an increase in the number of optical phonons emitted, with no increase in the drift velocity. The drift velocity is said to be saturated.

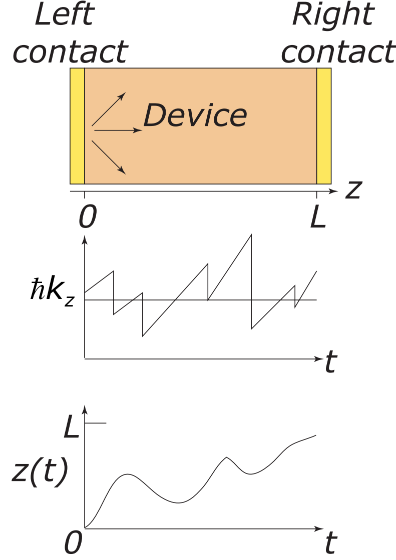

Figure 9: Figure depicting the actual transport process with time. The second shows how the z component of momentum increases linearly with time ($k_z(t) = \frac{F}{\hbar}t$) till it gets scattered to a random state. The third figure shows the progress of the particle in real space with time as it traverses the length of the device.

Thus, a fair back of the envelope estimate of the saturated drift velocity can be made by equating the kinetic energy at the saturated drift velocity $KE = \frac{1}{2}m^{\star}v_d^2$ to the optical phonon energy in the semiconductor -

$$v_d^{sat} \sim \sqrt{\frac{2\hbar\omega_{op}}{3m^{\star}}} \tag{81}$$

For GaN, $\hbar\omega_{op} = 92meV, m^{\star} = 0.2m_0$, and the drift velocity estimate is $v_d^{sat} \sim 4 \times 10^7 cm/s$. This is a good 'order of magnitude' estimate. The true value is $v_d^{sat} \sim 2 \times 10^7 cm/s$. Thus, the velocity-field curve saturates; this is shown in Figure 3, left graph.

In direct gap semiconductors (GaAs, GaN, etc) there are two CB minimum in the $E-k$ diagram. Γ valley is the lowest and X valley is the second lowest valley. The X valley has a smaller curvature and hence a larger effective mass. At high fields, when high k^s are achieved

between collisions, it becomes possible for electrons to transfer from the Γ tot eh X valley; the onset of this phenomena causes a region where the velocity *decreases* with increasing field. This is the origin of negative differential resistance (NDR) which is put to use to make Gunn oscillators which have very high frequency oscillations.

The velocity-field characteristics are of utmost importance in submicron devices. The reason is that for the same applied voltage, the field in a small length device will be large. Thus, saturation of drift velocity is achieved under fairly low voltages. With the miniaturization of transistors and decreasing gate lengths, it is very important for a device engineer to understand the hidden theories behind the velocity-field characteristics for innovation.

5.4 Transport Regimes

In a sufficiently pure semiconductors, electron wavelengths are sufficiently delocalized and can spread over large distances (large $\lambda \rightarrow$ small k). For such cases, transport occurs in the bands, this regime of transport is called band transport. Figure 10 shows schematically this form of transport. For low field regime, the transport proceeds by scattering from impurities. For high field regime, there are optical phonon emissions, but all transport is in the conduction (or valence) band. If there is a lot of disorder in a semiconductor (say

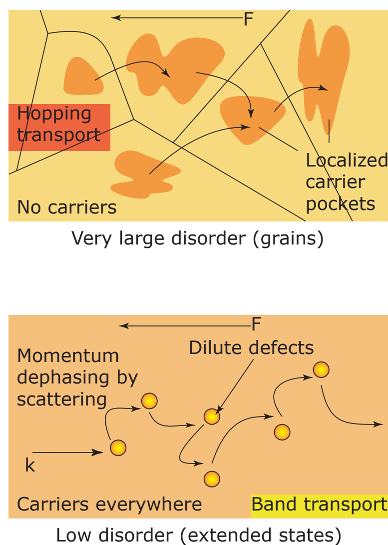

Figure 10: Band transport and Hopping regimes of transport. Band transport has been discussed in this work. Hopping transport requires activation to surmount the barriers between nearby localized pockets of carriers.

an polycrystalline semiconductor with grain boundaries), then there will be small localized pockets of carriers which will have to surmount the potential barriers between them to carry current. This process requires an activation energy, and the activation energy can be measured. This form of transport is very different in characteristic from band transport. The transport is called 'activated' or 'hopping' transport because the carriers hop from one

localized state to the other. Theoretical treatment of transport in such disordered systems requires percolation theory and sophisticated techniques such as Green's functions, which I leave for you to explore.

6 Formal Transport theory

Much of the following summary is collected from textbooks and research articles. The main references for this section are Seeger [1], Wolfe et. al. [2], and Davies [3]. No claim to originality is made for much of the material. The subsection on *generalization* of mobility expressions for arbitrary dimensions and arbitrary degeneracy is original, though much of it is inspired from the references.

6.1 Boltzmann transport equation

A distribution-function $f(\mathbf{k}, \mathbf{r}, t)$ is the probability of occupation of an electron at time t at $\mathbf{r}$ with wavevectors lying between $\mathbf{k}, \mathbf{k} + \mathbf{dk}$. Under equilibrium ($\mathbf{E} = \mathbf{B} = \nabla_r f = \nabla_T f = 0$, i.e., no external electric ($\mathbf{E}$) or magnetic ($\mathbf{B}$) field and no spatial and thermal gradients), the distribution function is found from quantum-statistical analysis to be given by the Fermi-Dirac function for fermions -

$$f_0(\varepsilon) = \frac{1}{1 + e^{\frac{\varepsilon_{\mathbf{k}} - \mu}{k_B T}}}, \tag{82}$$

where $\varepsilon_{\mathbf{k}}$ is the energy of the electron, μ is the Fermi energy, and k_B is the Boltzmann constant.

Any external perturbation drives the distribution function away from the equilibrium; the Boltzmann-transport equation (BTE) governs the shift of the distribution function from equilibrium. It may be written formally as [2]

$$\frac{df}{dt} = \frac{\mathbf{F_t}}{\hbar} \cdot \nabla_{\mathbf{k}} f(\mathbf{k}) + \mathbf{v} \cdot \nabla_{\mathbf{r}} f(\mathbf{k}) + \frac{\partial \mathbf{f}}{\partial \mathbf{t}}, \tag{83}$$

where on the right hand side, the first term reflects the change in distribution function due to the total field force $\mathbf{F_t} = \mathbf{E} + \mathbf{v} \times \mathbf{B}$, the second term is the change due to concentration gradients, and the last term is the local change in the distribution function. Since the total number of carriers in the crystal is constant, the total rate of change of the distribution is identically zero by Liouville's theorem. Hence the *local* change in the distribution function is written as

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial t}|_{coll} - \frac{\mathbf{F_t}}{\hbar} \cdot \nabla_{\mathbf{k}} f(\mathbf{k}) - \mathbf{v} \cdot \nabla_{\mathbf{r}} \mathbf{f}(\mathbf{k}) + \frac{\partial \mathbf{f}}{\partial \mathbf{t}}, \tag{84}$$

where the first term has been split off from the field term since collision effects are not easily described by fields. The second term is due to applied field only and the third is due to concentration gradients.

Denoting the scattering rate from state $\mathbf{k} \to \mathbf{k}'$ as $S(\mathbf{k}, \mathbf{k}')$, the collision term is given by

$$\frac{\partial f(\mathbf{k})}{\partial t}|_{coll} = \sum_{\mathbf{k}'}[S(\mathbf{k}', \mathbf{k})f(\mathbf{k}')[1 - f(\mathbf{k})] - S(\mathbf{k}, \mathbf{k}')f(\mathbf{k})[1 - f(\mathbf{k}')]]. \tag{85}$$

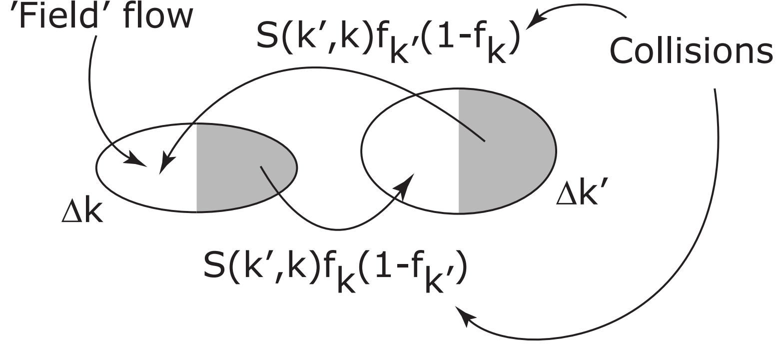

Figure 11: Scattering term of Boltzmann transport equation depicting the inflow and outflow of the distribution function.

Figure 11 provides a visual representation of the scattering processes that form the collision term. The increase of the distribution function in the small volume $\Delta\mathbf{k}$ by particles flowing in by the field term is balanced by the net flow out by the two collision terms.

At equilibrium ($f = f_0$), the 'principle of detailed balance' enforces the condition

$$S(\mathbf{k}', \mathbf{k})f_0(\mathbf{k}')[1 - f_0(\mathbf{k})] = S(\mathbf{k}, \mathbf{k}')f_0(\mathbf{k})[1 - f_0(\mathbf{k}')], \tag{86}$$

which translates to

$$S(\mathbf{k}', \mathbf{k})e^{\frac{\varepsilon_{\mathbf{k}}}{k_B T}} = S(\mathbf{k}, \mathbf{k}')e^{\frac{\varepsilon_{\mathbf{k}'}}{k_B T}}. \tag{87}$$

In the special case of *elastic* scattering, $\varepsilon_{\mathbf{k}} = \varepsilon_{\mathbf{k}'}$, and as a result, $S(\mathbf{k}', \mathbf{k}) = S(\mathbf{k}, \mathbf{k}')$ irrespective of the nature of the distribution function. Using this, one rewrites the collision term as

$$\frac{\partial f(\mathbf{k})}{\partial t}|_{coll} = \sum_{\mathbf{k}'} S(\mathbf{k}, \mathbf{k}')(f(\mathbf{k}') - \mathbf{f}(\mathbf{k})). \tag{88}$$

One can rewrite this collision equation as

$$\frac{df(\mathbf{k})}{dt} + \frac{f(\mathbf{k})}{\tau_q(\mathbf{k})} = \sum_{\mathbf{k}'} S(\mathbf{k}, \mathbf{k}')f(\mathbf{k}'), \tag{89}$$

where the *quantum scattering time* is defined as

$$\frac{1}{\tau_q(\mathbf{k})} = \sum_{\mathbf{k}'} S(\mathbf{k}, \mathbf{k}'). \tag{90}$$

A particle prepared in state $|\mathbf{k}\rangle$ at time $t = 0$ by an external perturbation will be scattered into other states $|\mathbf{k}'\rangle$ due to collisions, and the distribution function in that state will approach the equilibrium distribution exponentially fast with the time constant $\tau_q(\mathbf{k})$ upon the removal of the applied field. The quantum scattering time $\tau_q(\mathbf{k})$ may be viewed as a 'lifetime' of the particle in the state $|\mathbf{k}\rangle$.

Let us now assume that the external fields and gradients have been turned on for a long time. They have driven the distribution function to a *steady state* value f from f_0. The perturbation is assumed to be small, i.e., distribution function is assumed not to deviate far from its equilibrium value of f_0. Under this condition, it is common practice to assume that

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial t}|_{coll} = -\frac{f - f_0}{\tau}, \tag{91}$$

where τ is a time scale characterizing the relaxation of the distribution. This is the relaxation time approximation, which is crucial for getting a solution of the Boltzmann transport equation.

When the distribution function reaches a steady state, the Boltzmann transport equation may be written as

$$\frac{\partial f}{\partial t} = -\left(\frac{f - f_0}{\tau}\right) - \frac{\mathbf{F_t}}{\hbar} \cdot \nabla_{\mathbf{k}} f(\mathbf{k}) - \mathbf{v} \cdot \nabla_{\mathbf{r}} \mathbf{f}(\mathbf{k}) = \mathbf{0}, \tag{92}$$

where the relaxation time approximation to the collision term has been used. In the absence of any concentration gradients, the distribution function is given by

$$f(\mathbf{k}) = f_0(\mathbf{k}) - \tau \frac{\mathbf{F_t}}{\hbar} \cdot \nabla_{\mathbf{k}} f. \tag{93}$$

Using the definition of the velocity $\mathbf{v} = 1/\hbar(\partial \varepsilon_{\mathbf{k}}/\partial k)$, the distribution function becomes

$$f(\mathbf{k}) = f_0(\mathbf{k}) - \tau \mathbf{F_t} \cdot \mathbf{v} \frac{\partial f(\mathbf{k})}{\partial \varepsilon}, \tag{94}$$

and since the distribution function is assumed to be close to f_0, we can make the replacement $f(\mathbf{k}) \to f_0(\mathbf{k})$, whence the distribution function

$$f(\mathbf{k}) = f_0(\mathbf{k}) - \tau \mathbf{F_t} \cdot \mathbf{v} \frac{\partial f_0(\mathbf{k})}{\partial \varepsilon} \tag{95}$$

is the *solution* of BTE for a perturbing force $\mathbf{F_t}$.

6.1.1 Electric field

The external force $\mathbf{F_t}$ may be due to electric or magnetic fields. We first look for the solution in the presence of only the electric field; thus, $\mathbf{F_t} = -e\mathbf{E}$.

Using Equation 95, for elastic scattering processes one immediately obtains

$$f(\mathbf{k}') - f(\mathbf{k}) = \underbrace{e\tau \frac{\partial f_0}{\partial \varepsilon} \mathbf{E} \cdot \mathbf{v}}_{f(\mathbf{k}) - f_0(\mathbf{k})} (1 - \frac{\mathbf{E} \cdot \mathbf{v}'}{\mathbf{E} \cdot \mathbf{v}}) \tag{96}$$

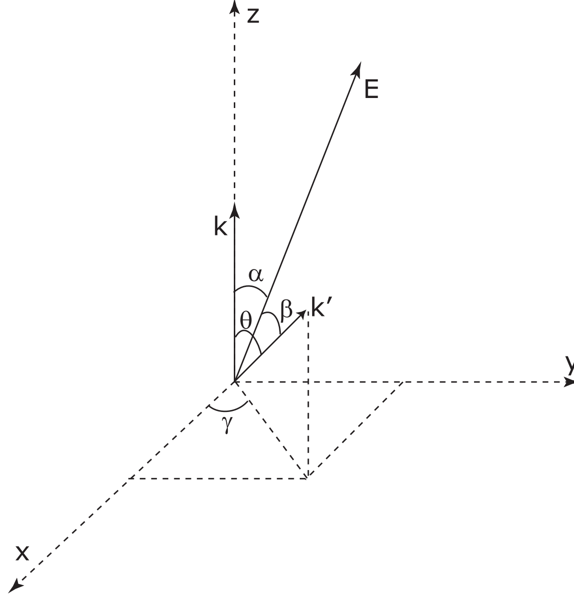

Figure 12: Angular relations between the vectors in Boltzmann transport equation.

for a parabolic bandstructure ($\mathbf{v} = \hbar\mathbf{k}/m^\star$). Using this relation, the collision term in the form of the relaxation time approximation becomes

$$\frac{\partial f(\mathbf{k})}{\partial t} = \sum_{\mathbf{k}'} S(\mathbf{k},\mathbf{k}')(f(\mathbf{k}') - f(\mathbf{k})) = -\frac{(f(\mathbf{k}) - f_0(\mathbf{k}))}{\tau_m(\mathbf{k})}, \tag{97}$$

where a new relaxation time is defined by

$$\frac{1}{\tau_m(\mathbf{k})} = \sum_{\mathbf{k}'} S(\mathbf{k},\mathbf{k}')(1 - \frac{\mathbf{E}\cdot\mathbf{k}'}{\mathbf{E}\cdot\mathbf{k}}). \tag{98}$$

This is the *momentum relaxation time.*

Let the vectors $\mathbf{k}, \mathbf{k}', \mathbf{E}$ be directed along random directions in the 3−dimensional space. We fix the $z-$axis along $\mathbf{k}$ and the $y-$axis so that $\mathbf{E}$ lies in the $y-z$ plane. From Figure 12, we get the relation

$$\frac{\mathbf{k}'\cdot\mathbf{E}}{\mathbf{k}\cdot\mathbf{E}} = \cos\theta + \sin\theta\sin\gamma\tan\alpha, \tag{99}$$

where the angles are shown in the figure.

When the sum over *all* $\mathbf{k}'$ is performed for the collision term, the $\sin(\gamma)$ sums to zero and the momentum relaxation time $\tau_m(\mathbf{k})$ becomes

$$\frac{1}{\tau_m(\mathbf{k})} = \sum_{\mathbf{k}'} S(\mathbf{k},\mathbf{k}')(1 - \cos\theta). \tag{100}$$

We note here that this relation can be generalized to an arbitrary number of dimensions, the three-dimensional case was used as a tool. This is the general form for momentum

scattering time, which is used heavily in the text for finding scattering rates determining mobility. It is related to mobility by the Drude relation $\mu = e\langle\tau(\mathbf{k})\rangle/m^\star$, where the momentum scattering time has been averaged over all energies of carriers.

The quantum scattering rate $1/\tau_q(\mathbf{k}) = \sum_{\mathbf{k}'} S(\mathbf{k},\mathbf{k}')$ and the momentum scattering rate $1/\tau_m(\mathbf{k}) = \sum_{\mathbf{k}'} S(\mathbf{k},\mathbf{k}')(1-\cos\theta)$ are both experimentally accessible quantities, and provide a valuable method to identify the nature of scattering mechanisms. The momentum scattering time $\tau_m(\mathbf{k})$ measures the average time spent by the particle moving along the external field. It differs from the quantum lifetime due to the $\cos\theta$ term. The angle θ is identified from Figure 12 as the angle between the initial and final wavevectors upon a scattering event. Thus for scattering processes that are isotropic $S(\mathbf{k},\mathbf{k}')$ has no angle dependence, the $\cos\theta$ term sums to zero, and $\tau_q = \tau_m$. However, for scattering processes that favor small angle ($\theta \to 0$) scattering, it is easily seen that $\tau_m > \tau_q$.

6.2 Mobility- basic theory

We will now arrive at a general expression for the drift mobility of carriers of arbitrary degeneracy confined in d spatial dimensions. d may be 1,2 or 3; for $d = 0$, the carrier in principle does not move in response to a field. Let the electric field be applied along the ith spatial dimension, ($\mathbf{E} = E_i\mathbf{i}$) and the magnetic field $\mathbf{B} = 0$. We assume an isotropic effective mass $m^\star$. Starting from the Boltzmann equation for the distribution function of carriers $f(k,r,t)$, and using the relaxation-time approximation solution, we write the distribution function as

$$f(\mathbf{k}) = f_0(\mathbf{k}) + eF_i\tau(k)v_i\frac{\partial f_0}{\partial\varepsilon}, \tag{101}$$

where $\tau(k)$ is the momentum relaxation time and v_i is the velocity of carriers in the ith direction in response to the field.

The total number of carriers per unit 'volume' in the $d-$dimensional space is

$$n = \int \frac{d^dk}{(2\pi)^d} f(k) = \int d\varepsilon f(\varepsilon) g_d(\varepsilon), \tag{102}$$

where the generalized $d-$dimensional DOS expressed in terms of the energy of carriers is given by

$$g_d(\varepsilon) = \frac{1}{2^{d-1}\pi^{\frac{d}{2}}\Gamma(\frac{d}{2})}(\frac{2m^\star}{\hbar^2})^{\frac{d}{2}}\varepsilon^{\frac{d}{2}-1}. \tag{103}$$

Here $\hbar$ is the reduced Planck's constant and $\Gamma(...)$ is the gamma function. Using this, and the parabolic dispersion we can switch between the k-space and energy-space.

The current in response to the electric field along the ith direction is given by

$$\mathbf{J} = 2e\int \frac{d^dk}{(2\pi)^d}\mathbf{v}f(\mathbf{k}). \tag{104}$$

Using the distribution function from the solution of the BTE, we see that the f_0 term integrates out to zero, and only the second term contributes to a current.

For a particle moving in $d-$dimensions the total kinetic energy ε is related to the average squared velocity $\langle v_i^2 \rangle$ along *one* direction by the expression $\langle v_i^2 \rangle = 2\varepsilon/dm^\star$. Using this result, we re-write the current as

$$J_i = en \underbrace{(-\frac{2e}{dm^\star}\frac{\int d\varepsilon \tau_m \varepsilon^{\frac{d}{2}}\frac{\partial f_0}{\partial \varepsilon}}{\int d\varepsilon f_0(\varepsilon)\varepsilon^{\frac{d}{2}-1}})}_{\mu_d} F_i, \tag{105}$$

where the mobility in the $d-$dimensional case is denoted by the underbrace. τ_m, the momentum relaxation time due to scattering events calculated in the Born approximation by Fermi's golden rule using the scattering potential, turns out to depend on the energy of the mobile carrier and the temperature. Let us assume that it is possible to split off the energy dependence of the relaxation time in the form

$$\tau_m = \tau_0 \left(\frac{\varepsilon}{k_B T}\right)^n, \tag{106}$$

where τ_0 does not depend upon the energy of the carriers. Using this, and the fact that $f_0(\varepsilon) \to 0$ as $\varepsilon \to \infty$ and $\varepsilon^m \to 0$ as $\varepsilon \to 0$, the expression for mobility can be converted by an integration by parts to

$$\mu_d = \frac{e\tau_0}{m^\star} \cdot \left(\frac{\Gamma(\frac{d}{2}+n+1)}{\Gamma(\frac{d}{2}+1)}\right) \cdot \left(\frac{F_{\frac{d}{2}+n-1}(\zeta)}{F_{\frac{d}{2}-1}(\zeta)}\right), \tag{107}$$

where $F_j(\zeta)$ are the traditional Fermi-Dirac integrals of the jth order defined as

$$F_j(\zeta) = \frac{1}{\Gamma(j+1)}\int_0^\infty dx \frac{x^j}{1+e^{x-\zeta}}. \tag{108}$$

Equation 107 may be viewed as a *generalized* formula for mobility of carriers in $d-$dimensional space.

Note that this is a *general* expression that holds true for an *arbitrary* degeneracy of carriers that are confined in *arbitrary* (d) dimensions. We now proceed to use this form of the expression for determining the mobility for two extreme cases. The strongly non-degenerate ('ND') case, where $\zeta \ll -1$ and the strongly degenerate ('D') case, where $\zeta \gg +1$.

For the non-degenerate case, the Fermi integrals can be shown to reduce to $F_j(\zeta) \approx e^\zeta$. This reduces the expression for mobility to the simple form

$$\mu_d^{ND} \approx \frac{e\tau_0}{m^\star}\left(\frac{\Gamma(\frac{d}{2}+n+1)}{\Gamma(\frac{d}{2}+1)}\right). \tag{109}$$

For the strongly degenerate case, we make another approximation of the Fermi-Dirac integral. We re-write it as

$$F_j(\zeta) = \frac{e^\zeta}{\Gamma(j+1)}[\int_0^\zeta dx \frac{x^j}{e^\zeta + e^x} + \int_\zeta^\infty dx \frac{x^j}{e^\zeta + e^x}]. \tag{110}$$

In the first integral, $e^x \ll e^\zeta$ and in the second integral, $e^x \gg e^\zeta$ since $\zeta \gg 1$. Using this and retaining only the *leading* power of ζ, the Fermi-Dirac integral can be approximated as

$F_j(\zeta) \approx \zeta^{j+1}/\Gamma(j+2)$. Further, $\zeta = \varepsilon_F/k_BT$ where ε_F is the Fermi-energy that is known for the degenerate case if one knows *only* the carrier density. So the expression for degenerate carrier mobility finally reduces to the simple form

$$\mu_d^D \approx \frac{e\tau_0}{m^\star}\left(\frac{\varepsilon_F}{k_BT}\right)^n. \tag{111}$$

The validity of the degenerate and non-degenerate limits rests on the accuracy of the approximations made to the Fermi-Dirac integrals. For strong degeneracy and non-degeneracy, the approximations for the three-dimensional case are shown with the exact Fermi-Dirac integrals in Figure 13.

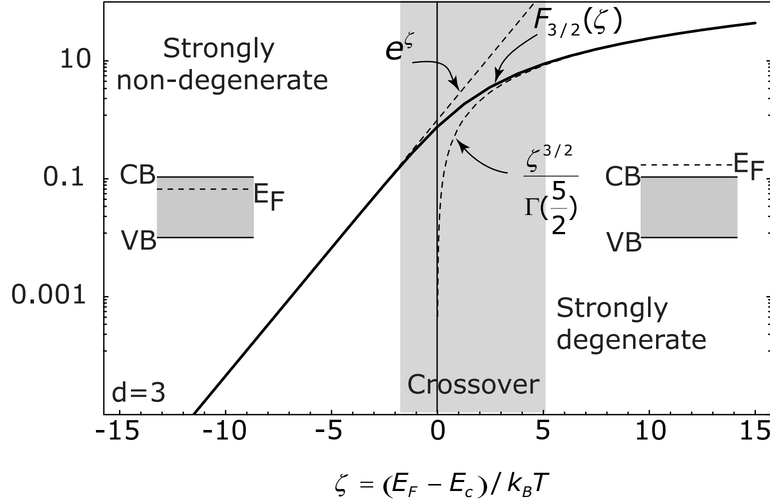

Figure 13: Accuracy of the approximations to the Fermi-Dirac integral in extreme degeneracy and extreme non-degeneracy.

6.3 Statistics for two- and three-dimensional carriers

The concentration of free carriers in the conduction band determines the location of the Fermi level. The carrier density for the $d-$dimensional case is given by

$$n = \int d\varepsilon g_d(\varepsilon) f(\varepsilon), \tag{112}$$

where $g_d(\varepsilon)$ is the $d-$dimensional density of states and $f(\varepsilon)$ is the distribution function. The distribution function is the solution of the Boltzmann transport equation. From the Boltzmann transport equation, the perturbation term in the distribution function has a $\partial f_0/\partial k$ term that is odd in k and integrates to zero. So the only term contributing to the carrier density is $f_0(\varepsilon)$, the equilibrium value of the distribution function given by the Fermi-Dirac function. This is saying nothing more than the fact that the carrier density does not change from the equilibrium value upon application of a field. Thus, the carrier density for

the $d-$dimensional case is evaluated using the generalized $d-$dimensional density of states to be

$$n = \frac{1}{2^{d-1}} \left(\frac{2m^\star k_B T}{\pi \hbar^2} \right)^{d/2} F_{\frac{d}{2}-1}(\zeta), \tag{113}$$

where $F_{d/2-1}(\zeta)$ is the Fermi-Dirac integral.

For the three-dimensional case, it reduces to the form

$$n_{3d} = \underbrace{2 \left(\frac{m^\star k_B T}{2\pi \hbar^2} \right)^{3/2}}_{N_c^{3d}} F_{1/2}(\zeta), \tag{114}$$

where N_c^{3d} is the 3-d band-edge density of states. The result holds true for arbitrary degeneracy. Sometimes, ζ is needed as a function of the carrier density and temperature; this is achieved by inverting the above expression by a numerical technique (the Joyce-Dixon approximation [4])

$$\zeta \simeq \ln(\frac{n}{N_c}) + \sum_{m=1}^{4} A_m (\frac{n}{N_c})^m, \tag{115}$$

where the constants $A_m = 3.536 \times 10^{-1}, -4.950 \times 10^{-3}, 1.484 \times 10^{-4}, -4.426 \times 10^{-6}$ for $m = 1, 2, 3, 4$ respectively. The Joyce-Dixon approximation holds good for the *entire* range of degeneracies that are achievable in semiconductors.

Similarly, for the two-dimensional case, we get immediately

$$n_{2d} = \underbrace{\frac{m^\star k_B T}{\pi \hbar^2}}_{N_c^{2d}} \ln(1 + e^\zeta), \tag{116}$$

which is a well known result for 2-d carrier density. For the 2-d case, $\zeta = (\varepsilon_F - \varepsilon_i)/k_B T$ where ε_i is the lowest subband energy.

6.4 Screening: Semiclassical Theory

In an insulator, atomic charges rearrange to screen external potentials. That is why we have the dielectric constant K, and the vacuum Coulomb potential $V(r) = \frac{e}{4\pi\epsilon_0 r}$ is screened to a value $V(r) = \frac{e}{4\pi\epsilon_0 K r}$. In a doped semiconductor, there are mobile carriers moving around and assist in screening further. The K gets changed. How? Let us see.

To understand the effect of screening, we start with an analogy. Consider a river with a rough bed. If there is a lot of water in the river (see Figure 14), a boat moving on the surface would be unaffected by the rough bed. The rough bed is 'screened' by water. However, if there is insufficient water, the rough bed shows up and impedes motion of the boat. Screening in semiconductors is very similar in behavior. Consider a band diagram shown in Figure 14. There is a impurity potential δV that causes a perturbation of the flat band diagram. This perturbation would be smoothed out by the flow of mobile carriers.

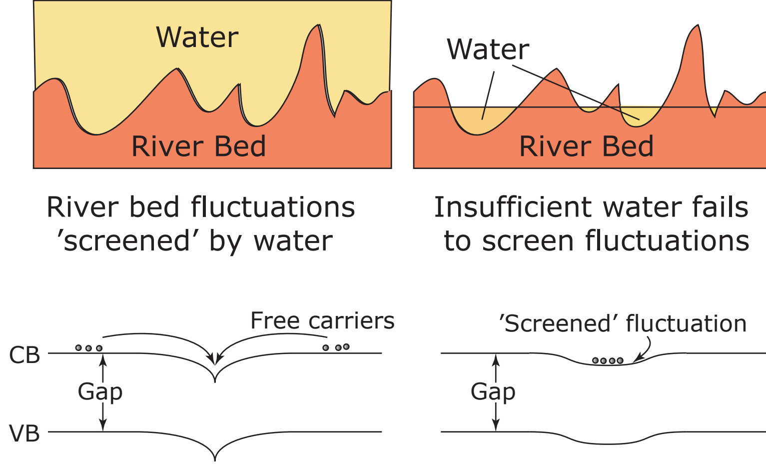

Figure 14: Figure to illustrate the phenomena of screening in semiconductors. The top figure gives an analogy of a river bed that is screened by water. The bottom figures show screening in action for a semiconductor.

We now treat the problem mathematically. The relation between the spatial carrier density variation and the conduction band edge is given by

$$n(\mathbf{r}) = N_C e^{-q\frac{E_C(\mathbf{r})-E_F}{k_B T}} \tag{117}$$

When there are no impurity potentials present, the conduction band edge is flat, $E_C(\mathbf{r}) = E_{C0}$. Then, the carrier density is uniform, and is given by $n_0 = N_C e^{-q\frac{E_{C0}-E_F}{k_B T}}$. Say the impurity potential causes a perturbation $E_C(\mathbf{r}) = E_{C0} + \delta V(\mathbf{r})$ in the conduction band edge. The carrier density then becomes

$$n(\mathbf{r}) = N_C e^{-q\frac{E_{C0}+\delta V(\mathbf{r})-E_F}{k_B T}} = n_0 e^{-q\frac{\delta V(\mathbf{r})}{k_B T}} \tag{118}$$

For *small* perturbations $\delta V(\mathbf{r}) \ll \frac{kT}{q}$, the exponential can be expanded using $e^x \simeq 1 + x$. Doing that, we get

$$n(\mathbf{r}) \simeq n_0 + \underbrace{\frac{qn_0}{k_B T}\delta V(\mathbf{r})}_{\delta n(\mathbf{r})} \tag{119}$$

For no extra charges in the semiconductor, the charge carriers are uniformly distributed and from charge neutrality,

$$\rho_u = q(p + N_D^+ - n - N_A^-) = 0 \tag{120}$$

holds. Let us consider just an $n-type$ semiconductor, where $p = N_A^- = 0$. The mobile charge changes to $n \to n_0 + \delta n(\mathbf{r}) = n(\mathbf{r})$ and the fixed ionized impurity occupations change to $N_D^+ \to N_{D0}^+ + \delta N_D^+(\mathbf{r}) = N_D^+(\mathbf{r})$. The $\delta(...)$ terms are in response to the perturbation

$\delta V(\mathbf{r})$, which is screening in action. The screened potential can be got by solving Poisson's equation that relates charges to potentials -

$$\nabla^2(E_{C0} + \delta V(\mathbf{r})) = -\frac{\rho(\mathbf{r})}{\epsilon(0)} = -\frac{q(N_D^+ + \delta N_D^+(\mathbf{r})) - (n_0 + \delta n(\mathbf{r}))}{\epsilon(0)} \tag{121}$$

Flatband conditions take out the constant terms in the equation, and we are left with only the $\delta(...)$ terms -

$$\nabla^2 \delta V(\mathbf{r}) = -\frac{q(\delta N_D^+(\mathbf{r}) - \delta n(\mathbf{r}))}{\epsilon(0)} \tag{122}$$

Using our approximations $\delta n(\mathbf{r}) \simeq \frac{qn_0}{k_B T}\delta V(\mathbf{r})$ and $\delta N_D^+(\mathbf{r}) \simeq \frac{-qN_{D0}^+}{k_B T}\delta V(\mathbf{r})$, we get

$$\nabla^2 \delta V(\mathbf{r}) = \frac{q^2 n^\star}{\epsilon(0) k_B T}\delta V(\mathbf{r}) = \frac{1}{\lambda_D^2}\delta V(\mathbf{r}) \tag{123}$$

Where we have defined $\lambda_D = \sqrt{\frac{\epsilon(0)k_B T}{q^2 n^\star}}$, the Debye screening length. Here, $n^\star$ is an effective carrier concentration that is effective in screening. It is generally not equal to the free mobile carrier density. For a n-type sample the effective electron screening concentration is given by

$$n^\star = n + \frac{n(N_D - n)}{N_D} \tag{124}$$

For completely ionized non-degenerate carriers, it is a good approximation to assume the screening electron density to be doping density, i.e., $n^\star = N_D$.

Now we have a equation that we can apply to solve for any fluctuating potential $\delta V(\mathbf{r})$. For the special case of a spherically symmetric fluctuation (example, Coulombic potential around a charged donor), Equation 57 becomes

$$\frac{d^2}{dr^2}(r\delta V(r)) = \frac{r}{\lambda_D^2}\delta V(r) \tag{125}$$

The solution for this equation for the Coulombic potential $V(r) = \frac{q^2}{4\pi\epsilon_0 K r}$ is the screened potential $V(r) = \frac{q^2}{4\pi\epsilon_0 K r}e^{-r/\lambda_D}$. The Coulombic potential is thus 'screened' by the Yukawa potential which reduces the potential very quickly within a few Debye lengths. We will use this result in evaluating ionized impurity scattering rates later.

For two-dimensional electron gases, it turns out that screening is much weaker. There is a surprising result that the screening length is a constant independent of the 2DEG density for all practical purposes. The 2DEG screening length is called the Thomas-Fermi length, and is given by

$$\lambda_{TF} = \frac{2\pi\epsilon_0 K \hbar^2}{q^2 m^\star} = \frac{a_B^\star}{2} \tag{126}$$

where $a_B^\star$ is the effective Bohr radius in the semiconductor.

From the form of FGR, we see that the scattering potential $U(r)$ always appears as a matrix element

$$U(q) = \langle k+q|U(r)|k\rangle = \underbrace{\int d^3 r e^{-iqr} U(r)}_{FT(U(r))} \tag{127}$$

where $\mathbf{q} = \mathbf{k'} - \mathbf{k}$. It is easy to see that this is nothing but the Fourier transform of the potential $U(r)$. Fourier transform is an extremely useful technique to solve such problems. To start off, let me state a very useful result[10] - the 3D Fourier transform of the bare Coulomb potential $U(r) = \frac{e^2}{4\pi\epsilon_0 r}$ is given by

$$U(q) = \frac{e^2}{\epsilon_0 q^2} \tag{128}$$

For a semiconductor without free carriers, this is scaled by the static dielectric constant K - $U(q) = \frac{e^2}{\epsilon_0 K q^2}$. Note that it diverges as $q \to 0$ - this is indicative of the long-range nature of the Coulomb potential. We derive the Fourier transform for a screened Coulomb potential $U_{sc}(r) = \frac{e^2}{4\pi\epsilon_0 K r} e^{-r/\lambda_D}$ in the following.

$$U_{sc}(q) = \int d^3 e^{i\mathbf{q}\cdot\mathbf{r}} U_{sc}(r) \tag{129}$$

Align $z - axis$ along $\mathbf{q}$.

$$U_{sc}(q) = \frac{e^2}{4\pi\epsilon^2 K} \int r^2 \sin(\theta) dr d\theta d\phi \frac{e^{-\frac{r}{\lambda_D}}}{r} e^{iqr\cos(\theta)} \tag{130}$$

Writing $1/\lambda_D = q_D$, and evaluating this integral gives us

$$U_{sc}(q) = \frac{e^2}{\epsilon_0 K (q^2 + q_D^2)} = \frac{U(q)}{\epsilon(q)} \tag{131}$$

where $\epsilon(q) = 1 + \frac{q_D^2}{q^2}$ is a scaling factor that the bare Fourier transform can be divided by to get the screened Fourier transform. $q_D = 1/\lambda_D$ is the Debye-Huckel wavevector. Note that screening has removed the divergence as $q \to 0$. Thus screening removes the long-range components of the Coulomb potential.

For a 2DEG, the Fourier Transform scaling factor after screening is given by

$$\epsilon(q) = 1 + \frac{q_{TF}}{q} \tag{132}$$

where $q_{TF} = 1/\lambda_{TF}$. Screening is weaker for a 2DEG than in 3D[11]. So the Fourier transform of any screened Coulombic potential will be the Fourier transform of the bare potential divided by this factor.

[10] pg 350, John Davies
[11] John Davies, pg 352

6.5 Screening by 2D/3D Carriers: Formal Theory

An important effect of the presence of mobile carriers in a semiconductor is screening. Since we are interested in scattering of mobile carriers from various defect potentials in the III-V nitrides, we summarize the theoretical tool used to attack the problem of screening in the presence of free carriers in the semiconductor.

The permittivity of vacuum is denoted as ϵ_0. If a material has no free carriers, an external d.c. electric field E will be scaled due to screening by movement of electron charge clouds of the atoms and the nuclei themselves - this yields the dielectric constant of the material, $\epsilon(0)$. The electric field inside the material is accordingly scaled down to $E/\epsilon(0)\epsilon_0$. If the electric field is oscillating in time, the screening by atomic polarization becomes weaker since the nuclei movements are sluggish, and in the limit of a very fast changing field, only the electron charge clouds contribute to screening, resulting in a reduced dielectric constant $\epsilon(\infty) < \epsilon(0)$. These two material constants are listed for the III-V nitrides in Table 1, and are related to the transverse and longitudinal modes of optical phonons by the Lyddane-Sachs-Teller equation $\epsilon(0)/\epsilon(\infty) = \omega_{LO}^2/\omega_{TO}^2$ [5].

The situation is more lively in the presence of mobile carriers in the conduction band [6]. In the situation where the perfect periodic potential of the crystal lattice is disturbed by a most general perturbing potential $V(r)e^{i\omega t}e^{-\Gamma t}$ (the potential may be due to a defect, impurity, or band variations due to phonons), *additional* screening of the potential is achieved by the flow of the mobile carriers. Lindhard first attacked this problem and with a random-phase approximation (RPA), arrived at a most-general form of the relative dielectric constant $\epsilon(q,\omega)$ given by [7]

$$\epsilon(q,\omega) = \epsilon(\infty) + (\epsilon(0) - \epsilon(\infty))\frac{\omega_{TO}^2}{\omega_{TO}^2 - \omega^2} + \epsilon(0)V_{uns}(q)\sum_{\mathbf{k}} \frac{f_{\mathbf{k}-\mathbf{q}} - f_{\mathbf{k}}}{\hbar\omega + i\Gamma + \varepsilon_{\mathbf{k}-\mathbf{q}} - \varepsilon_{\mathbf{k}}}. \quad (133)$$

Here, the first two terms take into account the contributions from the nuclei, the core electron clouds, and the valence electron clouds. The last term has a sum running over the free carriers only, and is zero for an intrinsic semiconductor. With this form of the dielectric function, the unscreened spatial part of the perturbation $V_{uns}(q)$ gets screened to $V_{scr} = V_{uns}(q)/\epsilon(q,\omega)$. Here $V_{uns}(q)$ is the Fourier-coefficient of the perturbing potential $V(q) = \int d^d r e^{iqr} V(r)$.

We are interested exclusively in *static* perturbations (defects in the material), and thus the time dependent part $\omega, \hbar\omega + i\Gamma \to 0$. With the approximations $f_{\mathbf{k}-\mathbf{q}} - f_{\mathbf{k}} \approx -\mathbf{q}\cdot\nabla_{\mathbf{k}} f_{\mathbf{k}}$ and $\varepsilon_{\mathbf{k}-\mathbf{q}} - \varepsilon_{\mathbf{k}} \approx -\hbar^2 \mathbf{q}\cdot\mathbf{k}/m^\star$, the dielectric function may be converted to [5]

$$\epsilon(q) = \epsilon_0(1 + V(q)\sum_k \frac{\partial f}{\partial \varepsilon}), \quad (134)$$

which is a very useful form that applies regardless of the dimensionality of the problem.

For 2-dimensional carriers, a Coulombic potential $V(r)$ which has the well-known Fourier transform $V_{2d}(q) = e^2/L^2\epsilon(0)\epsilon_0 q$, where L^2 is the 2DEG area and q is the 2DEG wavevector [3], the dielectric function may be written as

$$\epsilon_{2d}(q) = \epsilon(0)(1 + \frac{e^2}{q\epsilon(0)\epsilon_0}\frac{\partial(\sum_k f_k/L^2)}{\partial \varepsilon}) = \epsilon(0)(1 + \frac{q_{TF}}{q}). \quad (135)$$

Since the factor in brackets is the sheet density $\sum_k f_k/L^2 = n_s$, we get the 'Thomas-Fermi' screening wavevector q_{TF} given by

$$q_{TF} = \frac{m^\star e^2}{2\pi\epsilon(0)\epsilon_0\hbar^2} = \frac{2}{a_B^\star}, \tag{136}$$

$a_B^\star$ being the effective Bohr-radius in the semiconductor. Thus, the screening in a perfect 2DEG is surprisingly *independent* of the 2DEG density, and depends only on the basic material properties, within limits of the approximations made in reaching this result [8]. For quasi-2DEGs, where there is a finite extent of the wavefunction in the third dimension, the dielectric function acquires form-factors that depend on the nature of the wavefunction. Finally, the screened 2-d Coulomb potential is given by

$$V_{scr}(q) = \frac{e^2}{\epsilon_0\epsilon(0)(q + q_{TF})}. \tag{137}$$

Similarly, for the 3-d case, the Coulomb potential $V(q) = e^2/L^3\epsilon_0\epsilon(0)q^2$ leads to a dielectric function

$$\epsilon_{3d}(q) = \epsilon(0)(1 + \frac{q_D^2}{q^2}), \tag{138}$$

where Debye screening-wavevector q_D is given by

$$q_D = \sqrt{\frac{e^2 N_c F_{-1/2}(\zeta)}{\epsilon_0\epsilon(0)k_B T}}, \tag{139}$$

for an arbitrary degeneracy of carriers.

6.6 Mobility of two- and three-dimensional carriers

6.6.1 Two-dimensional carriers

The wavefunction of electrons for band-transport[12] in 2DEG is

$$\langle \mathbf{r}|\mathbf{k}\rangle = \frac{1}{\sqrt{A}} e^{i\mathbf{k}\cdot\mathbf{r}} \chi(z) u_{n\mathbf{k}}(\mathbf{r}), \tag{140}$$

where the wavefunction is decomposed into a plane-wave part in the 2-dimensional $x-y$ plane of area A and a finite extent in the $z-$direction governed by the wavefunction $\chi(z)$. $\mathbf{k}, \mathbf{r}$ are both two-dimensional vectors in the $x-y$ plane. $u_{n\mathbf{k}}(\mathbf{r})$ are the unit-cell-periodic Bloch-wavefunctions, which are generally not known exactly. The Kane model of bandstructure presents an analytical approximation for the Bloch-functions [9], which is not presented in anticipation of the cancellation of the Bloch-function for transport in parabolic bands.

Assuming that the defect potential is given by $V(\mathbf{r}, z)$, which depends on both the in-plane two-dimensional vector $\mathbf{r}$ and z perpendicular to the plane, time-dependent perturbation

[12] In the presence of heavy disorder, the wavefunctions are localized and transport occurs by hopping and activation. For such cases, we cannot assume plane-wave eigenfunctions for electrons. All samples studied here are sufficiently pure, localization effects are neglected.

theory provides the solution for the scattering rate of electrons in the 2DEG. Scattering rate from a state $|\mathbf{k}\rangle$ to a state $|\mathbf{k}'\rangle$ is evaluated using Fermi's Golden Rule [3]. The use of Fermi's Golden rule in the $\delta-$function form is justified since the typical duration of a collision in a semiconductor is much less than the time spent between collisions [7, 10]. The scattering rate is written as

$$S(\mathbf{k},\mathbf{k}') = \frac{2\pi}{\hbar}|H_{\mathbf{k},\mathbf{k}'}|^2\delta(\varepsilon_{\mathbf{k}} - \varepsilon_{\mathbf{k}'}), \tag{141}$$

where $H_{\mathbf{k},\mathbf{k}'} = \langle\mathbf{k}'|V(\mathbf{r},z)|\mathbf{k}\rangle \cdot I_{\mathbf{k},\mathbf{k}'}$ is the product of the matrix element $\langle\mathbf{k}'|V(\mathbf{r},z)|\mathbf{k}\rangle$ of the scattering potential $V(\mathbf{r},z)$ between states $|\mathbf{k}\rangle, |\mathbf{k}'\rangle$ and the matrix element $I_{\mathbf{k}\mathbf{k}'}$ between lattice-periodic Bloch functions. Owing to the wide bandgap of the III-V nitrides, the matrix element $I_{kk'} \approx 1$, the approximation holding good even if there is appreciable non-parabolicity in the dispersion [9].

By writing the scattering term in the form of Equation 141, we reach a point of connection to the Boltzmann-transport equation. Once the matrix element is determined, the momentum relaxation time $\tau_m(\mathbf{k})$ of the single particle state $|\mathbf{k}\rangle$ is evaluated from the solution of the Boltzmann-transport equation as

$$\frac{1}{\tau_m(\mathbf{k})} = N_{2D}\sum_{k'} S(k',k)(1-\cos\theta), \tag{142}$$

where N_{2D} is the total number of scatterers in the 2D area A and θ is the angle of scattering. Implicit in this formulation is the assumption that all scatterers act *independently* of each other, which is true if they are in a dilute concentration. If this does not hold (as in heavily disordered systems), then one has to take recourse to interference effects from multiple scattering centers by the route of Green's functions [11]. The impurity concentration in AlGaN/GaN 2DEGs is dilute due to good growth control - this is confirmed by the band transport characteristics.

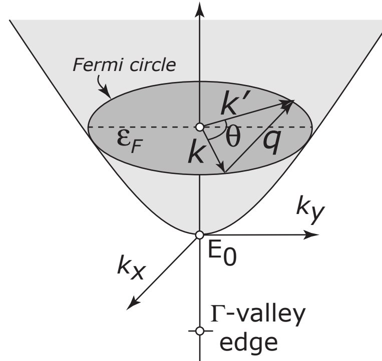

Figure 15: Visualization of the scattering process on the 2DEG Fermi-circle.

We write $\mathbf{q} = \mathbf{k} - \mathbf{k}'$ as depicted in Figure 15. Since states for the subband with $\varepsilon < \varepsilon_F$ are filled, they do not contribute to transport. Transport then occurs by scattering in the Fermi circle shown in the figure, and $|\mathbf{k}| = |\mathbf{k}'| \approx k_F$. From the figure, the magnitude of $\mathbf{q}$

is $q = 2k_F \sin(\theta/2)$ where θ is the angle of scattering. This makes $1 - \cos\theta = q^2/2k_F^2$. As a result, all integrals in the vector $\mathbf{q}$ reduce to integrals over angle θ.

Any *measurement* of transport properties samples over all state $|\mathbf{k}\rangle$ values. Converting the summation to an integral over the quasi-continuous wavevector states and exploiting the degenerate nature of the carriers for averaging $\tau_m(\mathbf{k})$, the measurable momentum scattering rate $\langle 1/\tau_m \rangle$ reduces to the simple form [3]

$$\frac{1}{\langle \tau_m \rangle} = n_{2D}^{imp} \frac{m^*}{2\pi\hbar^3 k_F^3} \int_0^{2k_F} |V(q)|^2 \frac{q^2}{\sqrt{1 - (\frac{q}{2k_F})^2}}, \tag{143}$$

where $n_{2D}^{imp} = N_{2D}/A$ is the areal density of scatterers and $k_F = \sqrt{2\pi n_s}$ is the Fermi wavevector, n_s being the 2DEG density.

The perturbation potential matrix element is given by

$$V_{nm}(q) = \frac{1}{A} \int dz \left(\chi_n^\star(z) \chi_m(z) \int d^2\mathbf{r} V(\mathbf{r}, z) e^{i\mathbf{q}\cdot\mathbf{r}} \right), \tag{144}$$

where n, m are the subband indices. This reduces to

$$V(q) = V_{00}(q) = \frac{1}{A} F(q) V(q, z_0) \tag{145}$$

when only the lowest subband ($n = m = 0$) is occupied. Here, $F(q)$ is a form factor that is unity when the 2DEG spread in the $z-$direction is a delta function. The scattering potential

$$V(q, z_0) = \frac{V_{uns}(q, z_0)}{\epsilon_{2d}(q)} \tag{146}$$

is the screened two-dimensional Fourier transform of the scattering potential of a scatterer located at z_0 for a perfect 2DEG (no $z-$spread), where the screened-dielectric function (Equation 135) was used.

For accurate evaluation of transport properties and scattering rates, the finite extent of the 2DEG along the z direction must be accounted for. The exact form of the wavefunction from the self-consistent Schrödinger - Poisson solution is very useful in determining the 2DEG sheet density and the shape of the wavefunction.

However, for analytic evaluation of scattering rates, the Fang-Howard variational wavefunction is a better candidate, and has been used successfully for transport calculations in the past [3]. The form of the wavefunction is

$$\chi(z) = 0, z < 0$$
$$\chi(z) = \sqrt{\frac{b^3}{2}} z e^{-\frac{bz}{2}}, z \geq 0, \tag{147}$$

where b is a variational parameter. The parameter is chosen such that it minimizes the energy; this is achieved when $b = \left(33m^* e^2 n_s / 8\hbar^2 \epsilon_0 \epsilon_b\right)^{1/3}$, where n_s is the 2DEG density. The centroid of the distribution (which is also a measure of the spread) is $\langle z \rangle = \int_0^\infty z|\chi(z)|^2 = 3/b$.

Thus, if the 2DEG $z-$dependence in not important in a scattering mechanism, the Fang-Howard function reduces to an ideal sheet charge $\rho(z) \rightarrow e n_s \delta(z)$ when $b \rightarrow \infty$.

The Fang-Howard variational wavefunction leads to a form factor

$$F(q) = \eta^3 = (\frac{b}{b+q})^3. \tag{148}$$

Screening by free carriers in the 2DEG is also affected due to the finite extent. This is reflected in another form factor $G(q)$ entering the 2D dielectric function

$$\epsilon_{2d}(q) = \epsilon(0)(1 + \frac{q_{TF}}{q} G(q)). \tag{149}$$

The screening form factor $G(q)$ is given by

$$G(q) = \frac{1}{8}(2\eta^3 + 3\eta^2 + 3\eta), \tag{150}$$

and q_{TF} is the Thomas-Fermi screening wavevector [8]. For a perfect 2DEG with no $z-$spread, $\eta \rightarrow 1$, and both form factors $F(q), G(q)$ reduce to unity. Thus, the Fang-Howard approximation along with particular scattering potentials can be used for evaluating the scattering rates of electrons in AlGaN/GaN 2DEGs.

6.6.2 Three-dimensional carriers

The generalized formula for mobility (Equation 107) applied to three-dimensional carriers (d=3) reduces to

$$\mu_{3d} = \frac{e\tau_0}{m^\star} \frac{\Gamma(n+\frac{5}{2})}{\Gamma(\frac{5}{2})} \frac{F_{n+\frac{1}{2}}(\zeta)}{F_{\frac{1}{2}}(\zeta)} = \frac{2e\tau_0}{3m^\star n_{3d}} \Gamma(n+\frac{5}{2}) F_{n+\frac{1}{2}}(\zeta), \tag{151}$$

where $n_{3d} = N_c^{3d} F_{1/2}(\zeta)$.

Once τ_0 is known for a scattering mechanism, ζ can be used from the Joyce-Dixon approximation (Equation 115) to calculate the mobility. Momentum scattering rates for various types of scattering mechanisms are well known ([1, 5, 2]). Mobility for some scattering mechanisms is now found used these scattering rates.

- *Deformation potential Scattering*

  The deformation potential acoustic phonon momentum relaxation time is given by

$$\tau_{ac} = \frac{2\pi\hbar^4 \rho v_s^2}{(2m^\star)^{\frac{3}{2}} a_C^2} \frac{\varepsilon^{-1/2}}{k_B T}, \tag{152}$$

  where the exponent $n = -1/2$, a_C is the conduction-band edge deformation potential, and ρ, v_s are the mass density and the sound velocity in the material respectively. Using this, and the fact that $F_0(\zeta) = \ln(1 + e^\zeta)$, acoustic-phonon scattering limited drift mobility becomes

$$\mu_{ac} = \frac{2e\hbar\rho v_s^2}{3\pi n_{3d} m^\star a_C^2} \ln(1 + e^\zeta), \tag{153}$$

where $\zeta = \varepsilon_F/k_BT$. This is a convenient closed-form expression for evaluation of mobility. The expression holds for an arbitrary degeneracy at arbitrary temperatures, which makes it much more useful than the traditionally used form of strongly non-degenerate limit. It is fortuitous that this form of mobility could be derived - it holds for all scattering mechanisms whose scattering rates have an energy exponent of $n = -1/2$.

The traditional expression for deformation-potential acoustic phonon scattering limited mobility in semiconductors in the strongly non-degenerate limit ($\zeta \ll -1$) is

$$\mu_{ac}^{ND} = \frac{2\sqrt{2\pi} e\hbar^4 \rho v_s^2}{3(m^\star)^{5/2} a_C^2 (k_BT)^{3/2}}, \tag{154}$$

with a $T^{-3/2}$ dependence. On the other extreme, for a strongly degenerate population of carriers, the mobility is

$$\mu_{ac}^{D} = \frac{(\pi/3)^{1/3} e\hbar^3 \rho v_s^2}{(m^\star)^2 a_C^2 n_{3d}^{1/3} (k_BT)}, \tag{155}$$

which is different from the non-degenerate case in the temperature dependence, and a carrier concentration dependence. Care needs to be exercised in using the expression for the strongly degenerate regime - if the acoustic phonon energy $\hbar v_s k$ is larger than the thermal energy k_BT which becomes very small at low temperatures, then the result does not hold since the phonon number cannot be approximated by $k_BT/\hbar v_s k$.

- *Piezoelectric phonon Scattering*

For piezoelectric phonon scattering, the relaxation rate is given by

$$\tau_{pz} = \frac{2\sqrt{2}\pi\hbar^2 \rho v_s^2}{(m^\star)^{1/2} (eh_{pz}/\epsilon)^2} \frac{\varepsilon^{+1/2}}{k_BT}, \tag{156}$$

where eh_{pz}/ϵ is the piezoelectric coefficient. Generalized mobility expression yields

$$\mu_{pz} = \frac{16 e\rho v_s^2 (k_BT)}{3\pi n\hbar (eh_{pz}/\epsilon)^2} \times F_1(\zeta), \tag{157}$$

which is in a form that may be evaluated numerically.

For the strongly non-degenerate case, we get the mobility as

$$\mu_{pz}^{ND} = \frac{16\sqrt{2\pi} e\hbar^2 \rho v_s^2}{3(eh_{pz}/\epsilon)^2 (m^\star)^{3/2} (k_BT)^{1/2}}, \tag{158}$$

which is the classical value quoted in texts [5] with the $T^{-1/2}$ dependence. In the strongly degenerate regime, we get

$$\mu_{pz}^{D} = \frac{2(3\pi^5)^{\frac{1}{3}} e\hbar^3 \rho v_s^2 n_{3d}^{1/3}}{(eh_{pz}/\epsilon)^2 (m^\star)^2 (k_B T)}. \tag{159}$$

- *Ionized Impurity Scattering*

Ionized impurity scattering time is given by [5]

$$\tau_{imp} = \frac{16\pi\epsilon^2 (2m^\star)^{1/2}}{e^4 N_I K_0} \times \varepsilon^{3/2}, \tag{160}$$

where N_I is the ionized impurity density and K_0 is a weak function of energy, and may be assumed constant in the calculations. Mobility is given by

$$\mu_{imp} = \frac{128 m^\star \epsilon^2 (k_B T)^3}{\hbar^3 e^3 N_I K_0 n} \times F_2(\zeta). \tag{161}$$

In the non-degenerate case, we recover the traditional Brooks- Herring formula

$$\mu_{imp}^{ND} = \frac{128\sqrt{2\pi}\epsilon^2}{e^3 N_I K_0 (m^\star)^{1/2}} \times (k_B T)^{3/2}, \tag{162}$$

with the $T^{3/2}$ dependence. For strongly degenerate carriers, we get

$$\mu_{imp}^{D} = \frac{24\pi^3 \epsilon^2 \hbar^3 n_{3d}}{e^3 N_I K_0 (m^\star)^2}, \tag{163}$$

which is weakly temperature dependent (through K_0), but now dependent on the 3d carrier density.

- *Alloy Scattering*

For alloy scattering, the momentum relaxation rate is given by [5]

$$\tau_{alloy} = \frac{2\pi\hbar}{V_0^2 \Omega[x] x(1-x)(2m^\star/\hbar^2)^{3/2}} \times \varepsilon^{-1/2}, \tag{164}$$

where V_0 is the alloy-scattering potential, x is the alloy composition, and $\Omega[x]$ is the volume range of the alloy-scattering potential. This is again with an exponent of $n = -1/2$. Using the techniques developed till now, we get the formally exact alloy scattering limited mobility to be

$$\mu_{alloy} = \frac{2e\hbar}{3\pi m^\star V_0^2 \Omega x(1-x)} \frac{k_B T}{n_{3d}} \ln(1 + e^{\zeta}). \tag{165}$$

We get the classical non-degenerate limit of mobility to be

$$\mu_{alloy}^{ND} = \frac{2\sqrt{2}e\hbar^4}{3(m^\star)^{5/2}V_0^2\Omega x(1-x)(k_BT)^{1/2}}, \tag{166}$$

and the degenerate limit as

$$\mu_{alloy}^{D} = \frac{(\pi/3)^{1/3}e\hbar^3}{(m^\star)^2V_0^2\Omega x(1-x)n_{3d}^{1/3}}, \tag{167}$$

where the degenerate nature of scattering removes the temperature dependence and adds a weak carrier density dependence.

- *Polar optical phonon scattering*

Polar optical phonon scattering limited mobility is given by

$$\mu_{POP} = \frac{4\pi\epsilon_0\epsilon^\star\hbar}{em^\star N_{bose}(T)q_0}, \tag{168}$$

where $1/\epsilon^\star = 1/\epsilon^\infty - 1/\epsilon^0$, the phonon-wavevector is given by $\hbar\omega_{op} = \hbar^2 q_0^2/2m^\star$, and $N_B(T) = 1/(\exp[\hbar\omega_{op}/k_BT] - 1)$ is the Bose-Einstein distribution function.

6.7 Material properties of III-V nitrides relevant to transport

For the calculation of transport properties of the III-V nitride semiconductors, it is essential to know various properties of the material. The bandstructure is of utmost importance, determining the effective mass of carriers and the allowed energy and momentum eigenvalues for the carriers. The band-alignment is essential for analyzing transport of quantum-confined carriers at heterojunctions. The electro-mechanical properties such as deformation potentials of the bands and acoustic wave (sound) velocities are essential in the electron-acoustic phonon coupling study. Optical deformation potential and optical phonon energies are essential in calculating optical phonon scattering rates. In what follows, these properties are surveyed for the III-V nitride semiconductors with an eye for application for transport studies.

Bandstructure and alignment

Since we are interested in analyzing transport in GaN, a close look at the bandstructure is in order. In Figure 2, the Brillouin zone, the theoretically calculated Bandstructure [12] and the simplified bandstructure [13] with only the lowest three conduction band valleys are shown.

Carriers residing in the lowest conduction-band valley (Γ-valley, direct gap) have an effective mass of $m^\star = 0.2m_0$. The next band minimum is in the $M-L$ direction of the Brillouin zone at 2.1 eV from the Γ point. The effective mass of carriers in this valley is $m_{ML}^\star = 0.4m_0$. From the Brillouin-zone picture, there are six such equivalent minima. The next minimum is close to the $M-L$ minimum in energy - it is at the A-point in the Brillouin-zone. The effective mass in this valley is $m_A^\star = 0.6m_0$. We note here that the effective masses and energy separations of the $M-L$ and A valleys are from theoretical bandstructure calculations [13] whereas that of the Γ valley has been verified experimentally ([14, 15]).

Table 1: Material properties of GaN for transport calculations [17]

Property	Symbol	GaN	AlN	InN	Units
Effective mass (Γ valley)	$m^\star$	0.2	0.5	0.1	m_0
Mass density	ρ	6.15	3.23	6.81	g/cm^{-3}
Static dielectric constant	$\epsilon(0)$	8.9	8.5	15.3	-
High frequency dielectric constant	$\epsilon(\infty)$	5.35	4.77	8.4	-
Optical phonon energy	$\hbar\omega_{op}$	92	100	89	meV
Deformation potential	Ξ	8.3	9.5	7.1	meV
Sound velocity (Longitudinal)	v_s	8	11	5.2	10^5 cm/s

The large separation of the direct-gap Γ valley minima from the other minima at $A, L-M$ ($\Delta\varepsilon \geq 1$ eV) bodes well for low-field transport analysis; a parabolic dispersion $\varepsilon = \hbar^2 k^2/2m^\star$ is a good approximation. The effective mass is $m^\star = 0.2m_0$, and has been reported to be isotropic [16], which simplifies the transport analysis. Figure 16 depicts the band discontinuities between the three members of the nitride family - InN, GaN, and AlN.

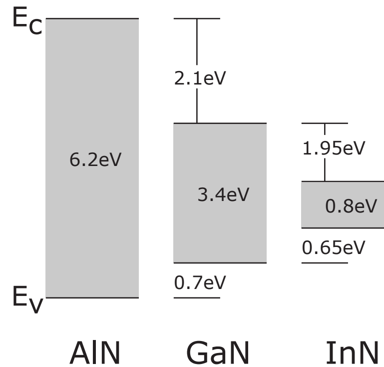

Figure 16: The band alignment in the III-V nitrides. Note that the fundamental bandgap of InN was previously believed to be 1.9eV - measured on heavily doped polycrystalline films. However, developments in growth techniques only recently made it possible to grow pure single crystals which lead to the new bandgap. The band alignment of InN is assumed to follow the same 3:1 ratio as in the case of GaN and AlN; it has not been experimentally measured yet.

The phonon dispersion curves of GaN have been calculated theoretically and measured experimentally as well. The experimental determination of phonon dispersion curves is traditionally done by the technique of neutron scattering. However, this technique requires unstrained samples of large sizes, which has not been possible till date. Bulk unstrained GaN samples of GaN can be grown by high pressure techniques, however, these samples are small in size (5×5mm). Ruf et. al. [18] overcame the size limitation by using the

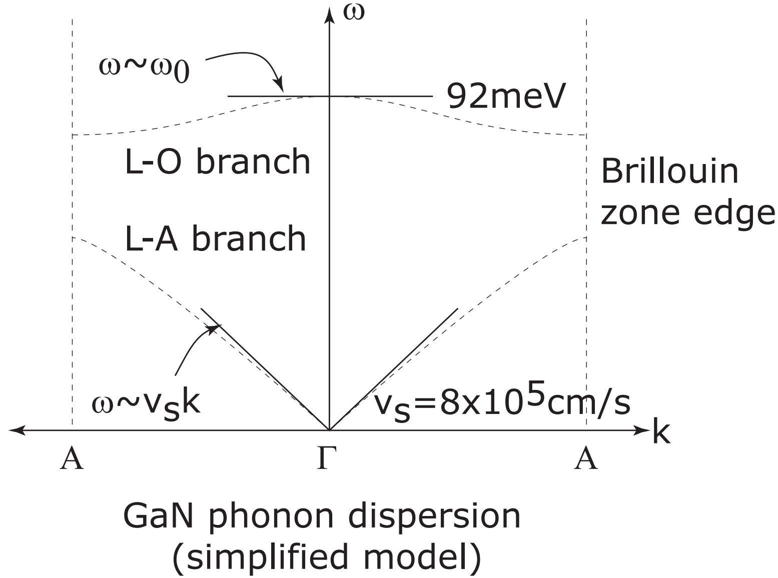

Figure 17: Phonon dispersion of GaN with the two parameters needed for transport studies.

similar technique of inelastic X-Ray scattering for determining the phonon-dispersion of the material. They found very close agreement between the measured phonon-dispersion and the calculated dispersions by ab-initio techniques.

Since our interest in this work is aimed towards determining the electron-phonon coupling for transport calculations, we use the simplified phonon-dispersion shown in Figure 17. The presence of both longitudinal and transverse modes of phonon propagation complicates the usage - we will assume only the longitudinal components since the effect of the transverse modes is much weaker.

The longitudinal sound velocities are given by $v_s = \sqrt{c_{11}/\rho}$ and $v_s = \sqrt{c_3/\rho}$ for the [1000] and [0001] directions of the wurtzite lattice. Since $c_{11} \sim c_{33}$ for GaN, AlN, and InN, it is a good approximation to assume that sound velocity is isotropic along the two directions; the sound velocities are tabulated in Table 1. Since acoustic phonons couple to electrons in the semiconductor through the deformation potential, the deformation potentials of the lowest conduction band valley are also listed in Table 1.

7 Current research, future directions

I list some topics of intense current interest. These are related to topics on transport theory covered in the class and you should be able to read papers related to these topics -

- Transport of electrons in 2DEGs - The integral and fractional quantum Hall effects. It has already generated four Physics Nobel prizes.(Von Klitzing, Stormer, Tsui, Laughlin).

- Transport in Disordered electronic systems - Metal-Insulator transistions in bulk and 2DEGs, Localization, etc. Has garnered at least 2 Physics Nobel prizes.(Mott, Anderson).
- Ballistic transport through quantum point contacts, quantized conductance.
- Transport in Quantum Wires, Dots.
- Transport in Nanotubes and Nanowires.
- Spin transport and spin scattering in dilute magnetic semiconductors.

References

[1] K. Seeger, *Semiconductor Physics, An Introduction*. Berlin: Springer Verlag, 6th ed., 1999.

[2] C. M. Wolfe, N. Holonyak Jr., and G. E. Stillman, *Physical Properties of Semiconductors*. Englewood Cliffs, New Jersey: Prentice Hall, 1st ed., 1989.

[3] J. H. Davies, *The Physics of Low-Dimensional Semiconductors*. Cambridge, United Kingdom: Cambridge University Press, 1st ed., 1998.

[4] W. B. Joyce and R. W. Dixon *Appl. Phys. Lett.*, vol. 31, p. 354, 1977.

[5] C. Hamaguchi *Basic Semiconductor Physics*, p. 280, 2001.

[6] J. Ziman *Theory of Solids*, Cambridge University Press, 1972.

[7] D. K. Ferry, *Semiconductor Transport*. London: Taylor & Francis, 1st ed., 2000.

[8] T. Ando, A. B. Fowler, and F. Stern *Rev. Mod. Phy.*, vol. 54, p. 437, 1982.

[9] B. M. Askerov, *Electron Transport Phenomena in Semiconductors*. Singapore: World Scientific, 1st ed., 1994.

[10] H. Kroemer, *Quantum Mechanics for Engineering, Materials Science, and Applied Physics*. Englewoods Cliff, New Jersey: Prentice Hall, 1st ed., 1994.

[11] G. D. Mahan, *Many Particle Physics*. New York: Kluwer Academic/Plenum Publishers, 3rd ed., 2000.

[12] S. Bloom, G. Harbeke, E. Meier, and I. B. Ortenburger *Phys. Stat. Solidi*, vol. 66, p. 161, 1974.

[13] U. Bhapkar and M. S. Shur *J. Appl. Phys.*, vol. 82, p. 1649, 1997.

[14] A. F. Brana, C. Diaz-Paniagua, F. Batallan, J. A. Garrido, E. Munoz, and F. Omnes *J. Appl. Phys.*, vol. 88, p. 932, 2000.

[15] W. Knap, S. Contreras, H. Alause, C. Skierbiszewski, J. Camassel, M. Dyakonov, J. L. Robert, J. Yang, Q. Chen, M. A. Khan, M. L. Sadowski, S. Huant, F. H. Yang, M. Goian, J. Leotin, and M. S. Shur *Appl. Phys. Lett.*, vol. 70, p. 2123, 1997.

[16] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan *J. Appl. Phys.*, vol. 89, p. 8815, 2001.

[17] E. archive New Semiconductor Materials Characteristics and Properties *http://www.ioffe.rssi.ru/SVA/NSM/.*

[18] T. Ruf, J. Serrano, M. Cardona, P. Pavone, M. Pabst, M. Krisch, M. D'Astuto, T. Suski, I. Grzegory, and M. Leszczynski *Phys. Rev. Lett.*, vol. 86, p. 906, 2001.